SEcure Cloud computing for CRitical Infrastructure IT

Contract No 312758

Deliverable D2.8
Final Ethics Report

AIT Austrian Institute of Technology • ETRA Investigación y Desarrollo • Fraunhofer Institute for Experimental Software Engineering IESE • Karlsruhe Institute of Technology • NEC Europe • Lancaster University • Mirasys • Hellenic Telecommunications Organization OTE• Ayuntamiento de Valencia • Amaris

Document control information		
Title	Final Ethics Report	
Creator	KIT legal	
Editor	Frank Pallas, Manuela Wagner, Silvia Balaban	
Description	This document identifies the main ethical problems and derives respective safeguards for the SECCRIT project	
Classification	☐ **Red** – Highly sensible Information, limited access for: ☐ **Yellow** – restricted limited access for: ☐ **Green** – restricted to consortium members ☒ White – public	
Reviewers	☐ AIT ☐ ETRA ☐ IESE ☒ KIT ☐ NEC	☐ ULANC ☐ MIRASYS ☒ OTE ☐ VLC ☐ AMARIS
Review status	☐ Draft ☐ WP Manager accepted ☒ Co-ordinator accepted	
Action requested	☐ to be revised by Partners involved in the preparation of the Project Deliverable ☐ to be reviewed by applicable SECCRIT Partners ☐ for approval of the WP Manager ☐ for approval of the Project Co-ordinator	
Requested deadline	31.12.2015	

Versions			
Version	**Date**	**Change**	**Comment/Editor**
1	26/06/2015	Initial version based on D2.5	Frank Pallas / Silvia Balaban
2	05/10/2015	Intermediate version with ch 5.1 and 5.2	Frank Pallas / Silvia Balaban
3	07/11/2015	First complete draft	Frank Pallas / Manuela Wagner / Silvia Balaban
4	07/12/2015	Full document review	Ioannis Chochlicouros / OTE
5	15/12/2015	Full document review	Santiago Caceres / ETRA

Abstract

The concept of cloud computing also affects the ethical domain. Whereas the law is setting regulations, framed by a governing power and which have to be accepted by everyone, ethics are the reflection on morals, forming the basis for human interrelations. Many concerns and doubts about cloud computing have been raised, particularly concerning the fact that unwanted consequences like, for example, the transmission to non-authorized third parties or even a more general "loss of control" may arise from this new paradigm of computing. Therefore, it not only becomes legally important to estimate the effects associated with the use of cloud computing but also ethically relevant to identify potential risks. This is the reason why D2.8 – as an extension of the initial ethics report D2.5 – focuses on ethical aspects that go beyond those covered in the legal deliverables D2.2 and D2.7. This was done under consideration of deliverable D2.4 on the results of National Data Protection consultation. In this vein, this document outlines the rather abstract ethical challenges of the project's activities, documents what measures were taken from the very beginning in order to adequately address them throughout the project (and, particularly, in the establishment of demonstrators), and relates identified ethical challenges to the technical and non-technical results of the project.
This document is an updated version of the initial ethics report (D2.5) submitted in M12. It contains the contents of this initial report and amends them in two respects: First, it lays out ethical challenges of cloud computing that were not foreseen in D2.5 but that emerged since the finalization of the initial ethics report either from the project activities or in international scientific discussions. Second, it describes how the challenges originally identified in D2.5 as well as those that emerged later were addressed within project SECCRIT, thus providing a comprehensive documentation on the overall approach taken within the project in matters of ethical aspects beyond legal questions (which are separately addressed in own deliverables – D2.2 and D2.7). In order to achieve logical consistency and to provide the possibility of a quick catch-up for those readers already familiar with the initial ethics report, the main chapters of D2.5 were adopted without any changes and the additional contents mentioned above are added in separate chapters.

Table of Contents

1 Motivation, Scope, Methodology and Structure

Throughout the whole runtime, it has been an integral aspect of the SECCRIT project that relevant ethical issues are identified and integrated from the very beginning and appropriately taken into account from the ground up during the design and implementation of the SECCRIT's technologies and methods. A sound analysis of the most relevant ethical issues beyond core legal aspects and their reflection within technical concepts was therefore conducted throughout the first year of SECCRIT's runtime and reported on in D2.5. By means of these early considerations, the project consortium significantly reduced the risk of developed technologies not being in line with fundamental ethical values, thus avoids essential ethical conflicts from arising during later project phases and, finally, reduces the risk of developed technologies not being applicable in practice.

The ethical safeguards initially developed for the project had to be applied during the design and development of technologies as well as during the establishment of demonstrators, calling for a concluding review and assessment. The document at hand therefore is an update of the initial ethics-related considerations from D 2.5 and extends these with delineations on how ethical aspects were addressed within the project.

In order to achieve logical consistency and to provide the possibility of a quick catch-up for those readers already familiar with the initial ethics report, the main sections of D 2.5 were adopted herein without any changes and the additional contents mentioned above are added in separate sections. In particular, sections 0, 3 and 4 and large parts of the remainder of this introduction are taken from the initial ethics report and mainly left unchanged except for minor editorial changes in matters of grammatical tense etc. In addition to these, section 5 then delineates how the safeguards and other ethics-related measures were practically applied in the project and retrospectively assesses the adopted ethics-related activities. Altogether, this document thus provides a comprehensive documentation on the overall approach taken within the project in matters of ethical aspects beyond legal questions.

Particularly with regard to cutting-edge information technology, data protection considerations usually feature prominently in ethical deliberations. These are, however, well-reflected in the respective legal data protection framework. This legal framework, in turn, is extensively covered in deliverables D 2.2 and D 2.7. Ethical considerations on issues of data protection in the strict sense as covered by data protection legislation will therefore not be addressed in the following parts of the present deliverable.

However, ethics also cover more general aspects of moral, giving a basis for human living. Ethics as a critical reflection on how to behave are – in contrast to the law – not explicitly codified. Moral and ethics derive from the Latin word *mores* and the Greek word *ethos*, which signify the beliefs and customs guiding the interaction and conduct amongst human beings. Even if every legislation is (or, at least, should be) based on ethical considerations, the law covers in most cases only a small part of ethical problems. It might therefore very well happen that obeying the law does not automatically imply that the acting is also ethically acceptable. Laws for instance can also change in line with changed governmental preferences and views, while the respective morals and customs remain unchanged. On the other hand, novel technologies might not always immediately be incorporated within legal frameworks while ethical considerations already provide a direction for socially acceptable conduct. Finally, legal provisions are enforced through governmental sanctions in case of noncompliance, while ethical misconduct does not always lead to explicit punishment (aside from societal disrespect).

In the context of SECCRIT, there are numerous ethical questions which have to be taken into account and which go well beyond data protection in the strict sense. In particular, these arise from the employed use-cases of video surveillance of critical infrastructures and urban mobility services. Especially with regard to video surveillance, there is already a well-understood body of knowledge on its ethical dimension which will therefore be reflected in the following sections.

Much less broad is the body of existing knowledge on the ethical dimension of (semi-) automated control of urban mobility, which will therefore also be addressed.

Different from other EU-funded projects such as INDECT[1], VideoSense[2], or the German Project CamInSense[3], SECCRIT is, however, not primarily a surveillance project. First and foremost, SECCRIT is a project about technologies for secure cloud computing for critical infrastructure IT. It would therefore be of questionable value to undertake extensive deliberations on the ethical dimension of (semi-) automated control of urban mobility and – in particular – video surveillance systems *as such* even though these are clearly not the primary subjects of the project. On the one hand, such considerations are to be left to projects focusing on research in these domains. On the other hand, video surveillance and urban mobility systems do clearly play a certain role as specific use cases for the cloud computing technologies developed in SECCRIT and should therefore not be factored out completely.

In a nutshell, the ethical considerations to be studied in the following sections must, on the one hand, pay regard to the ethical problems and conflicts possibly arising from the employed use cases of video surveillance and urban mobility systems while, on the other hand, keeping focus on the project's development of technologies that enable secure cloud computing for critical infrastructure IT. In order to serve these two partially conflicting goals best, this document employs a "triangulation approach" to identify ethical aspects that are specific to *cloud-based* video-surveillance and urban mobility services. In a first step, the most relevant ethical aspects of video surveillance and (semi-) automated urban mobility systems are summarized. This is done in section 0. Following these fundamental considerations, several characteristic qualities of cloud computing with particular relevance in matters of potential ethical impact are identified – again as discussed in the scientific literature. At the beginning of the project, however, research on "ethical aspects of cloud computing" was still in its infancy and had not yet reached a state of consolidation. A gathering of those ethical aspects of cloud computing that were already identified during the first year of the project is provided in section 3. The issues thereby identified were then used to form a "delta" between cloud-based and non-cloud systems in matters of the ethical dimension and served as a starting point for putting the rather generic ethical considerations on video surveillance and urban mobility systems into a SECCRIT-specific context. Section 4 does exactly this and identifies SECCRIT-specific potential ethical problems of *cloud-based* video surveillance and *cloud-based* urban mobility services from the perspective of the beginning of the project.

For each of the identified SECCRIT-specific potential ethical problems, we then identified measures for addressing it properly to ensure that the project is always acting in line with ethical values. Where necessary, this was also done for the rather generic ethical problems that arise from video surveillance, (semi-) automated traffic control and cloud computing. As a result of these discussions, safeguards ("SGs") were developed and employed throughout the of the project in addition to the measures taken in order to ensure legal compliance.

In addition to these aspects already covered in the initial version of this deliverable, the document at hand also takes a retrospective view from that last phase of the overall project and documents how the safeguards were actually applied and how ethical aspects were regarded in general during the various project activities. This is, complemented by some retrospective reflections, done in section 5. Finally, section 6 summaries and concludes this deliverable.

[1] See http://www.indect-project.eu/
[2] See http://videosense.eu/
[3] See http://www.iais.fraunhofer.de/5925.html

2 Ethical aspects in SECCRIT use cases

As depicted in the description of WP 6, two different demonstrators / use cases are inherent parts of the project in order to validate the practical applicability of the technologies to be developed in SECCRIT. The first demonstrator, particularly covering aspects of secure "storage and processing of sensitive data" in the cloud, will be based on video surveillance data. The second refers to "hosting critical urban mobility services" and especially focuses on questions of validation processes.[4] Following the "triangulation approach" outlined above, the main ethical aspects discussed for video surveillance and for (semi-) automated mobility services will therefore be gathered as a first step.

2.1 Ethical Problems of Video Surveillance

As outlined above, this section gives a compact overview of the most important ethical aspects of video surveillance in general. Even if SECCRIT is not a video surveillance project, these will have to be carefully taken into account in the course of evaluating potential technologies for secure cloud computing in the context of video surveillance, in order not to violate fundamental ethical values. Furthermore, these serve as a basis for identifying potential ethical challenges of cloud computing that would probably not have been seen without reflecting on the employment of this technology within a specific and sensitive application context.

Often-mentioned ethical aspects of video surveillance include the problem of people not being aware of being observed, the problem of well-informed statements of consent (or especially dissent) being hardly possible, the trend of surveillance being conducted excessively as compared to the original purpose, and the risk of video surveillance being conducted without a legitimate basis; for good reasons, these are addressed in existing national[5] data protection legislations. Such questions are therefore well covered by the principles-based considerations in the respective deliverables on legal aspects and will not be explicitly addressed here. It is, however, without any doubt that the respective legal givens have to be carefully taken into account during the implementation of the evaluation use case. The close coordination on the concrete implementation with responsible data protection authorities, reported in deliverable D 2.4, also requires that such aspects of data protection in the strict sense need to be appropriately addressed.

This being said, there are, however, further ethical aspects of video surveillance that go beyond questions of data protection. These have a long track of extensive scientific coverage that has been accompanying technological progress for decades. The history of ethical considerations on surveillance in general goes back even further, at least to Jeremy Bentham's well-known "**Panopticon**" concept (Bentham 1791), which refers to a specific architectural building structure explicitly targeted at manipulating inmates' behaviour and habits.

First described for prisons, manufactories, schools and other buildings where "*persons [..] are to be kept under inspection*", the concept consists of two factors. First is an architecture that allows any inmate to be seen at any time and without any exception from one central place of observation, for example a tower surrounded by circularly arranged cells with gratings facing to the centre ("*centrality of the Inspector's situation*", Bentham 1791, p. 23). This ensures that no "in-cell behaviour" can be hidden from an observer actually looking into the right direction. The second factor consists of a mechanism that prevents inmates from seeing whether an observer in the centre is currently looking into their direction or even whether an observer is actually present at all ("*seeing without being seen*", ibid., p. 23). This could, for example, be ensured by strong lights directed from the centre to the cells so that inmates are blinded when looking at the centre.

[4] „Electronic proofs" etc., see, e.g., D 2.2, section 2.2.2.

[5] In Germany, this is, for example, done in § 6b of the national data protection law. For a brief respective legal assessment of surveillance technologies as considered herein, see, for instance, Roßnagel, Desoi and Hornung (2012).

The ethically relevant impact of such surveillance situations is quite obvious: The mere knowledge of *possibly* being monitored makes people adapt their behaviour to the (assumedly) desired, normal, or least conspicuous-looking one and thereby leads to strong conformism. While Bentham himself intended exactly this for the context of prison inmates and other persons that are to be observed, and praised the so-called panopticon effect as a way for reaching a state of socially desirable behaviour. The effect is today usually ethically deprecated as it leads to the observed individuals having **zero privacy** (which is something completely different from "data protection") **even in most private situations** and because it raises the **risk of people negating themselves**, their beliefs and their needs in favour of being conformant with what the observer requests from them just to avoid punishments or even more subtle acts of enforcement.

The whole concept of the panopticon does, of course, apply to today's video surveillance in the same manner – and is of course the reason why most surveillance cameras are usually much larger and much more visible than technically required.[6] The **risk of people adapting their behaviour to a state of conformism** might be neglected as actually being a "risk" as long as punishment or other acts of enforcement are only exerted and expected in cases of behaviour that is definitely not accepted by society (e.g. in case of robberies, where the deterrent effect will usually be highly welcome). But depending on the political situation (which might, nota bene, be subject to change), the same systems can also be used to restrain people from, for instance, demonstrating against the government. The panopticon effect of people adapting to some (assumed) state of "target" behaviour is therefore one of the most important ethical risks of video surveillance in general, independently from the concrete technical design and capabilities.

Current surveillance technologies do, however, differ from the original panopticon concept in several regards. Under the term of "new surveillance", Marx (2002) identifies some specific characteristics of modern, **technology-mediated surveillance** as opposed to the traditional concept. First of all, modern surveillance is in most cases done by technological means, thereby **heightening the observer's capacity for recognition** (omnipresent video-cameras, heat sensors, etc. and, not to forget, the possibilities of modern data processing). Based on these technological means, current surveillance practices **break up the previously existing strong ties in matters of time, location and context** between the monitored behaviour, its actual recognition and interpretation, and the reactive response. While in the buildings described by Bentham, behaviour, notice, interpretation and reaction follow each other immediately and happen in direct local proximity, today's surveillance technologies allow the recording of plenty of manifold facts first and to analyse them later – be it on a periodic basis or in the case of specific events. And finally, it is a core characteristic of the so-called "new surveillance" that data (or **information of multiple kinds and from multiple sources) is integrated and combined**, leading to a much more encompassing basis for in-depth analysis. The fact of data being shareable and often shared among different parties strengthens this effect even further. This, in turn, lets the observers discover connections between different bits of information that would otherwise have been unnoticed and thereby potentially derive all-embracing images about personality and activities of those under surveillance – or rather, those under analysis.

Beyond these, there are plenty of further dimensions that distinguish current surveillance practices and technologies from former ones. Marx (2002) alone mentions 27 of them and there will presumably be more. An ethical assessment of all aspects of this shift is hardly doable in a

[6] The opposite problem of unnoticed, secret video surveillance, in turn, refers to traditional risks from the field of data protection. As mentioned above, these will not be addressed explicitly here as they are well covered in the respective legal deliverables (see, for instance, the respective sections on transparency in D 2.2). Nonetheless, the noteworthy dilemma of any planned use of video surveillance between possibly infringing upon individuals' data protection rights (unnoticed surveillance) and unwillingly pushing individual behaviour towards conformism (clearly visible surveillance) will not be ignored. For the cases covered in project SECCRIT, however, this dilemma does not arise practically as data protection laws require any video surveillance to be announced and be made visible to the potential subject of surveillance.

non-extensive way, especially because any of the changed characteristics can, depending on the context, lead to ethically welcome as well as undesirable implications. Marx himself gives the example that, "*through offering high quality documentary evidence and audit trails, the new surveillance may enhance due process, fairness and legitimacy*" (p. 22) while the same evidence could unquestionably also be used for ethically undesirable purposes. Furthermore, the ethical assessment of a concrete application of a certain technology also has to take into account the respective local cultural values.[7]

Due to these highly context-dependent implications, Marx (1998, p. 174) proposes an exhaustive "*set of questions to help determine the ethics of surveillance" without even trying to make strict normative arguments. Instead, he highlights* (ibid., p. 182):

> *„in matters so complex and varied we are better served by an imperfect compass than a detailed map. Such a map can lead to the erroneous conclusion that ethical directions can be easily reached or to a statement so far in the stratosphere that only angels can see and apply it."*

Therefore, these "compass questions" provided by Marx can very well guide the process of reflecting on the ethical dimension of a concrete application of technology-based surveillance and will therefore be applied throughout the establishment of the video demonstrator. They are reproduced in section 7 (Annex).

This leads us to the first safeguard that is to be applied throughout the establishment of both demonstrators in order to ensure that fundamental ethical values are kept in mind and appropriately addressed:

- **SG1 – Compass Questions: The "compass questions" reproduced in section 7 (Annex) will be applied in the design and implementation of the demonstrators.**

The strong context dependence notwithstanding, one aspect of the "new surveillance" as described by Marx is unquestionable: through the possibilities for strong integration of different sources of information, for their combined analysis and for their exchange amongst different observing parties, the fact that surveillance is increasingly based on technology leads to a significant strengthening of the observers' informational position, as opposed to the observed. This is even truer for the next qualitative shift of surveillance practices as noted by Graham and Wood (2003) under the term **"*digital* surveillance".**

Graham and Wood argue that novel digital surveillance technologies and practices, as opposed to former video surveillance technologies and practices, not only increase the amount of video data being collected and of areas being covered significantly but that these technologies and practices also lead to changed social practices and a novel category of ethical problems: Of particular relevance in this regard is the shift towards **automated analysis** ("algorithmic video surveillance", "algorithmic CCTV", Graham and Wood, 2003, p. 231, 235ff) with backend systems fed by digital cameras (and potentially other kinds of sensors) automatically analysing recorded data in order to recognize individuals and their movement, thereby allowing for far more **"meaning" to be automatically extracted** within video surveillance systems. This general concept can come in different flavours from comparing faces, or gaits of people walking along the street, against a "watch list" database of to-be-recognized individuals[8] over the automated movement tracking of single persons across areas covered by different cameras and other sensors to the automatic "detection" of "suspicious" or otherwise unwanted behavioural patterns that can, for instance, be used for social segregation within privately controlled, semi-public areas like a mall (Kang and Cuff, 2005, sect. IV.A).

[7] For a vivid depiction of significant differences in the perception of social acceptability in matters of privacy even between the closely related cultures of Europe and North America, see especially Whitman (2004).
[8] See, for example, Graham and Wood (2003, p. 236).

In this model, decisions based on the video data captured by surveillance systems thus increasingly shift from being made by human operators under (conscious or unconscious) comprehension of social values as well as human experience and discretion, to a mode where decisions are made or at least "suggested" automatically by means of algorithms explicitly directed at the detection of certain "patterns" that define, for instance, "usual", "unwanted" or "suspicious" behaviour. It is, however, **questionable to what extent such codified patterns actually match with what a well-trained human observer would identify as noteworthy**. Even worse, any algorithmic assessment is, at least to a certain extent, made on the basis of "*social and political assumptions that software producers embed (consciously or unconsciously) into their algorithms years before and thousands of miles away from the site of application*" (Graham and Wood, 2003, p. 242), provoking considerable "misinterpretations". And finally, the actual capabilities of technology are all too often strongly overestimated by the responsible parties. In the end, this leads to the "recognition" of a person as behaving "unacceptably" or any other process of "meaning extraction" **not being critically questioned in the light of human weighing and the social assumptions applicable to the respective context**. Instead, the technological assessment is all too often simply taken as a "given fact", leading to a **significant risk of mistreatment**.

To conclude, modern video surveillance technology in particular induces the risks of

- significantly reducing privacy of the observed,
- fostering conformism instead of individualism,
- leading to strong inequalities in matters of informational power between observer and observed through integration of multiple sources of information and their collection over long periods,
- automated analysis being performed on the basis of potentially defective or inappropriate algorithms without critical human reflection in the light of the application context's social givens, and of
- results of automated analysis being taken "as an objective fact" instead of "as a hint", potentially inducing significant mistreatment.

Even if the voluminous literature on surveillance, technology and their ethical implications discusses innumerable further aspects emerging from domains such as workplace surveillance (e.g., Hansen 2004), urban planning (e.g., Koskela 2000) or Gender Studies (e.g., Monahan 2009), these are the most important ethical aspects of video surveillance that serve as the basis of discussion for the remainder of this deliverable.

Moreover, it has been pointed out that the ethical impact of surveillance technology significantly depends on the concrete application scenario and context. As one and the same technology can prove ethically questionable as well as welcome depending on its application context, generic considerations about ethical implications of a given technology as such can never be exhaustive and must always be complemented by a context-specific assessment. This has been done throughout the establishment of the demonstrator based on the table of the "compass questions" and is reported on in section 5.1.

2.2 Ethical Problems concerning (semi-) automated Traffic Management Services

Different from the intended evaluation demonstrator of cloud-enabled video-surveillance, the subject of the second demonstrator holds – at least at first sight – no obvious ethical challenges going beyond those already discussed above. Of course, (semi-) automated traffic management services could also include surveillance equipment monitoring the current traffic situation, for example. This would in turn induce all the risks already mentioned above, including the generation of **movement profiles** based on recorded individuals or number plate recognition.

Beyond video surveillance, comparable problems could in the future also emerge from advanced models of **Car-2-X-communication being integrated with traffic control systems.** Within such systems, there exists an inherent risk that single and explicitly identifiable vehicles (and, thereby, their users) are subject to **tracking and monitoring**. An early form of such mechanisms which could very well also be integrated into a traffic control system are congestion detection mechanisms based on the density and movement speed of mobile phones like the one already employed by Google: Based on a multitude of Android mobile phones repeatedly sending their location data to a centralized platform, it is possible to derive a current traffic situation in potentially higher accuracy than with specialized sensors being installed near to the roads and nearly without any specific investment necessary. While this provides a technically interesting option for gathering current information about traffic flow, it is obvious that such a model induces the significant risk of the movement of individual persons being tracked without their consent and/or knowledge – at least as soon as these information are not strictly **anonymized or pseudonymized**. Were such technologies integrated in future traffic control systems without strong anonymization techniques, this would significantly strengthen a whole group of risks emanating from public bodies being able to track and locate individuals in real time as well as in retrospect. Such aspects would, however, already be covered either by existing data protection laws or by the ethical considerations on video surveillance laid out above and would therefore not call for specific considerations about the second demonstrator here.[9]

Beyond aspects that are reasonably attributed to the fields of data protection and video surveillance, the first aspect of (semi-) automated traffic management that could be extracted as being of possible ethical relevance from existing literature refers to the risk of **social discrimination / unequal treatment** based on rules that are ethically questionable. Graham and Wood (2003, p. 238f), for example, mention London's so-called "*congestion charge*" as well as "*private premium highways*" being erected in Toronto, Los Angeles, etc. as examples of traffic infrastructure that are practically accessible only to a selected portion of all drivers (those who can afford to regularly pay for using certain roads, for example). An example pointing in the opposite direction is that of so-called "*high-occupancy highway lanes*" which are well-known in the US and may only be used by cars with a minimum of two, three, or even four occupants, thereby introducing an incentive not to use cars alone and thus not to contribute to traffic congestion more than necessary. Generally speaking, both directions of differential treatment raise questions of ethical acceptability.

Furthermore, ethical questions also arise from the possibility of **enforcing the respective rule sets automatically through technology**. If, for instance, a (semi-) automated traffic management system would block any driver not able to pay an extra fee or not equipped with specific technical devices from entering a city centre during certain "high-congestion times" by means of physical barriers (comparable to the toll stations present on some European highways), this would induce the risk of **exceptions** which would traditionally have been unquestionably accepted **not being possible anymore** due to the usual **strict mode of technology-based enforcement**. As long as no alternative path that allows "overriding" of technical constraints exists, this would raise severe ethical questions regarding, for example, system behaviour in unforeseen cases of emergency.

Besides the risks of technology-based differentiation, and of unquestionably accepted rule-breaking made impossible because of absolute technological enforcement, no ethical challenges specific to the demonstration use case of traffic control could be identified that go beyond surveillance aspects already covered in section 2.1. Depending on the concrete design for the evaluation demonstrator, the first-mentioned aspects will have to be kept in mind and, if applicable, to be addressed appropriately throughout the evaluation phase to prevent ethically questionable uses of the demonstration system. The latter aspects from the domain of data protection in the strict sense as well as from the broader area of surveillance, in turn, will in all

[9] It is, however, also clear that the approach of ethical assessment on the basis of the "compass questions" developed by Marx (1998) should also be applied to this demonstrator.

likelihood play a key role for the concrete demonstrator and will therefore be addressed comparably to that of the video-surveillance demonstrator, i.e., by means of sound legal assessment and on the basis' of Marx's compass questions. Finally, the close cooperation with responsible data protection authorities reported on in D 2.4 also ensures at least the most relevant ethical and data protection issues to be properly accounted for.

3 Characteristics of Cloud Computing with Potential Ethical Impact

After the generic ethical aspects relevant for the concrete evaluation demonstrators planned in SECCRIT have been laid out so far, we now come to the main subject of the project and its potential ethical impact, namely cloud computing. As outlined in the DoW, these considerations will – besides also ensuring that all project activities are in line with established ethical values – provide "first high-level considerations that shall serve as a starting point for discussing the further development of the regulatory framework relevant for cloud computing." Furthermore, some first "concretizing deliberations regarding the transformation into the relevant regulatory framework" are to be derived. This will be done in the subsequent sections.

The identification of concrete characteristics of cloud computing that have potential ethical impact is subject to two main constraints: First, the whole field of cloud computing is – as compared to other technological domains – still in its infancy and especially lacks broad and consolidated coverage in the socio-technical domain. Different from, for example, surveillance technology, there is thus no well-established body of scientific discussion, categorization or classification of ethical aspects having to be considered in the context of cloud computing. Even if taking into account some early findings that are currently emerging in this regard, the following considerations are therefore necessarily of explorative nature, too.

Second, cloud computing is also subject to the well-known problem of general-purpose technologies being hardly assessable in matters of their expectable implications, including those of ethical relevance. Technologies like cloud computing, which basically do not carry an inherent "target application" but can rather be employed for a broad variety of ethically welcome or objectionable purposes, are particularly subject to the control dilemma prominently described by Collingridge (1980): as long as a certain technology is in its infancy, its (social) implications can hardly be assessed reliably. On the other hand, once the technology is mature enough for making reliable estimations on its implications, it is usually too late to actually influence development in order to achieve (socially) welcome outcomes. As ethical considerations do, like legal ones, always require a concrete application scenario (or "a case") for being meaningful, ethical considerations about new technologies are usually conducted on the basis of prognostic scenarios ("*scenario-based technology assessment*"). This, however, turns out to be of limited explanatory value with regard to general-purpose-technologies due to the prognosis- and control-dilemma, leading to the problem of (social) implications of general purpose technologies being hardly assessable at all in a reliable manner (Weber 2010). Nonetheless, a structured analysis of the most significant characteristics of a certain technology may at least allow for an identification of *possibly* expectable implications – albeit with the risk of significantly overestimating some aspects and overlooking others.

In the light of these two constraints, we will in the following identify *possible* problems of ethical relevance that could – given the technological characteristics of cloud computing – be induced by a switch from traditional computing models to cloud-based ones. In so doing, we will explicitly exclude aspects that are already covered in deliverable D 2.2 on legal fundamentals. In particular, this refers to concrete questions of data protection law which are already subject to current legislation. Instead, we will concentrate – in line with the intentions from the DoW –on higher-level aspects that will presumably cause problems not adequately addressed by the current regulatory status quo.

In the subsequent sections, these generic considerations will then be projected onto the scenario-specific deliberations already outlined above in order to generate more concrete ideas on the social implications of cloud computing being used within ethically relevant applications.

3.1 Loss of Control

The first foreseeable potential problem of ethical relevance induced by cloud computing emanates from the fact that the outsourcing of computing resources leads to a shift of control over these resources. Whereas data, processes etc. had formerly been residing on a local infrastructure, under cloud computing they will have relocated to the cloud provider. This necessarily implies at least a partial loss of control over them as the cloud user can exert control only via well-defined interfaces defined by the cloud provider. The cloud provider, in turn, is at least basically able to take notice of, or – depending on the concrete technical givens – even manipulate data as well as processes without being observed.[10] In the end, it is basically the cloud provider that has the ultimate control over data and processes and, in addition, ultimately determines the further (technical and non-technical) conditions under which data processing takes place.

The risk of control loss becomes even more significant in cases of interconnection between multiple cloud service providers. The ongoing trend towards end-user services being composed from other services provided by different cloud providers leads to the apparent problem of actual responsibility for the ultimate outcome being vaguely distributed across a multitude of involved parties.

Given this possible shift of control, forcing the cloud user to be the only one accountable for any kind of malfunction or misbehaviour by its contract partner (as is basically done by the current legal framework) raises concerns of appropriateness and fairness. Especially for cases of possible intrusion by non-authorized third parties, or of possible failures or, indeed, corruption[11], it would surely be important to prove what exactly went wrong. This is particularly true for cases with services from different providers being integrated with each other where, without technical mechanisms of proof, it would be impossible to attribute unwanted outcomes to the party that actually caused it. Ultimately we want to avoid unfair risk distributions and, in consequence, economically inefficient outcomes.

The potential loss of control, however, also induces ethical risk in areas much more closely bound to the individual. Whitman (2004), for example, points out that at least the European understanding of "privacy" for large parts refers to the individual's "rights to control your public image, rights to guarantee that people see you the way you want to be seen." With regard to personal data as covered by European data protection laws, this implies the need for exerting control over data that can unquestionably influence this "image" that is well-covered in the term of "informational self-determination": Any person is then basically in control over data that might influence the image that others have about her or him. Of course, this individual control following pure self-determination is already limited in traditional models of data usage and is – within well-defined boundaries defined by data protection laws – to a large extent exerted by the respective holder of the data (in legal terms, the controller). But within this traditional model, the data subject also has a couple of rights against the controller; and the controller, in turn, has a couple of responsibilities and must follow several comparably strict obligations to ensure that the individual has the largest possible control over the ultimate handling of data that forms her or his "person image".

[10] Notably, this also includes the risk of the cloud provider secretly exploiting the respective data for own purposes. See, for instance, Cavoukian (2008).
[11] See Paquette, jaeger and Wilson (2010).

In the case of cloud computing and especially in those cases involving the complex interconnection of services provided by different parties[12], this control over data that influences the person's image tends to decrease significantly. Instead of actually being in control the data subject will therefore increasingly have to rely on other mechanisms like "trust" regarding the question of whether their data falls into the wrong hands. The "right to control one's public image" thereby could become significantly harder, if not impossible, to be actually exerted for the individual. In order to prevent such a technology-deterministic fading of enforceability of fundamental rights, mechanisms should be developed that render the aforementioned "trust" back into actual control over the dissemination of personal data.

In any case, the generic risk of control loss can be identified as one possible source of changed socio-technical givens introduced by an increased use of cloud computing. This change risks the emergence of new ethical challenges, whereas questions about fairness and about the possibility of controlling the use and dissemination of personal data will presumably be of particular relevance.

3.2 Lack of transparency

Closely coupled with possible loss of control is the risk that processes of data handling increasingly lack transparency because of cloud computing. Transparency, as understood here, refers to the ability to know and retrace what actually happens within a system or service being used. Cloud computing, however, often assumes that services are used as "functioning black boxes" that exempt the user – be it a private person or an organization – from having to care about the internal details.

This reduction, or loss, of transparency again has potential implications with regard to the above-mentioned European understanding of privacy. If an individual cannot know and retrace how her or his personal data will be handled, it becomes nearly impossible to exert control over this data handling. Generally speaking, the same problem also applies to organizations using cloud services. Without possibilities to assess the internals of a used service, it can hardly be ensured that obligations of any kind (legal, business, corporate ethics, …) are met in practice.

Again, mechanisms must therefore be developed in the future that do, at least to a certain extent, re-establish transparency of used services on demand. In particular, such mechanisms must ensure that the cloud user can check what is actually happening to certain data transferred to the cloud and if the cloud is actually performing as intended. Key requirements for such mechanisms are the establishment of transparency about how the data will be handled and the exposure of different responsibilities, especially for cases with multiple actors being involved in a certain process (see above discussion on "loss of control"). This goes in the direction of one of the core challenges of the SECCRIT project as evidence and data protection have opposing goals when recorded actions cannot be revealed to any person, unless data protection regulations are totally ignored and therefore violated. Consequently, it seems plausible to distinguish between recording and exposing transparency information, whereas simultaneously evidence and data protection law must be adhered to.

3.3 Transfer across cultural and legislative boundaries

As cloud services can be located in any part of the world, cloud computing increases the risk of data being exposed to different cultural settings.[13] This raises the risk of an individual's data not being treated in conformance with his/her own cultural and ethical values, but rather in the light of the social context in which the storage or processing is taking place. In a comparable way, this is also true for the legislative setting, where individuals can no longer rely on the national legal framework with which they are familiar, but rather bear the risk of their data being handled on the

[12] See also Timmermans, Stahl, Ikonen and Bozdag (2010, p. 5).
[13] See, for example, Timmermans, Stahl, Ikonen and Bozdab (2010, p. 5).

basis of regulations that they do not know and probably would not accept if they were aware of them.[14]

In both variants, the trend towards increased data transfers across cultural and legislative boundaries heightens the risk of significant incompatibilities between what is expected or deemed acceptable by the individual with respect to data and what is actually done with that data. From an ethical perspective, this can be of particular relevance with regard to privacy, as the respective cultural conceptions significantly differ across the world.[15] The same risk could, however, also exist with regard to other domains shaped by cultural or legal differences whenever data is transferred across the respective boundaries.

Particularly with regard to the legal framework, the transfer across boundaries not only bears the risk of individuals' rights being significantly limited[16] but also of these individuals not being able to exert their rights properly. This could, for example, be due to changes of the parties that the respective individual would have to approach, due to specific requirements having to be met in order to make a request eligible, due to language barriers, etc.

The question of how data (and, in the future, processes) crossing cultural and legislative boundaries at massive scale could be properly addressed has to remain open at this point. Of course, we may consider approaches such as "unifying legal frameworks", but when discussing this option it has to be kept in mind that legal frameworks are always rooted in the respective cultures in which they are to be "operated". Alternatively, one could consider "extraterritorial islands" where a cloud service within one cultural or legislative framework is, under certain preconditions, operated in accordance with cultural values and legal regulations from another one. The latter would reflect the current approach of Europe, which tries to ensure that European data protection law is to be applied to "European personal data" even when stored and/or processed in cloud services outside of Europe. It is, however, obvious that this approach raises conflicts because it would logically result in two different legal frameworks being applicable to one and the same piece of data or processing instance.

For now, the identification of the expectable conflict arising from increased transfer across cultural and legislative boundaries must suffice. The question how it can effectively be dealt with shall, however, be further addressed throughout the remainder of the project.

3.4 Inherent Risk of Monopolies and Lock-Ins

Finally, one more aspect of cloud computing that bears certain indirect ethical risks will be discussed in brief, namely the inherent tendency of cloud computing services to develop so-called natural monopolies and the related risk of significant vendor lock-ins. Generally speaking, cloud services feature typical characteristics of information goods / network goods as depicted, for instance, by Shapiro and Varian (1998).

First, cloud services are subject to significant economies of scale; the provision of the first unit (an hour of computing time or a gigabyte of storage being offered through a sophisticated service interface, one user served with a complex SaaS application etc.) induces high initial development costs, while once the first unit has been provided, the costs for providing another unit are significantly lower or even negligible. This cost structure usually leads to a single vendor being

[14] In this regard, the current conflict between Europe and the US about data transfer obligations for US corporations and even their European subsidiaries is an example with high relevance in the context of cloud computing.
[15] See, for example, Capurro (2005) or again Whitman (2004).
[16] Again, just think about the current conflict between the US and Europe and, in particular, the obvious fail of the so-called „Safe Harbor Regulations", which were explicitly intended to build a bridge between two fundamentally different cultures and legislations in matters of data protection.

able to serve a whole market at lower costs than would be the case with a multitude of competing vendors and therefore stimulating the emergence of monopolies.

Second, cloud services can usually be seen as so-called network goods which are characterized by providing a value to any single user that increases with additional users joining the network. This value-increase can emanate from more possibilities for interaction with other members being present, from more complementary goods being available, or from a higher availability of specific qualification on the market, etc. Especially with regard to complementary goods and specific skills, it is plausible to assume this network effect to be present with regard to IaaS, PaaS and SaaS cloud services. In addition, SaaS cloud services will presumably also feature benefits from more possibilities for communication and data exchange among different users. These network effects also provoke market monopolization.

And third, cloud services can also be characterized as being subject to lock-in effects. These basically arise whenever a certain specific good or service is integrated into a customer's processes in way that would make it so expensive for the customer to switch to an alternative supplier that it is economically rational to stick with the current one even if the alternative offer came at zero costs. As can be prominently seen in the market for office software, a dominant vendor can of course consciously design his product in a way that heightens these switching costs and thereby strengthens its dominant position. The risk of this also happening in the context of cloud services is obvious: IaaS and PaaS providers could make arrangements that prevent customers from easily migrating virtual machines, data or application logic to another provider while SaaS providers could make their products incompatible to other ones in a multitude of further ways. In the end, all three characteristics supply the emergence of monopolistic structures and work against a functioning market in the traditional sense.[17]

Even if this effect should at first sight be subject to considerations on traditional market regulation, it could also have ethical relevance. Within a monopolistic market, the dominant player enjoys significant bargaining power that can be (mis)used to force customers to accept terms of use, prices or other contractual conditions that are clearly detrimental to them. In lack of realistic alternatives, this could also lead to ethically relevant "decisions" being made under the constraint of no viable alternative being available. Bearing in mind that cloud services are often realized on the basis of other cloud services, this monopoly power of one player may furthermore also extend to relations between the customers of the single player and their customers, respectively. In such situations, the concept of "individual consent", which is, for example, often referred to in the domain of data protection, is clearly of only limited value. Similarly, concepts like "privacy as competitive advantage"[18] – even if discussable in other contexts – would under such market conditions hardly change anything. Instead, ways must be found that ensure compliance with ethical values (to which "choice" might also belong) within monopolistic markets, too.

Finally, the monopolistic character of cloud computing could also lead to ethical risks on the societal level: Given that the cloud market is driven by economic mechanisms that foster monopolistic structures, there is the certain risk that more and more services of various sorts are realized on the basis of the fundamental services provided by one single player. This could then lead to a single player posing a "systemic risk" – a factor that is so omnipresent and so interwoven into so many aspects of daily life that it becomes indispensable for societies as a whole. Again, this holds tremendous extortion potential which might, for instance, also be exploited in matters of not accepting "overly strict" regulatory restrictions regarding data protection, security, etc. Even if questions of market regulation are unquestionably far beyond the subject of SECCRIT, this potential shall be kept in mind throughout future considerations on the further development of the existing regulatory framework relevant to SECCRIT topics.

[17] See also Gentzoglanis (2012).
[18] See, e.g., Hoeren (2000).

4 "Cloudification" of SECCRIT Use-Cases and Safeguards for Project Implementation

In a third step of the "triangulation approach" laid out in section 1, we now have to discuss which specific ethical problems could potentially result from the "cloudification" of the use-cases employed in SECCRIT. In order to identify these possible issues, we interweave both dimensions – the ethical aspects specific to the two use cases as identified in section 0, and the ethically relevant characteristics of cloud computing as laid out in section 3 – with each other, and we examine where the concurrence of two effects either considerably strengthens an already identified risk or even changes its nature significantly. Due to the fact that ethical assessments strongly depend on the concrete application context and in line with the above-mentioned problems regarding the impact assessment of general purpose technologies, we will structure these considerations along the two use-cases. Basically, we thus look for the changes implied by the ethically relevant characteristics of cloud computing upon the ethical problems already identified for the use-cases.

In so doing, we will also lay out how the respective risks – cloud specific as well as the generic ones identified in section 0 – will be addressed throughout the remainder of SECCRIT in order to ensure that all activities within the project are not only in line with legal requirements but also with ethical imperatives.

4.1 "Cloudification" of Video Surveillance

As the first ethically relevant aspect of video surveillance, we identified the risk of the privacy of those under observation being significantly reduced, that is, the risk of hardly any behaviour being definitively unobserved, depending on the space under consideration (a prison, the workplace, semi-public or public spaces, …). With strict regard to this risk, the identified characteristics of cloud computing (control loss, lack or loss of transparency, cross-cultural and cross-legislation transfer, risk of monopolies and lock-in) have no relevant impact. The amount of conduct being potentially observed solely depends on the extent to which the space under consideration is covered by surveillance equipment and not on the mechanisms and technologies employed for handling and analysing the respective recordings.

The risk of significant privacy-reduction being implied by video surveillance in general must, however, also be addressed throughout the project and, in particular, throughout the establishment and operation of the video surveillance demonstrator. This leads us to the following safeguards:

- **SG2 – No surveillance of private or public spaces: In order to prevent privacy invasions, no evaluation activities will be conducted based on the surveillance of (semi-) public or even private spaces.**
- **SG3 – Experimental environments: Evaluations and tests of the secure cloud technologies that are to be developed in the project will, with regard to video surveillance, only be conducted on the basis of experimental environments that are explicitly set up for the project.**
- **SG4 – Explicit Consent (Video): Evaluations and tests will, with regard to video surveillance, only be conducted on the basis of recordings that are specifically produced for this purpose with explicit, well-informed and written consent of the persons being recorded.**

Through these safeguards, we prevent any ethical conflicts with regard to privacy invasion from emerging through the evaluation of our to-be-developed technologies applied to the scenario of video surveillance. Furthermore, this also ensures legal compliance of the respective evaluation activities.

With regard to the second ethically relevant aspect of video surveillance – the risk of fostering conformism – the above-mentioned explanations on the absence of specific consequences of cloudification similarly apply. The risk emanates from the individual's knowledge of (possibly) being observed and not from a specific way of data being handled or analysed. The above-mentioned safeguards therefore also ensure that conformism-related risks do not arise from the evaluation use-case being implemented.

Different from the first two aspects, there is a risk of ever-increasing inequalities of informational power between the observer and the observed because of the possible integration of multiple sources, as already described by Marx (2002) under the term of "*new surveillance*". This is affected by the cloud-specific aspects of ethical relevance as identified in section 3. In particular, the risk of "control loss" could have a reinforcing effect on the inequalities of informational power, as cloud-technologies make the integration of a multitude of different data sources, the sharing of such data and their long-time storage even cheaper (and thus, more likely) while they at the same time decrease the observed individuals' possibilities for exerting influence on what is actually done with recordings showing them. Furthermore, this potential loss of control is not only confined to the observed but might also apply those parties that operate surveillance systems on the basis of cloud technologies. For example, a security firm monitoring a critical infrastructure like a subway station by means of a cloud-based surveillance system, can (without further technologies being used) not be sure that video recordings are not passed on to other parties by the cloud provider.[19]

Even if the technologies that are to be developed within SECCRIT explicitly counteract the risks associated with this control loss, through the establishment of advanced mechanisms for data flow control, for the generation of reliable digital evidence, etc., it cannot be guaranteed that these mechanisms are already well-functioning and complete during the evaluation phase. The outlined risk must therefore be addressed properly. In order to prevent potentially adverse and uncontrolled transfer, integration, analysis, etc. of video recordings, we will thus only use well-controlled cloud environments that are separated from the public Internet through strong state-of-the-art security mechanisms for our demonstrators. These well-controlled cloud environments are provided by one of the project partners (Amaris) and exclusively hosted in Europe. Different from other well-known cloud providers, we thereby employ more sophisticated and effective separation and security mechanisms and prevent any onward transfer of the respective data to third countries.

- **SG5 – Controlled cloud: Evaluations will only be conducted within well-controlled cloud environments separated from the public Internet through strong state-of-the-art security mechanisms and hosted within the European Union.**

The risks arising from the characteristic of lack of transparency also tends to reinforce the risk of significant inequalities of informational power as the observed individuals not only cannot influence which parties are able to access video recordings showing them, but they cannot even know who actually has this access. And depending on the concrete system design, even the operator of a cloud-based video surveillance system might not be aware where the respective

[19] Besides the ethical dimension, this would also be problematic from the legal perspective as in this case, *the security firm* would have to ensure that data is handled correctly and not passed on to unauthorized parties. Such considerations are, however, subject to the legal deliverables D 2.2 and D 2.7 and shall therefore not be discussed in detail here.

data will be stored or processed. Without this knowledge, however, any attempt to fill the idea of "informational self-determination" with life is doomed. Again, the technologies that are to be developed within the project shall mitigate this lack of transparency faced by the different parties today[20], but it cannot be guaranteed that these are already in full function during the evaluation phase. The confinement to a well-controlled and well-protected cloud setting together with the explicit, well-informed and written consent introduced above should, however, counteract ethical risks that would otherwise possibly have emerged in matters of (lack of) transparency.

A potential impact of cloud-specific risks on the risk of automated analysis being performed on the basis of potentially defective or inappropriate algorithms could emerge from the context of transfer across cultural and legislative boundaries. If, for example, a cloud-based analysis service should in the future be used ad-hoc within something like a "service based video surveillance system", this would pose a significantly higher risk of the employed analysis service not matching the social expectations of the observed (and possibly also those of the observers) than it can be assumed within consciously designed and well-conceived integrated systems. Such models are, however, far beyond the scope of the SECCRIT and will therefore not be employed within the evaluation demonstrator. Instead, the confinement to a well-controlled and well-protected cloud setting exclusively hosted within the European Union will prevent this risk from actually gaining relevance at all.

Regarding the risks of cloud computing posing the inherent risk of natural monopolies and lock-in-effects in turn, no expectable interferences with other risks could be identified for the video surveillance demonstrator. Also, no specific safeguards could be identified as being necessary for preventing negative effects of this rather general risk throughout the project. Nonetheless, this potential risk exists and shall therefore be kept in mind during further activities. In particular, this will be done whenever it is reflected on the further development of the existing regulatory framework in Task T 2.4.

As already stated in SG1, we will furthermore conduct ethical assessments of the demonstrators during their establishment based on the table of "compass-questions" in order to prevent potential negative impacts that were missed in the analysis so far. During this assessment, the risk of results from technological analysis being inappropriately taken as objective fact will also be covered, even if the risk of "mistreatment" being done on the basis of such a misinterpretation is rather limited within our evaluation setting.

A tabular overview on how the different risks and their identified interdependencies are counteracted within the project will be given below. Before doing so, we must, however, also analyse the potential risks arising from the "cloudification" of (semi-) automated traffic management in the same way we already did for the "video surveillance" use-case.

4.2 "Cloudification" of (semi-) automated Traffic Management

Concerning the (semi-) automated Traffic Management scenario, we identified the potential risks of possibly tracking individuals and establishing movement profiles of the observed persons by recording faces of those individuals or the number plates of their cars. The identified cloud-characteristics of losing control, lack of transparency, cross-cultural and cross-legislation transfer, risk of monopolies and lock-in have partially a direct impact on these risks.

This is especially the case for the loss of control, lack of transparency and cross-legislation transfer. In contrast to the video surveillance scenario, where we can build our own "experimental

[20] And again, there are also strong legal motivations for doing so which shall not be discussed in detail here. See, for instance, the sections of „transparency", „user rights" or on the provision of digital evidence in deliverable D 2.2.

environment", we are for this evaluation demonstrator forced to take recourse on the infrastructure of the city of Valencia. It is therefore not possible to make our assessment without really observing a "real" and not a simulated volume of traffic in order to manage it properly as required within the SECCRIT project.

In order to nonetheless avoid privacy invasions as well as the generation of movement profiles, social discrimination and further risks already identified in the context of video surveillance, it is therefore absolutely necessary that individuals cannot be identified within our demonstrator setting – neither on the basis of their faces or other characteristics nor through by other means like cars.

- **SG6 – Obviate identification: It is technically ensured that faces, number plates and other features possibly allowing the identification and tracking of individuals are made unrecognizable whenever visual data (video, still images) is used in the traffic control demonstrator – be it through blurring or through a respective reduction of image resolution.**

Beyond this, as cited in the section on the cloudification of video surveillance, the use of a **well-controlled cloud environment** for the storage and processing of the data (**SG5**) is correspondingly required.

Regarding a possible integration of Car-2-X communication technologies into the demonstrator, corresponding mechanisms for preventing the tracking of individuals and any possible discrimination are necessary, too. This can be done by means of strong technical mechanisms of anonymization or pseudonymization. As, however, identification of individuals could nonetheless be possible directly after data gathering and before their anonymization, well-informed and written individual consent is still to be obtained for any possible integration of Car-2-X data into the demonstrator.

- **SG7 – Anonymization/Pseudonymization: Whenever advanced technologies of Car-2-X communication are to be integrated into the traffic control demonstrator, the respective data will only be used in anonymized or pseudonymized form, preventing any attribution of data to individuals.**
- **SG8 – Explicit Consent (C2X): Any potential integration of data from Car-2-X communication and similar systems is only done if the respective individuals have given their (well-informed and written) consent in advance.**

The risk of cross-cultural and cross-legislation transfer will again be addressed by employing a well-controlled secure cloud environment exclusively hosted within the European Union (**SG5**), so that it will not gain any practical relevance for the traffic control demonstrator, too.

Potential problems with regard to monopolies and lock-ins will not be addressed specifically as they are not a threat for the concrete scenarios of the SECCRIT project, but rather shall be kept in mind during future deliberations on the development of the regulatory framework.

Concerning the selection of "*who is permitted to drive at a certain time*" this risk of discrimination/unequal treatment (congestion charges, extra lanes) emerges from the (semi-) automated traffic system itself and is not significantly affected by "cloudification". Even if the respective data and processes are relocated to external servers, this would lead to a problem only if the data makes it possible to draw conclusions about the individual from it which could then form as a basis for discriminations. For our demonstrator, these risks are, however, counteracted through the employment of alteration mechanisms (**SG6, SG7**) which are aimed at avoiding conclusions about individuals as well as through the strict confinement to well-controlled cloud environments specifically set up for the project (**SG5**).

Last, but not least, the automatic enforcing of rule sets through technology is also emanating from the (semi-) automated traffic management itself without significant cloud-specific aspects. The only imaginable impact of cloudification could arise from the cloud provider arbitrarily manipulating the technology to the disadvantage of the road user. Due to the well-controlled cloud environments used in the demonstrator (**SG5**), this risk will, however, play no significant role within the SECCRIT project.

4.3 Consolidated Overview

In order to provide a consolidated overview of the identified ethical risks, their potential interrelations with particular relevance and the safeguards that are to be employed throughout the project, the following table summarizes the findings obtained and the implications developed in this document. Whenever no specific risks were identified for a certain combination, this is marked as "n/a".

TABLE 1: CONSOLIDATED OVERVIEW OF ETHICAL RISKS

	General	**Control Loss**	**Intransparency**	**Transfer across boundaries**	**Monopolies and lock-ins**
General	n/a	SG5	SG5	DPA involvement (see D 2.4)	Continuous consideration during further development of regulation
CCTV					
Reduced Privacy	SG1, SG2, SG3,SG4	SG2, SG3, SG4, SG5	SG2, SG3, SG4, SG5	SG5	n/a
Conformism	SG1, SG2, SG3, SG4	n/a	n/a	n/a	n/a
Inequalities of informational power	SG1, SG2, SG3	SG3, SG4, SG5	SG3, SG4, SG5	SG5	n/a
Defective / inappropriate analytical algorithms	SG1	SG5	SG5	SG5	n/a
Results mistaken "as objective facts"	SG1	n/a	n/a	SG5	n/a

Traffic Control					
Movement profiles / tracking	SG1, SG6, SG7	SG5, SG6, SG7	SG5, SG6, SG7	SG5	n/a
Social discrimination / unequal treatment	SG1, SG6, SG7	SG5, SG6, SG7	SG5, SG6, SG7	SG5	n/a
Automated enforcement / no exceptions	SG1	SG5	SG5	SG5	n/a

The safeguards developed herein and referred to in the above table are:

- **SG1 – Compass Questions: The "compass questions" reproduced in section 0 will be applied in the design and implementation of the demonstrators.**
- **SG2 – No surveillance of private or public spaces: In order to prevent privacy invasions, no evaluation activities will be conducted based on the surveillance of (semi-) public or even private spaces.**
- **SG3 – Experimental environments: Evaluations and tests of the secure cloud technologies that are to be developed in the project will with regard to video surveillance only be conducted on the basis of experimental environments that are explicitly set up for the project.**
- **SG4 – Explicit Consent (Video): Evaluations and tests will with regard to video surveillance only be conducted on the basis of recordings that are specifically produced for this purpose with explicit, well-informed and written consent of the persons being recorded.**
- **SG5 – Controlled cloud: Evaluations will only be conducted within well-controlled cloud environments separated from the public Internet through strong state-of-the-art security mechanisms and hosted within the European Union.**
- **SG6 – Obviate identification: It is technically ensured that faces, number plates and other features possibly allowing the identification and tracking of individuals are made unrecognizable whenever visual data (video, still images) is used in the traffic control demonstrator – be it through blurring or through a respective reduction of image resolution.**
- **SG7 – Anonymization/Pseudonymization: Whenever advanced technologies of Car-2-X communication are to be integrated into the traffic control demonstrator, the respective data will only be used in anonymized or pseudonymized form, preventing any attribution of these data to individuals.**
- **SG8 – Explicit Consent (C2X): Any potential integration of data from Car-2-X communication and similar systems is only done if the respective individuals have given their (well-informed and written) consent in advance.**

Based on these safeguards, we will ensure that ethical aspects are consciously taken into account and addressed properly during further activities within SECCRIT. The SECCRIT Security group will therefore be responsible that these safeguards are followed throughout the implementation phase. Once again, we note that this deliverable explicitly concentrates on ethical aspects that are not, not yet, or not yet sufficiently referred to within the legal and regulatory framework covered in other deliverables (in particular, D 2.2 and D.2.7). Furthermore, aspects relating to the interaction with relevant national data protection authorities are also covered in a separate deliverable (D 2.4) and therefore not covered herein, too. Altogether, these deliverables and the deliberations included here (and there) form the basis for ensuring that all activities of the project are in line with well-established societal values and will prevent avoidable adverse impacts as far as possible.

5 Application and Retrospective Assessment

Based on these deliberations, it shall now be described how the identified ethical issues were addressed during the concrete design, implementation and test of the two SECCRIT demonstrators. For both demonstrators, this is particularly done on the basis of the compass questions reproduced in section 7 and the additional safeguards developed in section 4. Some of these safeguards are aimed at multiple ethics-relevant characteristics of Cloud Computing (Control Loss, Intransparency) in the context of the demonstrators while others are only aimed at single ones. Furthermore, some safeguards apply to both demonstrators while others only refer to one. For a respective assignment table, see section 4.3.

5.1 Application in Video Surveillance Demonstrator

As delineated in more detail in deliverable 6.1, the video surveillance demonstrator has been implemented to demonstrate the capabilities of cloud technologies for surveilling public spaces. As compared to the state-of-practice in this field – purely local installations of backend facilities (storage, video analysis etc.) at every single site to be monitored – a cloud-based approach is highly promising in matters of manageability, cost-effectiveness and adaptivity. On the other hand, moving substantial functionality away from the place being monitored to offsite facilities, induces several ethical concerns that have to be addressed appropriately. In the following, we will describe how this has been done in SECCRIT's video surveillance demonstrator.

The fundamental architecture that was erected in this demonstrator consists of several building blocks described in more detail in section 2.2 of deliverable 6.1. In particular, the general structure includes IP-enabled cameras with own edge-storage capabilities being installed on-site as well as various backend components (storage, video content analysis, systems management, etc.) that are developed as cloud-based components and together provide a "Video Surveillance as a Service (VSaaS)" solution which can be employed for serving multiple customers and sites. Finally, this VSaaS-System provides a user-interface that can be accessed by the security personnel responsible for a certain site.

Even if the demonstrator was clearly designed with the Helsinki main railway station as the surveilled site in mind (and even if the Finnish data protection authority would have considered the collection of "real-world" video data legitimate as long as certain further safeguards are in place, see section 2.2.2 of deliverable 2.4), ethical considerations beyond mere legal compliance clearly prohibit to erect the demonstrator in such a real setting during the development, evaluation and demonstration phase of SECCRIT technologies. Instead, a clearly restricted environment was established in order to strictly comply with the ethical safeguards introduced above. Furthermore, the demonstration was strictly confined to pre-recorded video showing

people who have given their explicit, well-informed and written consent to participate in the demonstrator evaluation before.

Further details on how the above-defined safeguards were practically applied within the video surveillance demonstrator will be laid out below in section 5.1.2. In line with the overall concept for ensuring compliance with ethical principles pursued within project SECCRIT, however, we will first analyse this demonstrator on the basis of Marx' (1998) compass questions explicitly designed for identifying ethically relevant aspects of concrete surveillance systems.

5.1.1 Application and Addressing of Compass Questions

As proposed section 0, the "compass questions" laid out in section 7 served as primary guidance for the ethically acceptable design and implementation of the demonstrators and their application is thus the first and most important safeguard (SG1) in project SECCRIT. In the following, we will therefore explicitly refer to each of these questions and lay out how they were assessed and – where applicable – addressed in the context of the video surveillance demonstrator. For an in-depth-understanding of the background of the questions, the respective explanations given in Marx (1998) might be helpful.

1. Harm: Does the technique cause unwarranted physical or psychological harm?

Physical harm is not to be expected from the technique used in the video surveillance demonstrator – neither in the demonstration setting explicitly erected nor in the real-world application to the Helsinki Central Railway station. Video-surveillance based on cameras is a physically non-invasive technique. Even if one might construct a physical harm resulting from an action of the security service based on the VSaaS System (e.g. a physical harm for a suspect restrained by the security service based on surveillance material), this risk can hardly be attributed to the specifics of a cloudified surveillance system as opposed to a traditional one. Unwarranted physical harm can thus be disregarded for the first demonstrator.

Psychological harm, in turn, might to a certain extent very well emanate from a video surveillance system in general. First of all, the panopticon-effect of feeling monitored and thus restricted in the individual freedom of action prominently described by Bentham (1791) and explained in section 2.1 is to be mentioned here. Independently from whether this effect does actually represent a "harm", however, it does again arise for any video surveillance system independently from the technical design. Following the "triangulation approach" pursued herein, it does thus not call for further examination.

Another risk of psychological harm resulting from video surveillance systems, in turn, emanates from the risk of incriminating or defaming images or recordings from the video surveillance system being made available to individual persons or even the public against the interest of the person being shown. Depending on the nature of the images or recordings, this might cause severe psychological harm to affected persons – particularly emanating from strong feelings of shame and from detrimental third parties' reactions to the images and recordings. Different from the other potential harms mentioned above, this risk very well depends on the concrete technological design of the surveillance system because of the direct effect on the probability of images and recordings to be disclosed. A system design that heightens this probability as opposed to another design also heightens the risk of the mentioned psychological harm actually coming true.

Basically, a strongly interconnected, non-local video surveillance system integrating video coverage of multiple sites through a cloud- and service-oriented model like the one developed in the video surveillance demonstrator heightens this risk because of the increased number of actors involved and the increased overall complexity as opposed to a traditional, isolated and

locally operated system. This, in turn, strongly calls for particular carefulness and for special measures to be taken in the concrete system implementation in order to limit the risk of unacceptable image and recordings disclosure to an acceptable level. In the demonstrator, this is done by the use of well-controlled cloud environments separated from the public Internet through strong state-of-the-art security mechanisms (SG5) as well as through the explicit confinement to pre-recorded video data gathered only after explicit, well-informed and written consent of the parties being shown. The specific technologies developed within project SECCRIT and prototypically deployed in the video surveillance demonstrator, in turn, are explicitly aimed at significantly limiting the risk of unwarranted images and recordings for real-world applications of cloud- and service-based video surveillance solutions. The technologies developed within project SECCRIT would therefore allow to profit from the technical and economic advantages of networked video surveillance solutions as opposed to traditional ones while still maintaining a low risk of disclosure and thus countervailing the originally heightened risk of psychological harm.

Physical or psychological harm arising directly from the process of collection, as originally considered by Marx, are not to be expected for this demonstrator.

2. Boundary: Does the technique cross a personal boundary without permission (whether involving coercion or deception or a body, relational, or spatial border)?

Regarding the violation of borders, a cloud-based video-surveillance system does not cross borders of the body and does not extend human senses in a way that would reveal information which could not be observed by a human person, neither in the demonstrator setting nor in the respective real-world applications. The same is also true for violations of expected non-observability as no recordings are made of private spaces in the demonstrator (see also SG2 in this regard) and can hardly be imagined for future real-world scenarios because of strict legal regulations in this field. The only risk of a “natural” border being potentially violated through the introduction of a highly connected, cloud-based video surveillance system as opposed to a traditional one might refer to the revealing of inner feelings based on advanced automated face and expression recognition technologies. Such technologies are, however, neither used in the system erected in the demonstrator and thought of for real-world applications, nor is the respective risk particularly subject to the *cloudification* of IT-based surveillance systems. Following the triangulation approach followed herein, a specific risk of natural borders being crossed does thus not arise for the technologies to be developed in project SECCRIT.

Social borders of assumed revelation protection in the sense of special relations between doctors and patients, between lawyers and clients or between family members can hardly be violated on the basis of a cloud-based video surveillance system. In contrast, violations of spatial and temporal borders could very well emanate from the establishment of such systems. At least in principle, images and recordings could be kept for an indefinite time and without being naturally bound to the place where they had been taken. However, strict legal regulations (see the respective deliverables D2.2, D2.4 and D2.7 in this regard) forbid to keep respective video surveillance data for longer than it is actually needed. Even if the risk of spatial and temporal borders being violated is thus basically heightened within highly connected and cloud-based video surveillance systems as opposed to traditional ones from the technological perspective, the strict regulatory framework of European data protection law prevents this risk from actually becoming relevant for the subjects of respective recordings. The foundational technologies developed in project SECCRIT, in turn, are explicitly designed to foster the enforcement of this regulatory framework through, e.g., technologically enforced deletion or geo-location policies (see the respective deliverables, particularly D3.2 - D3.4, D4.4, D5.1, D5.3).

3. Trust: Does the technique violate assumptions that are made about how personal information will be treated, such as no secret recordings?

To what extent recordings made by means of video surveillance systems are actually "secret" or in violation of the assumptions of those being monitored might be highly debatable. In many countries, legal regulations require video surveillance installations to be clearly announced with respective signs in order to make people aware of the fact that the specific area they are entering is monitored. This would speak against stating that the fact of recordings being made is actually secret. On the other hand, the qualitative properties of the surveillance system (are faces recognizable, can single persons be tracked while moving through a larger area, etc.) are often unclear to those being monitored and may thus very well raise contradictions between assumptions and factual givens. This is, however, not specific to *cloud-based* surveillance systems but rather would require broader socio-technical considerations on the acceptability of video surveillance of public spaces in general. Besides being not cloud- and thus not SECCRIT-specific, such considerations are also strongly bound to cultural expectations and values which strongly differ across Europe and can thus only be discussed and weighed locally, regarding the later practical application of SECCRIT-technologies. For the demonstrator, however, the risk of recorded persons not being aware of the surveillance system and thus of respective expectations being violated is strictly counteracted through the clear confinement to "recordings that are specifically produced for this purpose with explicit, well-informed and written consent of the persons being recorded" (SG3).

As opposed to traditional video-surveillance installation, highly interconnected cloud- and service-based surveillance systems do, however, introduce another risk regarding the violation of assumptions of information treatment: Basically, the cloud- and service-based model leads to the involvement of more actors in the overall data handling process and in a higher complexity of the overall system. This involvement of multiple parties (which can basically be assumed to be unknown by the recorded persons) and the concrete processes regarding the internal processing and use of respective data will hardly match the typical subjects' expectations. Furthermore, the respective systems and processes are typically too complex for communicating them to the surveillance subjects (e.g. via respective signs) in order to align expectations and actual givens. This, then, raises the dilemma that a technology which is clearly beneficial from the efficiency perspective, does, due to its complexity, necessarily raise unsolvable expectation-related conflicts. The only viable approach here is to consider expectations on a higher level of abstraction and refer to expectations like "recordings will only be accessed by trustworthy parties having a clear and proven need to do so, are always treated responsibly and are strictly protected against unreasonable disclosure", for example. The fulfilment of such expectations can, again, be facilitated by means of SECCRIT-technologies for real-world applications.

For the demonstrator, in turn, subjects to surveillance were provided with extensive explanations on the internal workings of the system and could directly ask respective questions before giving their well-informed and written consent. For the demonstrator only Mirasys employees which were well-informed were subject of the video data collection and gave their consent (the respective documents are reproduced in section 8). The risk of assumption-violations did thus, as opposed to later real-world applications, not apply to the demonstrator installation.

4. Personal relationships: Is the tactic applied in a personal or impersonal setting?

In the demonstrator as well as in later real-world applications, the cloud-based video surveillance system is clearly applied in non-personal relationships between data collector and subject with lower trust expectations than, say, within a family. For such impersonal relationships, Marx (1998) suggests more formal means of data collection governed by clear rules. Such rules, particularly covering ex-ante weightings of interests on the acceptability of certain areas being monitored or

on acceptable storage durations, are typically given by the different national frameworks regulating video surveillance based on the respective cultural givens etc. (see above). These formalizing rules do of course also apply to cloud-based surveillance systems, thus ensuring a certain formalization of the data collection. For the demonstrator, in turn, clear rules on the collection and handling of the recordings have been defined and communicated to the subjects before recordings were actually made. Last but not least, the Finnish national data protection officer was also involved from the very beginning, thus ensuring that the taken formal measures and safeguards actually meet national expectations and requirements.

5. Invalidity: Does the technique produce invalid results?

As laid out by Marx (1998), this question refers to whether the technical mechanism actually measures what it claims to measure. In particular, respective risks of unjustified treatment arise in the context of certain tests that might be invalidated by additional factors not taken into account of (a drug test producing invalid results because of a certain medicine being taken, for example). In the context of the cloud- and service-based video surveillance system, such “faulty” results could hardly be imagined in the context of taking the recordings themselves (this would, for example, require “a wrong person being shown” or “a person being shown in the wrong place” on a recording).

However, downstream processes of data processing, analysis and storage might very well give rise to a broad variety of faults leading to erroneous treatment of persons being wrongly identified as offenders, for example. Such invalid results could result from multiple factors including technical errors and, in particular, wrongly designed analytical algorithms. While technical errors can and have been worked against by technical means to a certain extent based on respective technologies and processes (including those developed in project SECCRIT), the risk of inappropriate results of algorithmic analyses is a severe threat with no established countermeasures beyond carefulness and diligence. Both have been applied in the project by experienced Mirasys engineers based on the long experience in image and video analysis. However, the increased technical complexity present within interconnected, cloud-based video surveillance systems as opposed to traditional ones might introduce a broad variety of additional sources for such errors that have not been foreseen and could thus not be addressed as a matter of principle.

A certain risk of newly introduced faults of automated analyses is thus unavoidable – even if not observed during the demonstrator-based evaluation. In such cases, however, it must at least be ensured that faulty results can be invalidated – be it by affected persons themselves or during official investigations. In this regard, SECCRIT technologies for root-cause analysis and audit trails will prove highly valuable. Nonetheless, they can clearly not eliminate the foundational risk of individual treatment being based on overly trusted but wrong algorithmic analyses as such. This risk is, however, omnipresent in modern societies strongly relying upon algorithms and is by far not restricted to cloud-based video surveillance systems. Solving this dilemma is thus beyond the scope explicitly taken herein.

6. Awareness: Are individuals aware that personal information is being collected, who seeks it, and why?

To a certain extent, this question is related to question 3 above, regarding the assumptions about information treatment and, in particular, the secrecy of recordings. As regards later real-world applications, the above statements on the sufficiency of signs announcing that a certain area is surveilled, on the involvement of multiple parties in cloud-based video surveillance systems (including the role of SECCRIT-technologies for countering potentially adverse implications arising therefrom) and on the importance of local cultural expectations thus apply here, too.

For the demonstrator evaluation, in turn, the strict confinement to recordings made on the basis of written and well-informed consent (the respective documents are reproduced in section 8) ensures that individuals are clearly aware of the information collection, the involved parties and the purpose of the recordings.

7. Consent: Do individuals consent to the data collection?

For video surveillance systems monitoring public spaces in practical use, individually given consent is not a viable approach for achieving legitimacy. Instead, it is common sense that for such systems, legitimacy must be based on a weighting of various individual and public interests, leading to a societal decision codified in respective legal legitimations and restrictions. For later real-world applications of the cloud- and service-based video surveillance system under consideration here, individual consent can thus not be achieved. However, well-established legal regulations exist in every member state and ensure that later real-world applications do not lack sufficient legitimation.

On the other hand, the demonstrator evaluation was realized with a strict confinement to individually given informed consent of the individuals being recorded. In this regard, see again the consent-related documents reproduced in Annex II (section 8)

8. Golden rule: Would those responsible for the surveillance (both the decision to apply it and its actual application) agree to be its subjects under the conditions in which they apply it to others?

For later real-world applications of the service- and cloud-based video surveillance system, this question can hardly be answered. However, it can be expected that the decision on its application would hardly differ from the decisions already made for applying current on-site video surveillance systems implementing the same fundamental functionality – as long as it technically does not introduce significant additional risks of data leakage, for example. The latter, in turn, is the primary objective of the technologies that have been developed in project SECCRIT.

Being those responsible for shaping the technological design, consortium members had intensive, repeated and controversial discussions at various consortium meetings and beyond on the acceptability of the cloud-based video surveillance system that has been developed. In the end, this led to the various safeguards established in D2.5 and laid down herein as well as to the concrete design of the cloud-based video surveillance system, including the respective technological measures. Given the design that was ultimately reached and the additional technical and non-technical safeguards, consensus was finally achieved throughout the whole consortium that being subject to such a system in a real-world setting would be acceptable. For the demonstrator-setting with pre-recorded video material produced in dedicated settings, in turn, acceptability was always undoubted throughout the whole consortium.

9. Minimization: Does a principle of minimization apply?

The principle of data minimization is highly prominent in the legal deliverables D2.2 and D2.7 and was already highlighted throughout the consultation with national data protection officers reported on in D2.4. Basically, the cloud-based video surveillance system does not implement a highly strict data minimization strategy (based on video-data pre-processing within the cameras itself, for example) in order not to render it impossible to benefit from the advantages of analyses being performed in the cloud-based, much more powerful and cost-efficient backend. Nonetheless, data minimization was particularly applied with regard to the storage duration, following the respective legally binding minimization obligations. In particular, no video data produced for the demonstrator evaluation was kept for longer than actually required for fulfilling the respective purpose. The respective legal obligations for data minimization of course also apply for any real-

world application of the developed cloud-based video surveillance system and again, the technologies developed in project SECCRIT particularly help in technically implementing and enforcing respective deletion policies etc. For more details on data minimization, see also, in particular, the respective sections of the legal deliverables D2.2 and D2.7.

10. Public decision-making: Was the decision to use a tactic arrived at through some public discussion and decision-making process?

For later real-world applications, this question can hardly be answered in advance. It can, however, be assumed that the deployment of a particular cloud-based video-surveillance installation will hardly be preceded by an explicit public consultation process. Rather, public (and political) decision-making processes in this field will refer to the general acceptability of video surveillance, differentiating on the basis of several criteria (e.g. the nature of the area being monitored, the technical and non-technical safeguards to be implemented, etc.). On such a generalized level, it can be assumed that any concrete deployment indirectly rests upon respective discussions and decision-making processes. A dedicated consultation process for any single real-world application site is, however, unrealistic.

For the demonstrator evaluation, in turn, a public consultation process was dispensable as only individuals who previously gave their well-informed consent and who voluntarily participated in the evaluation were actually subject to the recordings.

11. Human review: Is there human review of machine-generated results?

The underlying risk was already addressed in the context of question 5 on invalid results: Basically, the risk of inappropriate results of algorithmic analyses is always existing and cannot be completely eliminated as a matter of principle. On the other hand, the cloud-based video surveillance system does not lead to consequences for a person subject to wrongful analysis by itself. Instead, parties like security service providers are the ones ultimately acting, albeit supported by the video surveillance systems. For these parties, in turn, it must be ensured that they do not blindly act upon (potentially faulty) analyses of the system but solely take the results from automated analyses as “hints” or “indications”.

Ensuring such a deliberate understanding among later real-world users of the surveillance system is for large parts beyond the scope of project SECCRIT. Nonetheless, the results of automated analyses can be presented to the respective users in different ways, nudging them into different directions. For example, placing a security professional confirmation behind a result instead of a bare computer algorithm makes a significant difference in the resulting behavioural bias of respective users. This fact has been carefully taken into account in the design of the cloud-based video surveillance solution in order to prevent users from overly taking analysis results as facts and to encourage respective critical human reviews. For example, Surveillance system is configured to observe any movement on the certain limited area e.g. in front of a door. System can also be configured not to detect movements of dogs and cats (based on the height of the moving object). Let’s say that an object taller than 1 meter is moving in front of the door, an alarm is raised to the security guard. He will then open a specific camera feed from his monitor and see what is causing the alarm. He will then decide what caused the alarm and what kind of actions are needed, whether it was caused by a person having a cigarette outside or a burglar trying to break the lock in the door.

In the demonstrator evaluation, in turn, human review of the results was the constitutive goal by itself and no consequences would have resulted for the monitored individuals from faulty results. Additional measures were thus dispensable here.

12. Right of inspection: Are people aware of the findings and how they were created?

Making every subject aware of the results of analyses or even of the technological processes that led to them is hardly imaginable for typical settings of video surveillance of areas like train stations. Explicitly planning for or even developing respective measures to be applied in later real-world deployments of the developed system did thus hardly make sense within project SECCRIT. For activities regarding the demonstrator evaluation, however, participants were always offered to have a look at the results of the different processing steps that the recordings showing them were subject to as well as they were offered to see what the results of the respective steps were. 2 out of 2 participants made use of these offers. Reactions were positive, excited about the experience and satisfied with the recording quality.

13. Right to challenge and express a grievance: Are there procedures for challenging the results, or for entering alternative data or interpretations into the record?

The data subjects' rights to challenge and complaint as well as those for correcting or deleting/blocking inaccurate personal data are guaranteed in European data protection law (see the respective sections of D2.2 and D2.7 in this regard) and are thus to be provided for in any later real-world deployment of the video-surveillance system. In particular, these will in the concrete case apply to analysis results that can be linked to the respective data subject (e.g. by means of face recognition). Even if the strict obligations for deleting data as soon as it is not required anymore given by European data protection law (see above) should prevent much of such inaccurate data from being stored for longer durations, mechanisms were implemented in the developed system for manually deleting single datasets in case of analysis results actually being inaccurate. Individual persons cannot be deleted from videos but video material can be deleted manually any time. Video material will also be automatically deleted after a defined time interval. Technically time interval can be between 1 hour and 365 days. SECCRIT-specific technologies, in turn, were used for ensuring that such a deletion is also executed at the side of downstream actors like a cloud storage provider. For such mechanisms to actually make sense subjects must also be aware of the existence of such inaccurate analysis results. Different from other fields like online social networks, it is, however, impossible for individuals to request a synopsis of data stored about them in a video surveillance system based on credentials like name, etc. Together with the fact of no direct relation existing between the operator of the system (or, rather, the system itself) and the data subjects, this raises the question how subjects of the system should actually become aware of the existence of inappropriate analysis results. Obviously, this awareness automatically establishes in case a person is wrongfully treated as suspect by security personnel, for example. For these cases, the above-mentioned mechanisms for deletion are indispensable. Without such detrimental treatments, however, faulty results will be invisible to the respective subjects and thus hardly be challenged at all. This, in turn, is much more generic problem that cannot be adequately addressed within the scope of project SECCRIT.

Inaccurate original recordings will presumably be of less relevance in matters of rights to challenge, complaint, correction and deletion/blocking. Instead of "wrong" video content itself, errors can here most likely be expected with regard to wrong metadata attributions (date/time, area, used camera, etc.), wrongly suggesting persons to have been at a certain place at a certain time. Again, the system here also provides functionalities for rectification as well as for deleting respective datasets and SECCRIT-specific technologies support ensuring that rectifications or deletions are also propagated to downstream entities like the storage provider. Master server follows NTP (Network Time Protocol) i.e. it is very accurate because its time is synchronized within a few milliseconds of Coordinated Universal Time (UTC). Then, other servers are again synchronized with the Master server. There is very small probability for a need of rectification. In any case, all datasets are automatically deleted after predefined maximum retention time.For

both analysis results as well as original recordings, the formal procedures structuring a concrete process of challenge, complaint, rectification or deletion/blocking must be set up for the concrete case, particularly also depending on the specific regulatory givens to be applied. These concrete procedures are again beyond the scope of project SECCRIT.

In the evaluation demonstrator, the challenges regarding inaccurate data are less relevant: First, people are by no means "treated" on the basis of analysis results and second, the more direct relation between the subjects and the operators of the system together with the above-mentioned opportunity for subjects to take an inside-view of all process steps and intermediate results makes it more likely that inaccurate data is actually identified as such. Nonetheless, the above-mentioned functionalities for deletion and rectification were of course available here, too. There was, however, no single case where they were needed because of a participant calling for respective rectifications etc.

14. Redress and sanctions: If the individual has been treated unfairly and procedures violated, are there appropriate means of redress? Are there means for discovering violations and penalties to encourage responsible surveillant behaviour?

This question can for large parts only be answered for specific real-world applications under specific jurisdictions and subject to further, case-specific procedural rules. Regarding the means for discovering violations (and, not to forget, for retracing violations to their root cause, be it a technical or a human failure/misbehaviour), technologies for audit trails and root cause analysis developed within SECCRIT are nonetheless to be mentioned: Based on these, violation discovery and assignments can be made significantly easier and more reliable, thus clearly disencouraging irresponsible behaviour.

15. Adequate data stewardship and protection: Can the security of the data be adequately protected?

Yes. Basically, the whole project SECCRIT is dedicated to security in cloud computing in a multitude of facets. Listing all respective instruments and technologies here would hardly make sense – please thus refer to the various technical deliverables describing them in detail. In any case, the security of the data can unquestionably be adequately protected based on SECCRIT technologies. The fact of some of these not having been finalized at the beginning of the evaluation is relativized by the fact that only pre-recorded data of voluntary and well-informed participants were used in this phase.

16. Equality-inequality regarding availability and application: (a) Is the means widely available or restricted to only the most wealthy, powerful, or technologically sophisticated? (b) Within a setting is the tactic broadly applied to all people or only to those less powerful or unable to resist? (c) If there are means of resisting the provision of personal information are these means equally available, or restricted to the most privileged?

This is another question of only limited relevance for the case of video surveillance. Given the clear focus of later real-world applications to public spaces like railway stations, aspects of individual wealth, power or technological sophistication become irrelevant. Similarly, the technology is not explicitly "applied to" specific (groups of) individuals but rather to a certain area, independently from who is currently present in this area. Means of resisting the provision of personal data is limited to covering faces from cameras and comparable approaches, which basically are similarly available to all individuals. One might, of course, construct a certain kind of discrimination based on arguments like "the wealthy can go by car while the poor are forced to enter railway stations", but this would have been too far-fetched to be actually covered within project SECCRIT.

The same is of course true for the participation in the demonstrator. If we had offered substantial allowances for participating in the study, this would probably have been an inappropriate incentive especially for underprivileged persons to participate and thus to expose themselves to a certain extent. But given that only Mirasys employees during their working hour participated in the demonstrator such an argument would seem quite far-fetched.

17. The symbolic meaning of a method: What does the use of a method communicate more generally?

Like every approach of surveilling public areas, the cloud- and service-based video surveillance system under consideration here communicates a perceived need for producing video coverage to be given within the monitored area because of a perceived level of insecurity – and particularly a perceived risk of becoming victim to a crime – within that area. The more visible cameras such a system involves, the more might this lead to people feeling unsafe within the respective area just because of the many cameras that seem to be necessary for making it a sufficiently safe place. Furthermore, it does – again like any surveillance system – put everybody entering the monitored area under general suspicion of probably being an offender. Last but not least, any video surveillance system including the one under consideration here raises the feeling that any behaviour *could* at least be seen and recognized by the surveillant. Following the “panopticon effect” as described in detail in section 2.1 and also referred to for question 1, this could lead to (societally undesirable) conformism.

All these effects are, however, not specific to cloud-based video surveillance systems. Meanings specifically communicated by such systems are, however, hardly identifiable.

18. The creation of unwanted precedents: Is it likely to create precedents that will lead to its application in undesirable ways?

Like many ethically relevant information technologies, the service- and cloud-based video surveillance system considered herein could, in principle, easily be employed by certain players (e.g. authoritarian regimes) in unquestionably undesirable ways, e.g. for suppressing rebellions or demonstrations or for chilling oppositional movements in general. Even if this is unquestionably the case for any video surveillance technology, the service-based model might here weaken respective hindrances erected by, for example, export restrictions. As the service-based model lessens the need for local installations to be performed by experts and might rather allow respective actors to erect comprehensive and powerful surveillance infrastructures by themselves, it becomes crucial to ensure that the service-based video surveillance system is only used by undoubtedly eligible parties. In an alleviated form, the same arguments of course also apply to other players like corporations (or even single members of, for example, corporate security departments), which could because of the simplified service-based model and the low obstacles to entry be tempted to “play around” with the powerful instrument of video surveillance without proper legitimation, safeguards etc. In order to prevent such inappropriate applications, which could unquestionably also redound upon the provider of the system itself, Mirasys will establish strong restrictions on the use of its system once bringing it on the market. In particular, this will include pre-validations of the appropriateness of a certain planned deployment (including the eligibility of the operator demanding access) and regularly monitor the actual use of the system (e.g. to prevent an eligible deployment from being expanded into an unacceptable one). In any case, the common self-service property of cloud computing is largely inappropriate for the application field covered here.

With regard to the foundational security technologies, techniques and methodologies developed within SECCRIT, the creation of unwanted precedents or undesirable (fields of) application can hardly be influenced at all. Without going more into the details of the long history of debates around export restrictions on cryptographical products, intentional weakening of security

technologies, etc., we confine ourselves to the statement that actually restricting or influencing the concrete application of foundational and general purpose technologies is hardly viable at all. Considerations in this regard were thus not subject to project SECCRIT.

19. Negative effects on surveillants and third parties: Are there negative effects on those beyond the subject and, if so, can they be adequately mediated?

Negative effects on surveillants as well as on third parties are not to be expected for the cloud- and service-based video surveillance system – neither for later real-world applications nor for the evaluation demonstrator.

20. Beneficiary: Does application of the tactic serve broad community goals, the goals of the object of surveillance, or the personal goals of the data collector?

Again, this question does for the cloud-based video surveillance system raise much broader questions on the beneficiaries of surveilling certain areas. Generally speaking, these considerations do in most cases boil down to a weighting between collective interests (e.g., the society of a city being interested in lowering and solving crime, all employees of a firm being interested in preventing unauthorized access to an office building, etc.) and individual privacy invasions. In most cases of video surveillance, the objects of surveillance are thus also part of the relevant community and, sometimes, of the data collector. Established models exist for weighting these interests against each other (including political law-making, mechanisms of corporate governance, etc.) but in general, the specific properties of cloud- and service-based video surveillance systems as opposed to traditional ones play a minor role in these regards. Following our usual triangulation approach, more in-depth considerations are thus obsolete here.

21. Proportionality: Is there an appropriate balance between the importance of the goal and the cost of the means?

The answer is clearly yes. From a cost-perspective, a cloud-based video surveillance system is significantly advantageous over a traditional, otherwise equivalent system as necessary up-front investments are significantly reduced and as running expenses are always limited to the actual requirements, particularly diminishing the costs arising from traditionally necessary overprovisioning etc. As compared to traditional ones, a cloud- and service-based video surveillance system thus exhibits the same cost advantages that led to the broad establishment of cloud computing as opposed to traditional in-house data centres in general. Assuming that an appropriate balance between goal importance and costs is already given for traditional video surveillance installations, it must thus even more be given for cloud- and service-based ones.

22. Alternative means: Are other, less costly means available?

Following the same argumentation as given for question 21 above, the answer must be no. Of course, it could very well be discussed whether the cost-benefit-ratio is positive for established video-surveillance installations in general (implying that video surveillance might also be in use for serving other goals, including individual ones of respective decision makers), but such a discussion would again go well beyond the scope of project SECCRIT in general.

23. Consequences of inaction: Where the means are very costly, what are the consequences of taking no surveillance action?

Again, this question is beyond the scope as it is directed the question whether video surveillance should be used at all while we herein explicitly confine ourselves to the differences between traditional and cloud- and service-based video surveillance systems.

24. Protections: Are adequate steps taken to minimize costs and risk?

With regards to the costs, this question must mainly be asked for individual deployments and can hardly be answered for the development of the fundamental system itself. However, Mirasys, the partner implementing the cloud- and service-based video surveillance system, has long experience and a very own interest in preventing the development costs from overly expanding. The same is true for typical development risks. Other risks principally include legal and ethical/societal risks (particularly in matters of privacy/data protection and evidence/liability law), which might hamper the later introduction to the market if not properly taken into account from the very beginning. These risks were thoroughly addressed throughout the whole project, beginning with the establishment of legal fundamentals (D2.2 in M6), consultations with national data protection authorities (D2.4 in M12) and the initial ethics report (D2.5, also in M12) and followed by constant activities of providing techno-legal and ethics-related guidance throughout the whole project duration as reported on in D2.7 (M36) as well as herein (M36). For the evaluation demonstrator, in turn, no significant costs and other risks were identified as having to be separately addressed in addition to the aforementioned general measures.

25. Appropriate vs. inappropriate goals: Are the goals of the data collection legitimate?

Again, this question would refer to the appropriateness of the goals behind using video-surveillance in general. This question is thus again beyond the scope and we assume the legitimacy to be given for the general application of video surveillance in certain areas like railways stations because of their crucial infrastructural role. Nonetheless, it must be mentioned that the general legitimacy of the goals pursued by means of video surveillance (e.g. preventing and solving crime) within a certain area can only be evaluated on the basis of respective legal regulations and the specific cultural givens applicable to a specific case.

26. The goodness of fit between the means and the goal: Is there a clear link between the information collected and the goal sought?

Another question rather directed at the use of video surveillance in general. We assume the goodness of fit as given. Otherwise, this would imply that there is no clear link between doing video surveillance and the intended reduction and solving of crimes. If this were actually the case (and there are several scientific voices actually claiming it), this would require to initiate a general discussion on video surveillance as such. This, again, would be well beyond the scope of SECCRIT, where video surveillance is just one demonstrator field for applying the developed technologies, techniques and instruments aimed at secure cloud computing in general.

27. Information used for original vs. other unrelated purposes: Is the personal information used for the reasons offered for its collection and for which consent may have been given, and do the data stay with the original collector, or do they migrate elsewhere?

In the video surveillance demonstrator, recorded video data is only used for the purpose of evaluation and demonstration and not for any other function. Due to the very nature of SECCRIT as a cloud-related project, collected data do, however, not stay with the original collector (Mirasys as the partner erecting and operating the demonstrator) but are rather transferred to an isolated cloud environment hosted by consortium member Amaris. Beyond this transfer, recordings are not transferred to other parties. Furthermore, no involved party uses recordings (or other relevant data derived from them) for other purposes beyond evaluation and demonstration.

28. Failure to share secondary gains from the information: Is the personal data collected used for profit without permission from, or benefit to, the person who provided it?

As outlined above, recordings will only be used for the purposes of evaluation and demonstration. It might be argued that the insights gained from these activities lead to better products being

created by Mirasys as well as by Amaris (clearly an intention of the EU funding application-related research) and that the respective profit resulting from such improved products would classify as "secondary gain" that should somehow be compensated. Nonetheless, we feel that such an interpretation would not match the original intention of the compass question, especially as only employees participated in the demonstrator.

29. Unfair disadvantage: Is the information used in such a way as to cause unwarranted harm or disadvantage to its subject?

The answer here is clearly not. As the recordings are not used for any purpose beyond project-internal evaluation and demonstration, there is no relevant risk of recorded persons experiencing any disadvantageous treatment based on the data.

5.1.2 Application and Addressing of further Safeguards

Besides the ethical appraisal based on the "compass questions", additional safeguards for ensuring ethically acceptable conduct of responsible research and innovation during the implementation and testing of the video surveillance demonstrator were also specified in section 4.1. For each of these safeguards, it shall also be documented that and how it was implemented.

SG2 – No surveillance of private or public spaces: In order to prevent privacy invasions, no evaluation activities will be conducted based on the surveillance of (semi-) public or even private spaces.

As repeatedly stressed above, evaluation of the video surveillance demonstrator was exclusively conducted within a dedicated environment ensuring that no person is recorded without previously given, well-informed and written consent. For this purpose, it was selected a quiet time of the day when there were no trains arriving or departing from the station and therefore low risk to have unintended people to walk around in the video. It was also selected a quiet area at the very end of the station. The video manuscript was planned to be very short and to the point enabling to complete the recording between trains (that run every 10 minutes).

This setting ensures that no people are unwillingly or accidentally recorded and thus does not represent a "public space" in the sense of the safeguard. Furthermore, the setting is also everything but a private one, not raising any concerns of intimate information being revealed, for example.

SG3 – Experimental environments: Evaluations and tests of the secure cloud technologies that are to be developed in the project will, with regard to video surveillance, only be conducted on the basis of experimental environments that are explicitly set up for the project.

The above description of the experimental setting for recording sample data within the video surveillance demonstrator also illustrates how this safeguard was taken into account during the evaluation and demonstration phase. The environment was set-up exclusively for the purpose of evaluation and demonstration within project SECCRIT and no further data (like, for instance, other, already existing recordings from other places) were used during evaluation and demonstration.

SG4 – Explicit Consent (Video): Evaluations and tests will, with regard to video surveillance, only be conducted on the basis of recordings that are specifically produced for this purpose with explicit, well-informed and written consent of the persons being recorded.

This safeguard is strongly in line with legal requirements and the suggestions given by the Finnish data protection authority. In order to comply with this self-given safeguard and with the various regulatory givens, Mirasys – as the consortium member operating the demonstrator – recruited voluntary participant from their own pool of employees. All these were familiar with the inner-workings of the system. Mirasys prepared a document explaining all necessary facts in a laymen-understandable language and offering participants to ask for further details. Beyond providing them with the necessary information for making well-informed decisions, this document was also used for collecting the explicit and written consent from participants before producing recordings of them in the experimental environment mentioned above. The document is reproduced in section 8.

SG5 – Controlled cloud: Evaluations will only be conducted within well-controlled cloud environments separated from the public Internet through strong state-of-the-art security mechanisms and hosted within the European Union.

Any employment of cloud computing within the video surveillance demonstrator was strictly confined to the controlled cloud environment exclusively provided for project activities by consortium member Amaris. This controlled environment consisted of VMware virtual servers running on physical Cisco Servers, secured by physical Cisco Firewalls and was separated from the rest of the Amaris-cloud through virtual networks and dedicated Servers. Separation from the public internet, in turn, was realized through virtual Firewalls for each Customer-Demo virtual Environment. There was no real cameras used transferring data over the internet but instead, camera was emulated and video clip was locally saved in the Amaris cloud. The whole environment was physically hosted in Amaris' data centre in Vienna Austria.
In the video demonstrator, this separated cloud-environment was used employed for testing the functionality of various technical outputs. Testing reports are available as D6.2 and D6.3.

Altogether, the safeguards were thus applied during the establishment of the video surveillance demonstrator as originally intended. The fact that recalling the safeguards as well as the compass questions during definition, implementation and use of the demonstrator did at several points of the process lead to fruitful discussions among project partners and to several re-alignments of respective planning and technological designs repeatedly reconfirmed the value of defining and agreeing upon such self-given constraints at early stages of development. Having them in place prevented the project from unintentionally overlooking relevant ethical factors especially in more hectic phases of development and implementation and fostered the establishment of an ethically acceptable implementation of the video surveillance demonstrator.

5.2 Application in Traffic Management Demonstrator

The traffic management demonstrator, which is outlined in more detail in deliverable 6.1, serves the goal of evaluating and demonstrating the capabilities of SECCRIT-technologies for supporting the migration of data-intensive traffic control systems to the cloud while still fulfilling the specific requirements of such systems in matters of responsiveness, reliability and security against external attacks. As traffic control systems like the one used for the demonstrator are subject to strong workload variance depending on the time of the day as well as on seasonal changes, migrating such systems to the cloud promises significant efficiency gains as long as all previously mentioned requirements are met. Given the sheer mass of data from various kinds of sensors being fed into the traffic management system under consideration, however, it had to be ensured

that ethical values are not – intentionally or unintentionally – violated. In the following, it will be described how this was achieved during the design and implementation of the demonstrator.

Architecturally, the demonstrator consists of several building blocks which are described in more detail in section 3.2 of deliverable 6.1. In particular, Valencia's actual traffic management system comprises components for data collection / sensing (including video footage from CCTV cameras), data management, data analysis, traffic control, and for various functions of internal system management and maintenance.

Of these, all functions were migrated to the cloud except – different from the initial planning – those parts of the original system that cover the collection of real traffic data and feeding those data into downstream processes as well as those that actually "interact" with the real traffic (traffic light control etc.). Instead, simulated data was used throughout the whole process of designing, establishing and evaluating the demonstrator in order to prevent ethical conflicts and the "interaction" with outside traffic was also simulated. For the same reason, all functions related to video data were (even if strongly used within Valencia's actual traffic control system) explicitly excluded from being subject to the demonstrator, too. With regard to the collected "real-world data" in the traffic management demonstrator, the SECCRIT project thus prevented ethical conflicts to a degree that goes well beyond mere legal compliance, given that the Spanish data protection authority legitimated even the collection of non-obfuscated plate numbers (See section 2.1.2, block I of deliverable 2.4). Even if this exclusion of video-related components and the use of simulated data as well as of simulated "interaction" components effectively avoid most ethics-related interferences from the ground up, Marx' compass questions as well as the safeguards initially defined in deliverable 2.5 shall be applied to this demonstrator, too.

5.2.1 Application and Addressing of Compass Questions

Like for the video surveillance demonstrator, the application of the "compass questions" reproduced in section 7 is also the first and most important safeguard (SG1) to be applied for the traffic management demonstrator. In so doing, we try to prevent overly repetitions of arguments and considerations already presented in section 5.1 and to concentrate on those aspects that clearly differentiate the traffic management demonstrator from the video surveillance demonstrator.

1. Harm: Does the technique cause unwarranted physical or psychological harm?

Due to the non-existence of interactions between the demonstrator and the real traffic flow, the risk of physical harm can practically be ignored for the demonstrator. For later real-life applications of the developed technologies within cloudified traffic management systems, one might argue that there is a higher risk of the traffic management system being successfully attacked and used for disturbing the real traffic flow of complete cities, thus causing a broad range of potential harms, ranging from increased stress to ambulances not reaching hospitals in time. Unquestionably, this would be a risk of high ethical relevance as soon as the risk of attack would actually be higher for a cloudified environment than for isolated, locally hosted systems as currently used. The technologies that are developed in SECCRIT do, however, just serve the goal of minimizing the additional risks that would emanate from a respective cloudification and are thus just directed at minimizing the above mentioned risks of actual harm to happen. In any case, actual cloudification of traffic management systems require respective security technologies to be sufficiently mature to not cause additional risks but rather benefit from the positive characteristics of cloud-computing in matters of, for instance, availability, scalability, etc.

2. Boundary: Does the technique cross a personal boundary without permission (whether involving coercion or deception or a body, relational, or spatial border)?

A significant crossing of a personal– natural, social or relational – boundary cannot be recognized, neither for the demonstrator employing only emulated data nor for later real-world-applications of cloudified traffic management systems with the functionality covered by the current one of the city of Valencia. Only with regard to temporal and spatial borders as soon as sensor- or video- data were kept for longer times in a form allowing individuals to be identified. In this regard, however, the same arguments already discussed for the video surveillance demonstrator above do apply again. Strict legal regulations forbid practices that would make such boundary-crossing relevant and SECCRIT technologies clearly support the enforcement of these regulations.

3. Trust: Does the technique violate assumptions that are made about how personal information will be treated, such as no secret recordings?

As neither video data nor real sensor data and thus no personal data are used, the answer is clearly “no” for the demonstrator. For later real-world applications, however, highly relevant ethical threats might arise in this regard as soon as individualized, person-related data are collected and used within the system in the form of video-data with number plate recognition or in the form of non-anonymized data originating from Car-2-X systems and including individual identifiers of cars etc. The collection and processing of such data would presumably happen without the clear knowledge of any traffic participant being guaranteed, thus violating assumptions about what personal data is collected and how it is processed “in the cloud”. Even if not relevant for the demonstrator because of the ethically motivated exclusion of such data, safeguards 6 (on obviating identification-prone visual information like faces and number-plates) and 7 (on anonymizing an pseudonymizing employed Car-2-X data) are thus highly relevant for later real-world applications of the foundational model employed in the traffic management demonstrator.

4. Personal relationships: Is the tactic applied in a personal or impersonal setting?

The setting that is subject to a traffic management system is primarily impersonal in the demonstrator as well as for foreseeable later real-world applications.

5. Invalidity: Does the technique produce invalid results?

An invalid result in the sense of Marx (1998) would be an unjustified treatment experienced by an individual because of wrong or misinterpreted measurements. In the demonstrator, the strict avoidance of any real-world-interference prevents such unjust treatments from happening at all. In later real-world-applications, however, invalid treatments could very well be expected as soon as the traffic management system influences single traffic participants individually, e.g. based on (wrongly) measured driving speed. The risk of such invalid treatments basically arises for any such traffic management system with individualized responses and shall thus not be discussed in more detail. In “cloudified” traffic management systems, however, there might be specific risks in this regard resulting from data being (intentionally or unintentionally) altered and from missing possibilities for proving the integrity of such data. The technologies developed in SECCRIT are explicitly aimed at reducing these drawbacks of cloud computing and thus help to prevent such mistreatments from actually happening.

6. Awareness: Are individuals aware that personal information is being collected, who seeks it, and why?

See question 3 above. In the demonstrator, no personal information is collected. In later real-world-applications, this might however very well be the case and ensuring every traffic

participant's awareness of video and sensor data being collected might be unreachable. Nonetheless, any introduction of a traffic management system collecting such data should be accompanied by measures (publicity campaigns, etc.) informing as many traffic participants as possible about the collection and about how that data is handled and used. This is, however, independent from whether the system is hosted locally or in the cloud and shall thus not be subject to further discussions here.

7. Consent: Do individuals consent to the data collection?

Again, no personal data is collected in the demonstrator, thus making the question for consent irrelevant in this regard. For later real-world-applications, however, it is hardly imaginable that every traffic-participant explicitly expresses his/her consent in advance. Like for the case of video surveillance, legitimacy of data collection must thus be ensured on other grounds like regulatory legitimations. These, in turn, have to ensure a just weighing of individual and public interests and a multitude of further measures, particularly including the concept of data minimization discussed in more detail in the legal deliverables.

8. Golden rule: Would those responsible for the surveillance (both the decision to apply it and its actual application) agree to be its subjects under the conditions in which they apply it to others?

Due to decision not to use real-world data during the project's runtime, this question is obsolete for the traffic management demonstrator. With regard to later real-world-applications, however, consortium members (which presumably resemble later responsible parties quite well) quickly agreed that being subject to a traffic management system of the kind described here is acceptable – irrespectively from the question whether it runs locally or in the cloud – as soon as the additional safeguards proposed herein (anonymization, obviation of faces and number plates, technological means for securing the cloud environment) are applied.

9. Minimization: Does a principle of minimization apply?

In the demonstrator, the use of emulated data perfectly applies the principle of minimization – no real personal data are used at all. For real-world applications, in turn, keeping the principle of minimization in mind is highly important. This again refers to anonymizing any involved Car-2-X data as well as to the obviation of number plates and faces from video data.

10. Public decision-making: Was the decision to use a tactic arrived at through some public discussion and decision-making process?

Like for the cloud-based video-surveillance system discussed in section 5.1.1, this question can hardly be answered in advance for later real-world applications of a cloud-based traffic management system and it can be expected that the process of public decision-making will rather refer to the establishment of advanced traffic management systems in general. For the demonstrator, however, a public consultation was dispensable.

11. Human review: Is there human review of machine-generated results?

As discussed for question 5, the underlying risks only arise in the case of traffic participants being treated individually within the traffic management system which is, at least for the moment, not intended. More importantly, only a small fraction of the underlying risks motivating human review of analysis results etc. arises from the traffic management system being "cloudified" whilst most of them appear independently from the deployment model (local vs. cloud-based). In any case, cloud-specific risks, which could for instance result from intentional or unintentional data changes regarding measured driving speeds, will also be subject to the same review procedures that apply to non-cloudified traffic management systems in later real-world applications. In the demonstrator,

in turn, no individuals are actually affected and the results will nonetheless be subject to intensive human review during the evaluation process.

12. Right of inspection: Are people aware of the findings and how they were created?

Like for the case of video surveillance discussed above, making every subject aware of the results of analyses or even of the technological processes that led to them is hardly imaginable for typical large-scale deployments of traffic management systems. On might, however, think about mechanisms for informing the public about the inner workings of such systems in general – independently from the deployment model. In the demonstrator, however, inspections were dispensable because no real data were used.

13. Right to challenge and express a grievance: Are there procedures for challenging the results, or for entering alternative data or interpretations into the record?

Respective rights etc. are only required if individualized datasets are actually stored for longer times in the respective system. At least for the functionality currently foreseen, we don't see a need for such a functionality within traffic management systems as no individualized data is to be stored in personalized form. For those parts of later real-world traffic management systems incorporating video data, the remarks given on this question in section 5.1.1 apply here, too.

14. Redress and sanctions: If the individual has been treated unfairly and procedures violated, are there appropriate means of redress? Are there means for discovering violations and penalties to encourage responsible surveillant behaviour?

Like for video surveillance installations, this question can only be answered for specific real-world applications of traffic management systems under specific jurisdictions and subject to further, case-specific procedural rules. Discovering and retracing them to their (technical or human) root cause, in turn, is supported by SECCRIT technologies for audit trails and root cause analysis which thus clearly discourage irresponsible behaviour. For the demonstrator, questions of redress etc. are not relevant as no real individuals are affected.

15. Adequate data stewardship and protection: Can the security of the data be adequately protected?

See again the respective question in section 5.1.1 – The whole project SECCRIT is dedicated to security in cloud computing in a multitude of facets. Listing all respective instruments and technologies here would hardly make sense but in any case, the security of the data can unquestionably be adequately protected based on SECCRIT technologies.

16. Equality-inequality regarding availability and application: (a) Is the means widely available or restricted to only the most wealthy, powerful, or technologically sophisticated? (b) Within a setting is the tactic broadly applied to all people or only to those less powerful or unable to resist? (c) If there are means of resisting the provision of personal information are these means equally available, or restricted to the most privileged?

This question is only of limited relevance for the case of a cloud-based traffic management system. Basically, such a system is applied to any traffic participant in the covered area, irrespectively of wealth, power, etc. It is not expected that mechanisms for anonymization and obviation will be applied on an individualized basis – respective means are thus always similarly available to all affected individuals. In matters of availability of the system itself, discussing inequalities between rich and poor cities, for example, would hardly make sense.

17. The symbolic meaning of a method: What does the use of a method communicate more generally?

Like discussed for the video surveillance system in section 5.1.1, the establishment of all-encompassing traffic-management-systems could produce a feeling of being under constant supervision for individual traffic participants and thus lead to a “panopticon” effect. This is, however, not specific to cloud-based traffic management systems and shall thus not be discussed in more detail.

18. The creation of unwanted precedents: Is it likely to create precedents that will lead to its application in undesirable ways?

Of course, any traffic management system of the kind considered herein could be misused for surveillance and tracking purposes – especially in the case of video cameras and individualized data from Car-2-X systems being incorporated. With a particular focus on cloud-related aspects, it could be expected that sophisticated methods of analysing traffic data can more easily creep into other application domains or become available to other users than it is the case for locally hosted traffic management systems (see the respective discussion for cloud-based video-surveillance systems in section 5.1.1). The city of Valencia will strongly keep this in mind in case it considers to provide the developed technologies to other players.

19. Negative effects on surveillants and third parties: Are there negative effects on those beyond the subject and, if so, can they be adequately mediated?

Negative effects on surveillants as well as on third parties are not to be expected for the cloud-based traffic management system – neither for later real-world applications nor for the evaluation demonstrator.

20. Beneficiary: Does application of the tactic serve broad community goals, the goals of the object of surveillance, or the personal goals of the data collector?

The traffic management system in general serves a broad variety of community goals (including shorter travel times, less congestion, less noise, better air quality, etc.) which are laid out in more detail in deliverable 6.1. In most cases, these will coincide with the goals of individual traffic participants who will – even if not in personalized form – be the subjects of data gathering in later real-world applications. Further goals of the data collector (the municipality rolling out the traffic management system) are not significant – at least as long as the traffic management system is not used for other purposes like controlling driving speeds etc. The benefit of “cloudifying” the traffic management system, in turn, will primarily be monetary ones (serving the municipality and, indirectly, all it’s taxpayers) and goals related to service quality (availability etc.) which primarily benefit individual traffic participants and the broad community in general.

21. Proportionality: Is there an appropriate balance between the importance of the goal and the cost of the means?

The proportionality of the relation between the costs for the traffic management system and its goals in general do not need to be discussed here. The proportionality between costs and benefits of migrating such a system to the cloud, however, can easily be assumed to be given as the main motivation is just reducing the costs of maintaining sufficient local hardware and ensuring an adequate level of availability etc.

22. Alternative means: Are other, less costly means available?

The only viable alternative of migrating a traffic control system to the cloud is keeping it local. As it is the foundational assumption of evaluating the cloudification of the traffic control system is that doing so will reduce costs, the answer to this question is no.

23. Consequences of inaction: Where the means are very costly, what are the consequences of taking no surveillance action?

Taking no "surveillance" action at all would in this case mean not to establish a traffic control system at all. This would result in a variety of drawbacks ranging from longer travel times to lower air quality (see again deliverable 6.1 in this regard). With a focus on the cloudification of the already existing system, the consequences would primarily be higher overall costs and efforts as well as lower capabilities for serving situations of extremely high system load.

24. Protections: Are adequate steps taken to minimize costs and risk?

With regard to cost and risk minimization, the same arguments given for the same question in section 5.1.1 apply here, too: The municipality of Valencia is quite aware of expectable cost structures and a multitude of activities (including extensive initial and constant legal assessment, external consultations etc.) have been pursued throughout the project in order to minimize risks of the ultimate solution not being applicable in practice.

25. Appropriate vs. inappropriate goals: Are the goals of the data collection legitimate?

The goals pursued with the traffic management system in general (shorter travel time, better air quality, …) appear to be perfectly appropriate from the perspective of a whole municipality as well as from the perspective of the individual traffic participant. The goals behind cloudifying this traffic management system (reduced costs, better availability/reliability, better ability to handle peaks, etc.), in turn, also seem perfectly appropriate – in particular from the perspective of the municipality and its taxpayers.

26. The goodness of fit between the means and the goal: Is there a clear link between the information collected and the goal sought?

Another question that can only be answered with regard to the traffic management system in general. In this context, the information collected from various kinds of sensors is clearly necessary for achieving the required overview about the current traffic situation and for managing traffic upon this basis. With regard to the integration of video data, however, the answer is less clear. Even if not relevant for the demonstrator, the appropriateness should be reassessed based on the intended purpose of this integration before any actual real-world application. Should later revisions of the – possibly cloudified – traffic management system furthermore be expanded by the collection of potentially individualized Car-2-X data, this question is to be critically reassessed. For the sake of completeness, the link between cloudification and the goal of lowering costs and heightening availability, reliability, scalability etc. is more than clear.

27. Information used for original vs. other unrelated purposes: Is the personal information used for the reasons offered for its collection and for which consent may have been given, and do the data stay with the original collector, or do they migrate elsewhere?

In the demonstrator, no personal information is collected. In later real-world applications, however, personal information might be (at least initially) collected in the form of video footage or Car-2-X data carrying individual identifiers. There is, of course, the foundational risk of these data being used for other purposes than traffic management (for issuing speeding tickets, for example) before being anonymized / obviated. In practice, however, such infringements of the principle of purpose limitation are usually prevented (or rather limited to certain legitimate cases) by strict regulatory constraints. In the later real-world applications of the cloudified traffic management system, collected data would not stay with the original collector but rather necessarily transferred to the respective cloud provider. In the best case, this would only happen in anonymized and obviated form. Should this not be possible in later real-world applications, SECCRIT-technologies

could at least help ensuring that data is handled in accordance with regulatory constraints in the cloud.

28. Failure to share secondary gains from the information: Is the personal data collected used for profit without permission from, or benefit to, the person who provided it?

While in the demonstrator no personal data is collected at all and while later real-world applications of a cloud-based traffic management system should for large parts only use anonymized data, even the collection and use of data in personalized form should be the absolute exception. For cases where such personalized data actually are indispensable, however, secondary uses will hardly appear reasonable. Any later real-world application should be critically assessed in this regard. This can, however, not be done in advance herein.

29. Unfair disadvantage: Is the information used in such a way as to cause unwarranted harm or disadvantage to its subject?

Using data originally collected for the purpose of traffic management for issuing speeding tickets, for example, would pose such an unfair disadvantage to the traffic participant. This would, however, be a case of unreasonable purpose expansion. Other categories of unfair disadvantages – especially those emanating from the specific givens of a cloud-based system as opposed to a traditional, locally hosted one, were not identified.

5.2.2 Application and Addressing of further Safeguards

In addition to the ethical aspects already addressed through the compass-questions, further ethical safeguards have also been specified for the traffic management demonstrator in section 4.2. As done for the video surveillance demonstrator, it shall also be documented how these were regarded in the second demonstrator's design and implementation:

SG5 – Controlled cloud: Evaluations will only be conducted within well-controlled cloud environments separated from the public Internet through strong state-of-the-art security mechanisms and hosted within the European Union.

This safeguard was applied to the traffic management demonstrator in the same way it was for the video surveillance demonstrator. All respective explanations given in section 5.1.2 do thus apply here, too. In the specific case of the city of Valencia demonstration, all virtual machines have been placed in three interconnected networks:

(1) one belonging to ETRA in Spain, which is hosting the user interfaces and simulation tools to emulate all sensors of the city of Valencia, hence it is not connected to the city network in any sense,

(2) the private cloud environment offered by AMARIS in Austria,

and (3) the OTE Lab environment in Greece.

In all three, we are using state of the art technologies for security, as e.g. communications through VPN. Furthermore, SECCRIT beyond state of the art security tools are deployed and running.

SG6 – Obviate identification: It is technically ensured that faces, number plates and other features possibly allowing the identification and tracking of individuals are made unrecognizable whenever visual data (video, still images) is used in the traffic control demonstrator – be it through blurring or through a respective reduction of image resolution.

Aspects of the traffic management system employing CCTV were (even if legitimated by the Spanish data protection authority, see deliverable 2.4, section 2.1.2) in the course of the project explicitly excluded from being subject to the demonstrator in D6.1, section 3.1. Without video recordings and still images being employed in the demonstrator, the need for obviating faces, number plates etc. from such videos and images does of course disappear and the safeguard is automatically fulfilled.

SG7 – Anonymization/Pseudonymization: Whenever advanced technologies of Car-2-X communication are to be integrated into the traffic control demonstrator, the respective data will only be used in anonymized or pseudonymized form, preventing any attribution of data to individuals.

Even if this was not clear at the beginning of the project, Car-2-X data was not employed in the demonstrator at all and the whole demonstrator was realized on the basis of emulated data. This safeguard is thus automatically fulfilled, too.

SG8 – Explicit Consent (C2X): Any potential integration of data from Car-2-X communication and similar systems is only done if the respective individuals have given their (well-informed and written) consent in advance.

For the same reasons as SG 7, SG8 is also fulfilled automatically.

5.3 Application to SECCRIT-Technologies in General

Besides the ethical aspects and safeguards that were applied to the SECCRIT specific demonstrators above, some foundational characteristics of cloud computing exhibiting particular ethical relevance were also identified in deliverable 2.5. In the following, we will pick up these characteristics and explain to what extent SECCRIT technologies may contribute to countervailing respective ethical drawbacks of cloud computing. Furthermore, we will identify potential starting points – beyond those already identified in the legal deliverable 2.7 – for discussing the further development of the regulatory framework governing the practical application of cloud computing.

The repeatedly stressed “triangulation approach” of primarily focusing on “cloudified” video surveillance and traffic management notwithstanding, we identified a general “control loss”, a need for mechanisms that heighten transparency, foundational issues of data transfer across cultural and legislative boundaries and indirect effects arising from an inherent tendency towards monopolies and lock-ins as aspects of particular ethical relevance in deliverable 2.5. These issues are so discussed in the following parts.

5.3.1 General “Control Loss” induced by Cloud Computing

Control loss directly results from the foundational principle of outsourcing computing resources that is the heart of cloud computing. Data and processes are under the primary control of the cloud provider and can typically only be controlled via well-defined interfaces by the cloud user while the cloud provider is at least basically able to take notice of or even to influence/alter them. Furthermore, the cloud provider ultimately determines the concrete technical details of the environment while the cloud user has only limited possibilities for exerting influence in this regard.

Especially in settings involving multiple, interconnected cloud providers, this leads to the apparent problem of actual responsibility for the ultimate outcome being vaguely distributed across a multitude of involved parties.

As already identified in deliverable 2.5, this leads to doubts in matters of fairness and appropriateness regarding the current legal situation of the cloud user being the only party responsible for any malfunction or misbehavior at the side of the cloud provider and his sub-contractors. Even if the cloud user might in a secondary step take recourse to the cloud provider (who might, in turn, take recourse to his sub-contractors in a third step), the problem of insufficient available information discussed in large detail in the legal deliverables 2.2 and 2.7 makes this approach primarily a theoretical one.

Two foundational options might be thought of in order to address this unfair and inappropriate distribution of responsibilities and resulting risks among business partners: realigning responsibility for certain kinds of malfunctioning to the cloud provider and his sub-contractors on the one and heightening the transparency about the actual conditions and activities provided and conducted by the cloud provider. The first option is discussed in some more detail in the legal deliverable 2.7 and the potential of transparency-heightening technical mechanisms is addressed in section 5.3.2 below.

From an ethical perspective, however, the problem of control loss is much more relevant with regard to individual data subjects and their possibilities for controlling how data relating to them is handled, to whom it is disclosed etc. in order to ultimately control – as quoted from Whitman (2004) above – their “public image” which is at the heart of at least the European understanding of “privacy”. As already pointed out above, exerting this “right to control one’s public image” could become significantly harder, if not impossible, under the conditions of cloud computing and especially in those cases involving the complex interconnection of services provided by different parties.

Over the past years, several approaches have been discussed in this regard. Technical concepts particularly include parameterizing data with an expiration date as suggested by Mayer-Schönberger (2007) and prototypically implemented by Backes et al (2011). Slightly related to this are approaches from the field of distributed usage control (e.g. Pretschner et al 2006), which allow data subjects to specify simple or even sophisticated policies on the use of certain pieces of data and which has seen constant development throughout the past years (Kelbert and Pretschner 2014). If implemented in the aspired way, such mechanisms would give back a certain grade of control to cloud users or – depending on the specific setting – even individual data subjects even in case of data being processes by unknown parties under uncertain conditions and would thus countervail the problem of control loss to a certain extent. However, as such mechanisms would require significant changes to be made to the underlying information infrastructure, they have, at least so far, not found their way into practical and sufficiently widespread applications.

Mechanisms based on fully homomorphic encryption (Gentry 2009), in turn, would tackle the problem of control loss at its roots. In this model, data is transferred from one party to another only after being encrypted in a specific manner, making it impossible for the receiver to decrypt the data while still allowing for certain operations to be exerted upon it before being sent back and decrypted. In the context of cloud computing, this concept could, for instance, be used between the cloud user and the cloud provider, excluding the provider from gaining any insight on the data he receives. This, in turn, would not tackle the problem of control loss but rather eliminate it completely. There are thus good reasons for the repeated discussion of fully homomorphic encryption in the context of cloud computing. Nonetheless, at least currently, the concept has a significant performance overhead which would more than nullify all the benefits of

cloud computing. Fully homomorphic encryption did therefore also not (yet) make its way into practical applications.

The methods and technologies developed in SECCRIT, in turn, prove valuable for practically counteracting the problem of control loss from the (business) perspective of cloud users as well as from the (individual) perspective of data subjects. The Cloud assurance profile and evaluation method developed in deliverable 5.2, for instance, does not address the control loss problem as such but might prove valuable for informing cloud users about the general security level of a certain provider. Even if not empowering cloud users to exert control over the concrete givens of data handling, it gives them at least an estimation on what security efforts and remaining risks to expect at the provider side.

The IND^{2}UCE framework (described in more detail in deliverable 4.4) goes a step further and provides a comprehensive platform for defining, enforcing and monitoring complex policies for data usage control. Within SECCRIT, this framework was extended to cover cloud-specific components of the software Stack: VMware as one of the most important systems for managing virtual machines on the one hand and HBase and Hadoop as highly relevant foundational systems for cloud storage on the other. Through this extension of the IND^{2}UCE framework, cloud users can significantly gain control over the inner processes affecting “their” virtual machines and data within the cloud as soon as their cloud provider actually integrates IND^{2}UCE technologies into his offers. For example, a cloud user can specify that “his” virtual machines may not be hosted out of Europe and must – for reasons of resilience – not share a common physical machine, given that his provider is “IND^{2}UCE-enabled”. This clearly heightens the level of control that can be exerted externally by the cloud user while still not requiring him to care about concrete single machines etc., thus maintaining the level of abstraction and opacity that constitutes much of cloud computing’s added value as opposed to traditional local hosting.

The ITF (described in more detail in deliverable 5.3), in turn, allows far more meaningful and trustworthy information on the details of data handling within the cloud to be provided to cloud users than it is the case for state-of-the-art mechanisms provided by cloud management solutions. In particular, the possibility of in-depth-analyses to be performed by trusted third parties upon trustworthy raw monitoring data allows for reliable inspections on the provider’s adherence to certain agreed-upon restrictions regarding the placement of virtual machines, for example. Through delegation of respective analyses (possibly spanning and integrating input from multiple providers employed by the cloud user) to a trusted third party, this can be done without burdening the cloud users with monitoring-related details that would possibly nullify the benefits arising from cloud computing as opposed to self-hosting. While the IND^{2}UCE-framework allows for the specification and execution of security policies, the ITF can be used for validation on the actual adherence to such policies, making the combination of the two frameworks highly reasonable in matters of gaining back a certain amount of control over processes outsourced to the cloud.

Even if these technologies will hardly be employable by individual data subjects for exerting control over the flow of data concerning them (and, thus, for controlling “who knows what about them”, see above), they can nonetheless play an important role in this regard: Given the legal situation lined out in more detail in deliverables 2.2 and 2.7, the cloud user is responsible for ensuring the data subjects’ respective rights to be fulfilled (particularly including the obligation to ensure that data is not used for other purposes than those originally intended and potentially agreed to by the data subject) and for providing data subjects with certain information on how their data is actually handled. With the SECCRIT technologies ITF and IND^{2}UCE being employed by cloud providers as well as their subcontractors, the cloud user has far more possibilities for actually fulfilling these duties and responsibilities, ultimately leading to a higher likelihood that the data subjects’ will is effectively matched by the actual data handling processes. Furthermore, later versions of these technologies might also be extended in a way that supports the automatic

enforcement and propagation of the increasingly discussed “right to be forgotten” on a dataset-basis.

Another indirect effect of the mentioned technologies is that data subjects or even consumer rights associations and data protection authorities being aware of the general technological possibilities could become cautious in case a certain controller of personal data (the cloud user, for example) does *not* provide respective trustworthy information proving that certain policies for data handling and data usage have been technologically implemented and enforced by him and his potential processors and their respective subcontractors. In this case, a broad variety of possible responses ranging from avoiding respective services to initiating more detailed and formal inspections can be thought of.

In order to prevent cloud providers from claiming that the implementation of respective technologies is (for whatever reason) not possible or would impose inappropriate competitive disadvantages, a regulatory obligation for cloud providers (in data protection law, this would for instance be “data processors”) to provide state-of-the art technical means for credibly demonstrating the adherence to given instructions and restrictions regarding the handling of data might be thought of. In data protection law, the traditional on-site inspections served the same goals in the context of “processing on behalf of the controller”. Such on-site inspections are, however, largely inappropriate and hardly doable in the context of cloud computing. To what extent the “remote control and auditing” capabilities provided by the mentioned SECCRIT-technologies might serve as a more appropriate replacement for those dysfunctional on-site inspections in the context of cloud computing within a future regulatory framework remains to be discussed. The general idea, however, is also picked up and discussed in some more detail in section 4 and 6 of the final legal deliverable 2.7.

In any case, especially the SECCRIT-technologies ITF and IND^{2}UCE have been shown to be highly valuable for overcoming the general problem of significant control loss in the context of cloud computing. Thereby, they serve ethical goals directly and indirectly: Directly through giving back some control to the cloud user while still maintaining the benefits of delegation and indirectly through making cloud users better able to ensure and prove their providers’ adherence to instructions and restrictions, also making them more likely to act in the interest of data subjects. To what extent such technologies might also be anchored in the regulatory framework remains to be seen, s. D2.7, section 4 and 6.

5.3.2 Mechanisms heightening Transparency

As identified in section 3.2, the aim of ensuring transparency as one of the core principles of responsible technology design is in stark contrast to the intended “black box” nature or “opacity” of typical cloud services. Nonetheless, transparency – understood as the ability to know and retrace what actually happens within a system or service being used – is strongly desirable from an ethics perspective for various reasons, particularly from the field of privacy, self-determination and informed decision-making. Furthermore, a certain level of transparency about how data are handled “in the cloud” is indispensable for many legal regulations to become relevant at all: Without practical possibilities for validating compliance (or, rather, for discovering non-compliance), any legal provision becomes practically effect- and meaningless.

Last but not least, this also applies to the fairness and reasonableness of responsibility- and risk-distribution among multiple parties being involved in the provision of a certain cloud service. Given that the cloud user is – at least currently – the only responsible party for any misbehavior or malfunctioning, trustworthy information about how and under what circumstances data is handled by a cloud provider and his sub-contractors would significantly mitigate the problem of insufficient information being available to the cloud user for taking recourse to the cloud provider

elaborated in more detail in deliverables 2.2 and 2.7. A higher level of transparency in these regards would thus lead to a more fair and appropriate attribution of risks among the parties involved in providing a certain service. Altogether, there are thus strong ethical reasons speaking in favour of technical mechanisms for re-establishing transparency of used cloud services on demand to a certain extent.

Besides achieving a higher level of control, the SECCRIT-technologies of ITF and IND^{2}UCE would, if broadly applied in practice, also have positive effects in this regard. While in current settings of cloud computing, the position of the cloud user is rather weak as compared to that of the cloud provider (and as this imbalance of power is even strengthened by the natural tendency of the cloud market to develop monopolies or oligopolies, see section 5.3.4 below), such transparency-heightening technologies would in all likelihood work towards the cloud user and the cloud provider to meet on equal footing. Besides leading to more fair and appropriate business relations in general, this particularly includes the attenuation of the rather perverse incentives currently faced by cloud providers in matters of security, service quality and honesty in general that were extensively analyzed in a separate study in SECCRIT (see Pallas 2014).

Finally, the same technologies also provide the cloud user (typically recognized as the controller under European data protection law) with a better and more trustworthy basis of available information which he can then give to the data subjects following his respective transparency obligations lined out in more detail in deliverables 2.2 and 2.5. If taken seriously, these obligations can only limitedly be fulfilled by the cloud user in the originally intended sense – allowing data subjects to retrace and comprehend how, where and by which parties data relating to them is stored and handled – on the basis of current cloud technologies. Indirectly, technologies like ITF and IND^{2}UCE would thus allow data subjects to gain better insights about how their data is treated, to what parties it is transferred etc.[21] At least to a certain extent, this would also countervail inappropriate handling, transfers or uses of personal data that would simply remain unrecognized under current technological givens.

Altogether, technologies like ITF and IND^{2}UCE are thus ethically desirable from the perspective of transparency, too. However, regulatory obligations to provide such trustworthy transparency-heightening mechanisms could, like for the general problem of control loss, also prove necessary from the perspective of transparency as providers will – besides eschewing additional investments in general – presumably have strong interests in keeping the respective information private in order not to reveal business internals and to keep their current (albeit unfair) business advantages in matters of risk and responsibility. Any consideration in this regard would, however, also have discuss to what extent this would require additional standardization activities in order to achieve compatibility, comparability interoperability and a common set of minimum coverage across different frameworks of transparency-heightening technologies.

5.3.3 Issues of Transfer across cultural and legislative Boundaries

In section 3.3, foundational ethical issues of cloud computing also emanate from the fact of data typically being transferred across cultural and legislative boundaries. The current uproar in the aftermath of the European Court of Justice's "Safe Harbor Decision"[22] is, for example, an obvious reminder that even between different countries from the western hemisphere conceptions of privacy and the resulting legislations can be highly incompatible in matters of what a data subject can expect with regard to the handling of data relating to her.

[21] Furthermore, one might also think about the involvement of experts like consumer rights associations or data protection authorities in the external evaluation of certain cloud services on the basis of credible transparency data.

[22] ECLI:EU:C:2015:650, http://curia.europa.eu/juris/document/document_print.jsf?doclang=EN&docid=169195

Furthermore, and apart from different legal approaches, such incompatibilities can also arise in a more intangible manner. If data originating from the Europe is, for example, transferred to a cloud-data centre in India, the ultimate handling of these data basically happens in the shadow of the Indian system of values – a fact that might easily render data subject's expectations about how the respective data are treated invalid. An administrator with a conception about "privacy" that significantly differs from the European one would then, for instance, have no feeling about certain procedures infringing upon a fundamental, yet un-codified, norm or value at all. On the other hand, European regulations are also (unconsciously) based upon certain core assumptions which must not be assumed as given for any cultural area, thus posing the risk of the regulation itself becoming dysfunctional. For example, cultural values have significant impact on the effectiveness of measures used and established within the western hemisphere for achieving information security.[23] With information security being a central instrument used by regulations for safeguarding ethical values with regard to personal data privacy, in turn, it becomes obvious that divergences between cultural value systems also have implications for the effectiveness of regulations that certain data are subject to.

For the implementation of SECCRIT's demonstrators, however, the transfer of personal data across cultural and legislative boundaries was no significant issue. In the traffic control demonstrator, no personal data was used at all and in the video surveillance demonstrator, the only boundary crossed was the one between the site of collection (located in Finland) and the data centre hosting the virtual machines (located in Austria). Due to European harmonization throughout the past years and because of most relevant fundamental values being shared by those countries, no significant incompatibilities between data subjects' expectations about data handling and the actual givens had to be considered. Furthermore, this demonstrator only used video footage of data subjects who had previously given their explicit, well-informed and written consent after being informed that video data showing them would be transferred across Europe, thus at least limiting the risk of respective misconceptions.

Beyond the concrete demonstrator implementation, however, SECCRIT technologies might be of particular relevance for the fundamental challenge of data transfer across legislative and cultural boundaries. Especially after the European Court of Justice's ruling on the invalidity of data transfer from Europe to the US under the heavily used "Safe Harbor"-framework[24], it can be expected that respective transfers will have to be governed much more rigid in the future. In this context, technologies capable of enforcing location- (and, thus, legislation- and culture-) related policies like the IND^{2}UCE partially developed within SECCRIT will in all likelihood play significantly strengthened role for the still growing field of Cloud Computing. In particular, especially smaller companies not capable of controlling all the internal shifts and adjustments being done by IaaS providers in order to heighten efficiency and effectiveness will presumably profit from such technologies allowing them to easily specify respective constraints depending on the type of data being affected and on further factors influencing the acceptability of such data being transferred across certain boundaries. Like it is the case for several other challenges, trustworthy transparency mechanisms as provided by the ITF (described in more detail in deliverable 5.3) perfectly complement such constraint-enforcement systems, providing higher certainty that data actually stay within certain predefined boundaries. Both technologies are thus highly valuable components for governing the handling of – particularly personal – data with regard to the transfer across legislative and cultural boundaries. Like for the other generic challenges discussed above, it might be argued that the provision of such functionalities should thus be anchored in the regulatory framework – i.e. as some kind of (explicit or implicit)

[23] See, for example, Glaser (2009).
[24] See, again, ECLI:EU:C:2015:650, http://curia.europa.eu/juris/document/document_print.jsf?doclang=EN&docid=169195

precondition for a cloud provider to qualify for the processing of personal data relating to Europeans.

Besides such technology-driven approaches, the increased relevance of boundary-crossing (both, legislative and cultural) data transfers implied by the unbowed trend towards cloudification might also call for additional measures that at least attenuate existing frictions implied by legislative and cultural differences. With cloud computing increasingly becoming the norm rather than the exception, it might be worthwhile to consider further approaches for ensuring that data are treated in accordance with the ethical values they were subject to during their initial creation. In the legislative domain, the “Safe Harbor" concept was an – albeit unquestionably dysfunctional – approach for achieving this: Companies handling personal data referring to Europeans within the US committed themselves to handle these data in accordance with a set of principles that basically resembled the main concepts of European data protection law. These main principles were thus “exported” to “extraterritorial islands” that otherwise did not provide adequate legal safeguards and thereby – at least to a certain extent – ensured that data subjects’ expectations were also met beyond the original legislative context.

Even if this concept was in practice undermined in a multitude of ways and thus rightly discarded by the European Court of Justice in the concrete case of the “Safe Harbor” framework, the fundamental approach itself might still prove valuable in the future. In particular, it might be discussed whether and to what extent the concept might also be extended to ethical values besides mere legal compliance. Having in mind the above example of personal data being transferred from Europe to a data centre in India, one could, for example, require the respective personnel to be familiar with the European system of ethical values relating to personal data as a precondition for qualifying to handle respective data. To what extent such approaches are actually realistic and how they could be implemented (on the regulatory side as well as technically) must, however, remain subject to future considerations. In any case, thinking about future possibilities for incorporating the original ethical context in boundary-crossing data transfers in more detail seems highly valuable for lessening foreseeable ethics-related frictions emanating from ever-increased international data transfers in the context of cloud computing.

5.3.4 Indirect ethical Implications of Monopolies and Lock-Ins

The ethical issue of potential adverse implications of monopolies and lock-ins differs from the ones discussed above in several regards. It does not directly result from the fact of Cloud Computing being employed instead of traditional models of IT-usage but rather stems from the inherent characteristics of Cloud-markets that foster monopolistic market-structures. This, in turn, would, if not properly addressed, limit choice, competition and the viability of “privacy as competitive advantage”.[25] The responses to the EU Court of Justice’s invalidation of the Safe Harbor Agreement vividly illustrate this risk: Already today, European corporations using Cloud Computing offers from US providers are so strongly bound to them that they face serious trouble when not allowed to use them (at least, under the currently employed legal construct) anymore for processing personal data of European citizens. Furthermore, the ethical risks implied by the inherent characteristics of Cloud markets go well beyond privacy/data protection aspects. With one single player becoming virtually omnipresent for all applications of cloud computing, this single player would have a tremendous extortion potential which he could not only use for coercing his customers into unfavourable agreements but also for enforcing regulatory provisions that primarily serve his business objectives instead of societal goals. It is therefore in the very interest of society and individuals that monopolistic tendencies of Cloud markets are kept track of during related regulatory activities and that activities emasculating these inherent monopolistic tendencies are promoted.

[25] See, e.g., Hoeren (2000).

Even if clearly being far beyond the main focus of SECCRIT, project activities have been clearly oriented towards choice and interchangeability of providers as well as of technologies. All activities were focused on two of the most widespread provider-independent frameworks for the provision of cloud infrastructure resources: VMware as a proprietary product which can be used by different providers and OpenStack as a broadly used open source framework.

The concrete demonstrators were established on the basis of these frameworks, thus documenting their suitability for critical contexts and constituting publicly visible reference cases for comparably provider-independent applications of Cloud Computing in the critical infrastructure domain. Technologies developed within SECCRIT (like, for example, the CloudInspector framework) were also realized on the basis of these frameworks and contribute to their functional enhancement with particular regard to the requirements in the context of critical infrastructures, thus fostering the use of interchangeable and interoperable technologies in this field. The principle of provider-independency and of counteracting monopolies and lock-ins was thus constitutive for the concrete demonstrators as well as on the level of technology-development for future applications of cloud computing in general within project SECCRIT. Especially the developed technologies contribute to the functional power of the above-mentioned frameworks and the ecosystems around them and thus help towards their broader adoption. This, in turn, will work against increasing monopolisation and lock-ins in the cloud infrastructure market and thereby indirectly also serve the societal as well as individual goals of fostering choice, independency and market competition. Last but not least, it might also contribute to the prevention of market structures with single providers posing “systemic risks” and being able to avoid regulatory restrictions that would contradict their business goals.

6 Conclusion

This deliverable is an updated version of the initial ethics report compiled at the beginning of the project. The initial ethics report identified and pointed out the ethical problems emerging from the risk of using new cloud technologies. Generic ethical aspects of the two demonstrators, CCTV and (semi-) automated traffic control, were identified and the concept of cloud computing was analysed in matters of ethically relevant characteristics. On this basis, the initial ethics report identified the specific risks that potentially arise from “cloudification” within the areas of CCTV and (semi-) automated traffic have been identified (following a “triangulation approach”) and laid out how these should be addressed within the SECCRIT project. In particular, a set of safeguards was defined to be applied during the establishment of demonstrators. Furthermore, a set of compass questions was specified that should later be used for assessing the appropriate consideration of ethical aspects during the establishment of the demonstrators. Furthermore, some rather general ethically relevant aspects of Cloud Computing were identified.

In this updated version, the compass questions have been extensively discussed in relation to the two demonstrators in sections 5.1.1 and 5.2.1 and the concrete implementation of the additional safeguards within the demonstrators was described and discussed in sections 5.1.2 and 5.2.2. As depicted there, the application of the pre-identified safeguards successfully circumvented ethical conflicts. For example, the strict confinement to specifically erected test environments for the video surveillance demonstrator prevented various ethically relevant questions from practically arising in the demonstrator at all. In some cases, the practical demonstrator implementation went even beyond the pre-defined safeguards. In the traffic control demonstrator, for example, only emulated test data were used, thus making any safeguards regarding number plate or face obfuscation obsolete. Notably, ethical considerations beyond mere legal compliance led to both demonstrator implementations going well beyond the requirements of responsible national data protection authorities as laid out in deliverable 2.4

Besides the concrete demonstrators, the discussion of compass questions also referred to potential latter applications in real-world use-cases. Of course, these are beyond the original scope of the ethics report but the respective discussions did very well show potential ethical implications of the cloudification of video surveillance and traffic control systems in general. As it was shown, various SECCRIT concepts and technologies might play an important role in the context of addressing these ethical issues in future real-world applications of cloudified video surveillance and traffic control systems.

Last but not least, the identified generic ethical implications of Cloud Computing ("control loss", "lack of transparency", "transfer across cultural and legislative boundaries" and "inherent risk of monopolies and lock-ins") have also been addressed without specific confinement to the concrete demonstrator use-cases in section 5.3 of this updated report. As the respective considerations show, the concepts and technologies developed within SECCRIT can clearly play a role in addressing these generic aspects in other application contexts, too. Mechanisms for heightening transparency like the ITF or the IND^{2}UCE framework, for example, can of course be employed for serving this goal beyond video surveillance and traffic control, too.

The strong integration between reflections on ethical aspects and concrete technology and demonstrator design pursued within project SECCRIT led to various starting points for future regulations that might satisfy the specifics of cloud computing better than current ones. For instance, a realignment of responsibilities for certain kinds of malfunctioning might prove necessary in the light of omnipresent control loss present in Cloud Computing scenarios and the "remote control and auditing" capabilities provided SECCRIT-technologies might serve as a more appropriate replacement the dysfunctional on-site inspections from current data protection law. These and several further initial starting points were picked up in legal considerations throughout the project and are discussed in some more detail in the legal deliverable 2.7.

Altogether, the approach taken for the consideration of ethical aspects within SECCRIT has proven constructive. Potentially critical issues could be identified in an early stage of the project based on theoretical considerations and triangulation even before the demonstrator specification was initiated. The results of these considerations could then be fed into technology and demonstrator development based on safeguards and compass questions, leading to demonstrators that are in line with relevant ethical aspects and to technologies that, besides other goals, might also serve the adherence to ethical principles in later real-world applications.

7 Annex I: Table of “Compass Questions”

TABLE 2: "COMPASS QUESTIONS" AS DEVELOPED BY MARX (1998)

A. The Means	
	1. Harm: Does the technique cause unwarranted physical or psychological harm?
	2. Boundary: Does the technique cross a personal boundary without permission (whether involving coercion or deception or a body, relational, or spatial border)?
	3. Trust: Does the technique violate assumptions that are made about how personal information will be treated, such as no secret recordings?
	4. Personal relationships: Is the tactic applied in a personal or impersonal setting?
	5. Invalidity: Does the technique produce invalid results?
B. The Data Collection Context	
	6. Awareness: Are individuals aware that personal information is being collected, who seeks it, and why?
	7. Consent: Do individuals consent to the data collection?
	8. Golden rule: Would those responsible for the surveillance (both the decision to apply it and its actual application) agree to be its subjects under the conditions in which they apply it to others?
	9. Minimization: Does a principle of minimization apply?
	10. Public decision-making: Was the decision to use a tactic arrived at through some public discussion and decision-making process?

	11. Human review: Is there human review of machine-generated results?
	12. Right of inspection: Are people aware of the findings and how they were created?
	13. Right to challenge and express a grievance: Are there procedures for challenging the results, or for entering alternative data or interpretations into the record?
	14. Redress and sanctions: If the individual has been treated unfairly and procedures violated, are there appropriate means of redress? Are there means for discovering violations and penalties to encourage responsible surveillant behaviour?
	15. Adequate data stewardship and protection: Can the security of the data be adequately protected?
	16. Equality-inequality regarding availability and application: (a) Is the means widely available or restricted to only the most wealthy, powerful, or technologically sophisticated? (b) Within a setting is the tactic broadly applied to all people or only to those less powerful or unable to resist? (c) If there are means of resisting the provision of personal information are these means equally available, or restricted to the most privileged?
	17. The symbolic meaning of a method: What does the use of a method communicate more generally?
	18. The creation of unwanted precedents: Is it likely to create precedents that will lead to its application in undesirable ways?
	19. Negative effects on surveillants and third parties: Are there negative effects on those beyond the subject and, if so, can they be adequately mediated?
C. Uses	
	20. Beneficiary: Does application of the tactic serve broad community goals, the goals of

	the object of surveillance, or the personal goals of the data collector?
	21. Proportionality: Is there an appropriate balance between the importance of the goal and the cost of the means?
	22. Alternative means: Are other, less costly means available?
	23. Consequences of inaction: Where the means are very costly, what are the consequences of taking no surveillance action?
	24. Protections: Are adequate steps taken to minimize costs and risk?
	25. Appropriate vs. inappropriate goals: Are the goals of the data collection legitimate?
	26. The goodness of fit between the means and the goal: Is there a clear link between the information collected and the goal sought?
	27. Information used for original vs. other unrelated purposes: Is the personal information used for the reasons offered for its collection and for which consent may have been given, and do the data stay with the original collector, or do they migrate elsewhere?
	28. Failure to share secondary gains from the information: Is the personal data collected used for profit without permission from, or benefit to, the person who provided it?
	29. Unfair disadvantage: Is the information used in such a way as to cause unwarranted harm or disadvantage to its subject?

8 Annex II: Consent-Related Documents used in Demonstrator Evaluation

MIRASYS

CONSENT TO PARTICIPATE IN TEST VIDEO RECORDING

I, _____, herewith declare my free and voluntary consent to participate in test video recordings to be produced and used for evaluation purposes in the EU funded research project SECCRIT - SEcure Cloud computing for CRitical Infrastructure IT, Contract No 312758.

As an employee of MIRASYS, I am aware of the functionalities, capabilities and inner workings of the evaluated IT based video surveillance system. Furthermore, I got an additional personal introduction to the intentions of the research project and the concreate evaluation test cases.

Helsinki, November ____ 2015

_____

Mirasys Ltd. Atomitie 5 C FIN-00370 Helsinki, Finland T: +358 9 2533 3300 info@mirasys.com www.mirasys.com

References

Backes, J., M. Backes, M. Dürmuth, S. Gerling and S. Lorenz (2011): "X-pire! { A digital expiration date for images in social networks", http://arxiv.org/pdf/1112.2649.pdf [22.09.2015]

Bentham, J. (1791): "Panopticon, or the Inspection House", Dublin/London, T. Payne.

Capurro, R. (2005): "Privacy. An intercultural perspective", Ethics and Information Technology 7(1), pp. 37-47.

Cavoukian, A. (2008): "Privacy in the clouds", Identity in the Information Society 1(1), pp. 89-108.

Collingridge, D. (1980): "The Social Control of Technology", New York, St. Martin's Press.

Gentry, C. (2009): „Fully homomorphic encryption using ideal lattices". Proceedings of the 41st Annual ACM Symposium on Symposium on Theory of Computing, pp. 169-178.

Gentzoglanis, A. (2012): "Evolving Cloud Ecosystems: Risk, Competition and Regulation", Communications & Strategies, 1(85), pp. 87-107.

Glaser, T. (2009): "Culture and Information Security: Outsourcing IT Services in China", Doctoral dissertation, TU Berlin. http://www.opus4.kobv.de/opus4-tuberlin/frontdoor/deliver/index/docId/2294/file/glaser_timo.pdf [16.10.2015]

Graham, S. and D. Wood (2003): "Digitizing surveillance: categorization, space inequality", Critical Social Policy 23(2), pp. 228-248.

Hansen, S. (2004): "From 'Common Observation' to Behavioural Risk Management – Workplace Surveillance and Employee Assistance 1914-2003", International Sociology 19(2), pp. 151-171.

Hoeren, T. (2000): "Datenschutz als Wettbewerbsvorteil – eine Fortsetzung früherer Überlegungen mit neuem Vorzeichen", E-Privacy 2000, pp. 263-279, Vieweg&Teubner.

Kang, J. and D. Cuff (2005): "Pervasive Computing: Embedding the public sphere", Washington & Lee Law Review 62, pp. 93-147.

Kelbert, F. and A. Pretschner (2014): "Decentralized Distributed Data Usage Control", In Cryptology and Network Security, LNCS vol. 8813, pp. 353-369. Springer.

Koskela, H. (2000): "'The gaze without eyes': video surveillance and the changing nature of urban space", Progress in Human Geography 24(2), pp. 243-265.

Marx, G. (1998): "Ethics for the New Surveillance", The Information Society, 14, pp. 171-185, dx.doi.org/10.1080/019722498128809

Marx, G. (2002): "What's New About the 'New Surveillance'? Classifying for Change and Continuity", Surveillance & Society 1(1), pp. 9-29.

Mayer-Schönberger, V (2007): "Useful void: The art of forgetting in the age of ubiquitous computing". KSG Working Paper No. RWP07-022. http://ssrn.com/abstract=976541 [22.09.2015]

Monahan, T. (2009): "Dreams of Control at a Distance: Gender, Surveillance, and Social Control", Cultural Studies – Critical Methodologies 9(2), pp. 286-305.

Pallas, F. (2014): "An agency perspective to cloud computing", 11th International Conference on Economics of Grids, Clouds, Systems, and Services (GECON), pp. 36-51, Springer.

Paquette, S., P.T. Jaeger and S.C. Wilson (2010): "Identifying the security risks associated with governmental use of cloud computing", Government Information Quarterly 27(3), pp. 245-253.

Pretschner, A., M. Hilty and D. Basin (2006): Distributed usage control. Communications of the ACM, 49(9), pp. 39-44.

Roßnagel, A., M. Desoi, and G. Hornung (2012): “Noch einmal: Spannungsverhältnis zwischen Datenschutz und Ethik – Am Beispiel der smarten Videoüberwachung”, ZD 10/2012, pp. 459-461.

Shapiro, C. and Varian, H. (1998): “Information Rules: A Strategic Guide to the Network Economy”, Boston, Harvard Business Press.

Timmermans, J., V. Ikonen, B.C. Stahl and E. Bozdag (2010): “The Ethics of Cloud Computing: A Conceptual Review”, Proc. IEEE CloudCom 2010, pp. 614-620.

Weber, K. (2010): “Reichweite und Grenzen der Technikfolgenabschätzung”, In: Lingner, Lutterbeck, Pallas (ed.) “Die Zukunft der Räume”, pp. 53-65, Europäische Akademie Bad Neuenahr-Ahrweiler.

Whitman, J.Q. (2004): “The Two Western Cultures of Privacy: Dignity Versus Liberty”, Yale Law Journal 113, pp. 1151-1221.