SEVENTH FRAMEWORK
PROGRAMME

FP7 – SEC- 2011-284725

SURVEILLE

Surveillance: Ethical issues, legal limitations, and efficiency

Collaborative Project

SURVEILLE Deliverable 3.6

Report on methodology and criteria for incorporating perception issues in the design phase of new surveillance systems

Due date of deliverable: December 31st, 2013 (month 23)

Actual Submission date: December 18th, 2013 (month 23)

SURVEILLE Work Package number and lead: WP3 led by Prof. Coen van Gulijk (Delft)

Author: Dr. Elisa Orrù (ALU-FR), with support from Prof. Hans-Helmuth Gander (ALU-FR) and Dr. Sebastian Höhn (ALU-FR).

Project co-funded by the European Commission within the Seventh Framework Programme		
Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission services)	
CO	Confidential, only members of the consortium (including the Commission Services)	

Abstract

In this deliverable we propose a methodology to take into account perception issues when designing new surveillance technologies. The proposed methodology builds on the basic assumption that interventions to address perception issues are meaningful and compatible with a non-paternalistic approach only when preceded by measures that address the background conditions affecting perceptions. As to negative perceptions, our methodology envisages three levels of intervention: Minimum Harm by Design, Transparency by Design and Accountability by Design. The first level aims to minimise the negative impact of technologies on individuals and societies, the second to make the way surveillance functions and its improvements transparent to the public and to the people affected by surveillance, and finally, the third level aims to enable the misuse of technologies to be held to account and its authors to be sanctioned. Pertaining to perceived effectiveness, our methodology foresees two levels of interventions: measures at the first level aim at improving effectiveness compatibly with legal, ethical and societal restraints, and measures at the second aim at making success rates and improvements in effectiveness transparent to the public and to people affected by surveillance. For both negative perceptions and perceived effectiveness, also measures at the institutional, societal and legal levels are required in order to make design interventions fruitful.

Executive summary[1]

In this paper we propose a methodology to incorporate perception issues in the design phase of new surveillance technologies. The proposed methodology should enable developers of new technologies to design them in a more perception-sensitive way.

The proposed methodology consists of two parts corresponding to the two broad groups of perception issues around surveillance: negative perceptions and perceived effectiveness. The two parts are based on the same basic idea and present similar structures. The common idea behind them expresses the demand to avoid manipulative interventions that aim at addressing only perceptions without substantively improving the technologies and their uses. The common structure of the two parts of the methodology derives from this founding idea and consists of successive steps, firstly addressing the background conditions from which negative perceptions or perceptions of poor effectiveness arise and, secondly, perception itself.

The part addressing negative perceptions adopts the analysis of perception-related effects and side effects of surveillance as its starting point. Its basic assumption expresses the need for individuating, addressing and, as far as possible, correcting the rationales for negative perceptions rather than simply making the particular technology or its particular use *appear* "better" than it is. Once the background conditions related to negative perceptions are identified, the proposed methodology envisages three levels of intervention:

1. "**Minimum harm by design" (MHbD)**. Implementing MHbD for surveillance technologies implies designing them in a way which makes their negative impact on individuals, their behaviour and society as small as possible. Although such measures overlap in part with the ones prescribed by Privacy by Design, in the paper we argue that it is more appropriate in this context to refer to MHbD.
2. **Transparency by design (TbD)**. Complying with TbD requires designing technologies in a way that makes as much information as possible accessible to the public or to the people affected by surveillance.
3. **Accountability by design (AbD)**. The claim for AbD expresses the idea that the way technologies are designed should make cases of misuse and their authors traceable, accountable and sanctionable.

The second part of the methodology concentrates on issues of perceived effectiveness. It rests upon the idea that interventions should first address the background conditions

[1] We would like to thank Katerina Hadjimatheou, Erik Krempel, Martin Scheinin, Sebastian Sperber, Sebastian Volkmann and the participants in SURVEILLE's Second Annual Forum for Decision Makers, which took place in Brussels on the 23th of September 2013 for their comments on a previous draft of this paper and their advice as to literature.

affecting perceived effectiveness and avoid measures inspired by the "security theatre". After identifying the background conditions of poor perceived effectiveness, the proposed methodology requires designing technologies in order to achieve:

1. **Higher effectiveness.** This requires improving the system's effectiveness as much as it is compatible with legal, ethical and social restraints;
2. **TbD.** In order to achieve TbD for addressing perceptions of effectiveness, technologies should be designed in a way that keeps track of their operations. Combined with further information, this data on system operations should make it possible to document the success rate of the system.

For the purpose of both addressing negative perceptions and effectiveness, design measures should be supported by a social, political, institutional and legal context that makes them fruitful.

We conclude the deliverable by combining the results of the two parts into common methodological guidelines.

Table of contents

1. Results from D3.2

In SURVEILLE Deliverable 3.2 "*Review of European level studies on perceptions of surveillance. Negative perception, effects, side effects and perceived effectiveness*" we traced, systematised and described the negative perception-related effects and side effects of surveillance and investigated the relationship between perception and the effectiveness of surveillance technologies. We summarise its main results below; For reference to the studies we drew upon, please see D3.2[2].

In the first part of the paper, we identified 12 effects and side effects of surveillance and organised them according to the way they connect to the negative perception of surveillance:

A) as potential sources of negative perception;

B) as potential consequences of negative perception that influence people's behaviour, or

C) as potential threats to democracy, the rule of law and solidarity that impact society and may therefore influence perceptions of surveillance negatively.

The following effects and side effects belong in the first group, *Potential sources of negative perception*:

1. Technologies perceived as threats themselves;
2. Security dilemma and surveillance spiral;
3. Fear of misuse (incl. function creep);
4. Fear of insufficient protection of personal data;
5. Fear of unlimited expansion and irreversibility.

As part of the second group, *Potential consequences of negative perception*, we identified the following three (side) effects:

6. Self-surveillance;
7. Chilling effect;
8. Conformism and loss of autonomy.

The third group, *Impact on society,* consists of the following four effects and side effects:

9. "Control society";
10. Social exclusion and discrimination;
11. Social homogenisation;
12. Decline of solidarity.

[2] SURVEILLE Deliverable 3.2 "*Review of European level studies on perceptions of surveillance. Negative perception, effects, side effects and perceived effectiveness*", <http://www.surveille.eu/index.php/research/publications/>.

The first side effect of surveillance technologies, which consists in these technologies being perceived as threats, refers to the fact that surveillance technologies can make people feel uncomfortable even when they are perceived as being used properly, i.e. in conformity with the stated goals and in accordance with legal requirements. For instance, they may be perceived as restricting people's privacy and freedom of movement or may make people feel that they are "under suspicion".

The second side effect, labelled as the security dilemma, stems from security technologies increasing people's feelings of insecurity rather than making them feel safer. This may lead to a further side effect of surveillance: in order to compensate for increasing insecurity, more surveillance is required which in turn may further increase insecurity. A surveillance spiral is thus triggered.

The fear of misuse of surveillance technologies consists of being afraid that the surveillance technologies are used in a different way than the one officially declared and/or permitted. An important common kind of fear of misuse of surveillance is "function creep", which occurs when the use of a technology expands gradually beyond its original scope and purpose.

A further side effect of surveillance is the fear that collected personal data may not be protected sufficiently from unauthorised access. It refers to the operators' possible carelessness in allowing third parties access to the information held by them.

A last side effect of surveillance that may influence people's perception negatively is the fear of unlimited expansion and irreversibility and has to do with the feeling that some protective barriers fall away once surveillance technologies are introduced. This may happen in two ways: First, while the initial introduction of a particular technology may be met with resistance, it is much easier to expand its use after overcoming initial opposition.
Second, there is a concern that once a technology has been introduced it will be almost impossible to make it disappear again, even if it emerges that the technology is misused, ineffective, unnecessary or dangerous.

A common basis of the effects and side effects belonging to the second group is the phenomenon known as self-surveillance. People who live in situations in which surveillance is ubiquitous and potentially continuous internalise the effects of surveillance. They behave as if they are under surveillance even when surveillance is actually not taking place.

Two further effects and side effects follow: the chilling effect and conformism. The former consists of people refraining from engaging in legitimate activities such as taking part in demonstrations or signing a petition which would attract the attention of the authorities. The latter, conformism, consists of people trying to avoid any "deviant" behaviour in order to avoid attracting attention and implies a loss of autonomy: people under surveillance do not behave in accordance with their "own" reasons but rather in accordance with what they think they are supposed to do in order to not be sorted out as "deviant".

The last four effects and side effects of surveillance (8-12) belong in the third group of surveillance and are ones which impact society rather than individuals. The first one, referred to by the expression “control society”, alludes to a society in which control over citizens is maximised on the basis of generalised suspicion. Maximisation of control and generalised suspicion impact the way security is perceived in society, suggesting that everybody is a potential risk.

The risk of social exclusion occurs when surveillance is used in a way that promotes the application of categorical suspicion: controllers tend to equate whole social categories, sorted out on the basis of, for instance, their appearance or their internet activities, with dangerous groups. This strengthens prejudices because it seems to confirm them and amplifies social exclusion.

Social homogenisation as an effect of surveillance derives from the influences on individuals’ behaviour described above: the chilling effect and conformism. They reduce the variety of political positions and points of view expressed in a society and may further lead to societal stagnation, since deviant and dissenting behaviour is considered to be an important driving force for social change.

Also the last effect of surveillance, the decline of solidarity, relates to surveillance’s influence on people’s behaviour, in particular to the chilling effect. It refers to the risk that people delegate their responsibilities towards others when they know that surveillance systems are installed.

As to the positive aspects, i.e. perceived effectiveness of surveillance, in the second part of SURVEILLE Deliverable 3.2 we identified three ways in which perception and effectiveness relate to each other.

First, we addressed the question whether surveillance, independent of its actual security improvements, increases perceived security. Studies show little evidence of a causal relationship between the deployment of surveillance technologies and a reduction in fear of crime and suggest that feelings of safety depend on elements like the actual reduction of victimisation, familiarity with people, situations and places and the presence of other people rather than on technical factors like the installation of a surveillance system.

The second way perception and effectiveness relate to each other refers to the relationship between actual and perceived security, i.e. to the question of whether an improvement in actual security brings about an increase in perceived security. The review of existing studies pointed out the so-called “fear of crime paradox”: the fear of crime seems to increase or decrease independent of, or even contrary to, crime rates.

The third variation of the relationship between perception and effectiveness is properly called perceived effectiveness. It relates to the question of whether people think surveillance is effective, typically in reducing crime and reducing the fear of crime. Most surveys consulted report that the majority of those interviewed do not think of CCTV as effective.

2. Proposed methodology

Building on the results of SURVEILLE D3.2, in this paper we propose a methodology to incorporate perception issues in the design phase of new technologies. The proposed methodology should enable developers of new technologies to design them in a more perception-sensitive way.

As we have seen, perception issues surrounding surveillance can be divided into two broad groups: negative perceptions and perceived effectiveness. Because the two kinds of perceptions relate to two distinct frameworks, we will develop the methodology to address each of them in two distinct parts.

However, the two parts of the methodology are based on the same basic idea and present similar structures. The common idea behind both parts of the methodology consists in the need to avoid manipulative interventions that aim at addressing perceptions only without substantively improving the technologies. The common structure of the two methodology parts derives from their common founding idea and consists of successive steps firstly addressing the background conditions from which negative perceptions or perceptions of poor effectiveness arise and secondly, perception itself.

After developing the two parts of the methodology, we will highlight their similarities and combine the results in common methodological guidelines.

2.a. Proposed methodology – first part: negative perceptions

As its starting point, the proposed methodology for addressing issues of negative perceptions adopts the analysis of perception-related effects and side effects of surveillance, instead of focusing primarily on technologies and their uses.

The undesirability of negative perceptions is twofold. On the one hand, in SURVEILLE, negative perception is considered to be a cost of surveillance technologies. On the other hand, the perception-related effects and side effects of surveillance impact individuals' behaviour and society in a way that threatens the background conditions and basic principles of democracy, the rule of law and solidarity[3].

2.a.I. B.A.: Addressing background conditions rather than surfaces

In looking for a methodology to incorporate perception issues in the design phase of new technologies, we will take into account both aspects of such undesirability. This means that we reject an approach aiming exclusively to act on the surface of the (side)effects by improving perceptions without tackling the actual problems (which often correspond to the effects and side effects of surveillance) negative perception

[3] See SURVEILLE D3.2 "*Review of European level studies on perceptions of surveillance*", cit.

is symptom or cause of. Such an approach would be incompatible with basic principles of the rule of law and democracy.

The following example illustrates what we mean by interventions aiming only to cosmetically address perceptions. Take the case of a CCTV system installed in a park which raises privacy concerns within the public, which, in turn, influences peoples' perceptions of CCTV surveillance negatively. A possible way to address perception issues could include measures such as an advertising campaign presenting CCTV systems as friendly to park-visitors, painting the cameras green in order to make them blend into the scenery better, or make covert use of them in order to make park visitors unaware of their existence. Independent of the question whether such measures would be effective in the short, middle and long run, our approach rejects them because of their paternalistic character. They conflict with the image of human beings as rational and autonomous persons that a democratic and rule-of-law oriented approach should adopt. Instead, we propose a methodology that takes people's concerns seriously, questions the actual problems behind them and seeks to address them effectively.

Our option for a background-oriented approach is expressed by the **basic assumption** of the proposed methodology:

> *In order to address perception issues in a way compatible with fundamental rights and democratic principles, the background issues affecting negative perceptions rather than perceptions only should be tackled in the first place.*

This does not mean that the background issues affecting perception negatively are always identical with the most apparent potential rationales for it. For instance, in our example on CCTV, the most obvious rationale for negative perceptions can be a real invasion of privacy. However, further conditions may potentially cause negative perceptions, including a lack of knowledge about the existing privacy preserving features of the CCTV system, or a lack of transparency in the way they are communicated to the public. The rationales for negative perceptions vary for each case and are context-dependent: identifying them is therefore a task to be carried out on a case-by-case basis.

Whatever the rationales for negative perceptions in a particular case are, the first basic assumption of the proposed methodology expresses the need of individuating, addressing and as far as possible correcting them rather than simply making the particular technology or its particular use *appear* "better" than it is in order to avoid or minimise negative perceptions. In our example with CCTVs, corrective measures could aim at reducing the installation of cameras to a minimum, providing extensive information on existing protective mechanisms and/or improving communications transparency.

2.a.II. The three levels of intervention

Once the background conditions related to negative perceptions are identified, the proposed methodology envisages three levels of intervention. In the design phase of new technologies, measures should be taken at each of the three levels in order to effectively address the background conditions of negative perceptions and the negative perceptions themselves:

- At the **first level**, measures should be adopted in order to achieve "**minimum harm by design" (MHbD)**;

- At the **second level**, measures should be adopted to implement **transparency by design (TbD)**;

- At the **third level** measures should be adopted that aim at enhancing **accountability by design (AbD)**.

Interventions at the first level put into effect the idea expressed in the basic assumption that it is necessary to *actually* improve the technologies and their uses in order to minimise negative perceptions and that a purely cosmetic intervention on the perception level is not sufficient. The further two levels specifically address perceptions: they express the idea that, once realised, actual improvements should also be made transparent and verifiable. Only in the rare event that negative perceptions arise only from a lack of transparency, from misinformation or from the wrong kinds of communication strategies can interventions take place primarily at the second and third level.

2.a.II.i. Minimum harm by design (MHbD)

Implementing MHbD for surveillance technologies implies designing them in a way that makes their negative impact on individuals, their behaviour and society as small as possible.

MHbD can be achieved, for instance, by designing the technologies in a way that makes them invade privacy as little as possible, or to project them in a way that minimises the possibilities of misuse. A type of technology or a system can be designed to reduce its privacy impact, for example, by making it collect as little personal data as strictly necessary for achieving its goals or by making people as unidentifiable as possible from collected data, or by elaborating the collected data in a decentralised way. How this applies to each different type of technology is a matter to be solved case by case and it depends not only on the technical characteristics of a type of technology or system but also on its destination, the context of deployment, etc.

Existing examples of how to implement such proposals focus on mechanisms to enhance data minimisation. Pioneering proposals date back to the mid-Eighties, when David Chaum proposed a large-scale transaction system like the ones used for electronic payment that provides security for organisations without requiring the identification of users[4]. More recently, Claudia Diaz et al. presented a system for

[4] David Chaum, 'Security without Identification: Transaction Systems to Make Big Brother Obsolete', *Commun. ACM*, 28 (1985), 1030–1044.

signing electronic petitions that allows both controllers to detect double signatures and signatories to protect their privacy through anonymity[5]. Moreover, Josep Balash et al. elaborated a prototype electronic toll pricing system that minimises the privacy impact principally by decentralising the processing of data, thus reducing the quantity of data transmitted to the central database. The presented toll pricing system is able to prove to the central system that information transmitted is genuine without disclosing fine-grained location data that would reveal sensitive information about the users[6]. Further examples include proposals to make smart CCTV systems at airports more privacy-aware[7] and default encryption of images of individuals collected by drones[8].

One recurrent though not necessary feature of such proposals is decentralisation: the proposals show that it is possible to leave a greater amount of information in the hands of the persons whose personal data are handled without jeopardising the functionality and security of the system. Hence, when meaningful, MHbD requires achieving as much decentralisation as possible.

The focus on this level of intervention can be particularly effective at mitigating negative perceptions arising from side effects of surveillance such as the ones referred to under number 1 (technologies perceived as threats themselves), 3 (fear of misuse) and 4 (fear of insufficient protection of personal data). However, also for negative perceptions deriving from these side effects, interventions at the other two levels are necessary.

The notion of MHbD overlaps in part with the notion of Privacy by Design (see info-box), nowadays a well established and often quoted set of principles[9].

Although we recognise the validity of the research done in the PbD realm, we prefer, nevertheless, not to refer to PbD here and elaborate instead the notion of MHbD for the following reasons:

[5] Claudia Diaz and others, 'Privacy Preserving Electronic Petitions', *Identity in the Information Society*, 1 (2008), 203–219.

[6] Josep Balasch and others, 'PrETP: Privacy-Preserving Electronic Toll Pricing', in *19TH USENIX SECURITY SYMPOSIUM* (USENIX Association, 2010), pp. 63–78.

[7] Christoph Bier, Pascal Birnstill, Erik Krempel, Hauke Vagts, Jürgen Beyerer
Enhancing Privacy by Design From a Developer's Perspective
1st Annual Privacy Forum 2012, Limassol, Cyprus.

[8] Cavoukian, Ann, *Surveillance, Then and Now: Securing Privacy in Public Spaces*, June 2013, http://www.privacybydesign.ca/index.php/paper/surveillance-then-and-now-securing-privacy-in-public-spaces/ [accessed 27 November 2013].

[9] For a set of principles aiming at the same purposes as to PbD applied to CCTV see European Forum for Urban Security: Charter for a democratic use of video surveillance, 2010, http://www.cctvcharter.eu/fileadmin/efus/CCTV_minisite_fichier/Charta/CCTV_Charter_EN.pdf and the SURVEILLE D3.2, "*Paper by local authorities end users*", http://www.surveille.eu/index.php/research/publications/.

- "PbD" misleadingly suggests that technologies complying with its requirements bring about an improvement of privacy. The expression "MHbD", on the contrary, signalises that surveillance technologies *always* bring about a negative impact on individuals and society and that impact can be minimised at best but will never be completely eliminated.

- PbD focuses on information privacy, i.e. privacy regarding the collection and use of personal information. However, it is not proven that (information) privacy is the only right threatened by surveillance, nor that threats to other rights and values necessarily depend on a previous violation of (information) privacy or that they would not take place if the invasion were to be removed. The answer to the question whether intrusions in information privacy are always the preconditions for further violations mostly depends on the definition of "privacy" adopted, which is itself a controversial matter[10]. Moreover, as the following example should illustrate, whether a violation of fundamental rights and values depends on a previous privacy intrusion is at least a matter of perspective. Take for example the effect "social exclusion and discrimination" and a CCTV-surveillance scenario. In our scenario, the installed CCTV system incorporates biometrical facial recognition, which allows for the identification of the people filmed. If people are considered to be suspect on the basis of any of the data collected by the CCTV system, they are singled out for further checks. To select suspects, the operators also use categorical suspicion based on appearance. By removing biometrical identification, the system would be made less intrusive to privacy. However, the effect "social exclusion and discrimination" would not diminish unless the skin colour of a person or the way she is dressed were also concealed in the output image of the CCTV system. Of course, one could object that concealing the particular features of people filmed is also a privacy-preserving measure. However, an approach that focuses not only on privacy like the one proposed here seems able to solve such problems in a more straightforward way. By referring to a "minimum harm" rather than only "privacy" we aim not to exclusively restrict *a priori* the field of intervention into privacy-related issues.

- PbD targets whole organisations' practices instead of kinds of technologies or technology systems. PbD, for instance, does not primarily or exclusively prescribe how a licence plate recognition system should be designed in order to minimise its impact on privacy. Instead, Cavoukian's approach targets the whole context in which such a system is adopted and prescribes measures regarding, say, the code of conduct for employees handling the data, or the legislative measures limiting the uses of the data. Although such a holistic approach is meaningful and technical aspects should not be addressed in

[10] Daniel J Solove, *Understanding Privacy* (Cambridge, Mass.: Harvard University Press, 2008) and Judith DeCew, 'Privacy', in *The Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta, Fall 2013, 2013 <http://plato.stanford.edu/archives/fall2013/entries/privacy/> [accessed 26 November 2013].

isolation from the organisational, societal, political and legal context in which they are used, for analytical purposes we find it to be more fruitful to keep the different stages separate. We therefore concentrate here on the *technological* aspects that: a) reduce technologies' impact on negative perception through reducing the impact on basic values of solidarity, democracy and the *rule of law*; b) make it possible and meaningful to adopt further strategies at the institutional, political, legal and societal level that further reduces the impact on negative perceptions and the above mentioned values.

- As pointed out in different contributions, current definitions of “PbD” are characterised by vagueness and do not provide guidelines for how to translate its principles into engineering practices for designing new technologies[11]. Given these shortcomings, PbD risks becoming a label with which to reassure consumers and the public without bringing about real improvements for privacy - exactly the opposite of our first basic assumption[12].

INFOBOX: PRIVACY BY DESIGN

(See Cavoukian, Ann, 'Privacy by Design. The 7 Foundational Principles', August 2009, <http://www.privacybydesign.ca/content/uploads/2009/08/7foundationalprinciples.pdf/>, accessed 27 November 2013)

Ann Cavoukian introduced the concept of PbD in the Nineties to address the growing challenges posed by new technologies to the protection of personal information[1]. PbD aims to make privacy assurance the default mode of operation for organisations and it may be attained by acting in accordance with the following seven foundational principles of PbD:

1. Proactive not reactive; preventive not remedial.
   This principle expresses the idea that PbD should act proactively, preventing privacy intrusions from happening instead of intervening after they have occurred.

2. Privacy as the default setting.
   No action should be required by users in order to protect their privacy: Personal data should be automatically protected as the default rule.

[11] Gürses, Seda, Carmela Troncoso, and Claudia Diaz 'Engineering Privacy by Design', <https://www.cosic.esat.kuleuven.be/publications/article-1542.pdf/> [accessed 27 November 2013], and Christoph Bier, Pascal Birnstill, Erik Krempel, Hauke Vagts, Jürgen Beyerer, Enhancing Privacy by Design, cit.

[12] Gürses, Seda, Carmela Troncoso, and Claudia Diaz 'Engineering Privacy by Design', cit.

3. Privacy embedded into design.
   Privacy should be embedded into the system from the beginning, it belongs to its core functionality instead of being added after the system has been already designed.

4. Full functionality – positive-sum, not zero-sum. No trade-offs between privacy and security are necessary: in the PbD approach it is possible to have both.

5. End-to-end security – full lifecycle protection.
   Because privacy protection is embedded into the system from the design phase, it is operative before personal information is collected. This should guarantee the protection of personal information throughout the whole lifecycle of data processing.

6. Visibility and transparency – keep it open.
   This principle aims at assuring stakeholders that the system is operating in a privacy-protecting manner and is subject to independent verification.

7. Respect for user privacy – keep it user-centric.
   Users' interests should have the highest priority both in the design and operating phases.

2.a.II.ii. Transparency by design (TbD)

Transparency by design means that technologies should be designed in a way that makes as much information as possible accessible to the public or to the persons affected by their use (typically people affected by surveillance).

The way technologies are designed should, for instance, enable the group of people targeted by surveillance to know:

- For what purposes were the technologies created;

- How the technologies are used and whether these uses correspond with the original/authorised purposes;

- How much and what kind of personal information is collected using the technologies, how it is used and for how long it is kept;

- Who operates the technologies and who has access to the data collected by them;

- What measures of MHbD are implemented in the technologies and what are their limits;

- What measures of AbD (s. below) are implemented in the technologies and how they can make use of them;

- How they can access the information listed above.

As the list above shows, transparency should cover also, but not only, the other two levels of intervention: MHbD and AbD.

The quantity and kind of information made available and to whom varies largely depending on the technology, its particular use, and the context of deployment.

In general, two broad categories of information about surveillance technologies exist: general and personal. They should be made available according to two different strategies: General information should be publicly available, whereas personal information should be made available only to the persons to whom it belongs. For instance, in the case of a licence plate recognition system on a motorway, the information available to the whole public could include: the purposes for which it was installed; the authority which authorised the installation and for what purposes and under what limitations; whether the system works in a covert, overt or opaque manner; what kinds of data are collected; how long they are kept before being definitively deleted; to whom they are communicated; who has access to them. Annual reports subjected to independent verification can be a way to reach publicity for such kinds of information[13]. Clearly, the general public should not have access to the actual database – however the persons whose data were collected should. In this case, then, the plate recognition system should be designed in a way that makes it possible for each individual to know whether her vehicle had been tracked by the system, when and where it happened, who/which organisation or agency accessed the data, what was done with the data, whether they where deleted at the right time or not, and so on.

The way to access these pieces of information should be straightforward and uncomplicated; no special skills or knowledge should be expected in order to have access to them: No more, say, than the ability to use a smart phone if the technology in question is a smart phone, or no more than the ability to browse the internet if the technology in question is an internet browser. For our example of a plate recognition system, no more skills than the ones necessary for obtaining a driving licence and carrying on the usual bureaucratic activities related to the possession of a car (such as stipulating an insurance and paying car taxes) should be necessary to access to the data.

If supported by appropriate measures of MHbD and AbD, the focus on this level of intervention can be particularly effective against side effects of surveillance such as number 5 (Fear of unlimited expansion and irreversibility), and numbers 6 to 8 (self-surveillance, chilling effect, conformism and loss of autonomy). Transparency about the

[13] Cavoukian, Ann, *Surveillance, Then and Now*, cit.

objectives and uses of surveillance technologies can, for instance, be effective in minimising the fear of unlimited expansion and irreversibility; while clear and precise information about where and when surveillance takes places and about the criteria for suspicion can minimise self-surveillance and the related side-effects of societal chill, conformism and loss of autonomy. These examples make the interdependence of the three levels clear: transparency of course can be contra productive if the objectives of surveillance are too broad, if surveillance is ubiquitous or if the criteria for suspicion are too vague.

2.a.II.iii. Accountability by design (AbD)

The claim for AbD expresses the idea that the way technologies are designed should make cases of misuse and their authors traceable, accountable and sanctionable.

Examples of misuse of surveillance technologies include:

- Deployment beyond the original purposes;
- Use in places or situations that are not authorised or not identical with the original ones;
- Use of the collected data for purposes that are not authorised or different from the original ones;
- Non-authorised circulation of the collected data;
- Use of the collected data beyond the authorised timeframe;
- Deployment of the technologies and use of the collected data in a discriminatory way.

Proposals of design forms that enable accountability focus on logs registering access and handling of personal data within a system and have been applied to e-mail service providers handling e-mail users' data, bank operators handling the personal data of bank customers, or operators accessing data collected by drones[14]. Existing literature shows that it is possible to design systems in a way that makes *a posteriori* checks about compliances with the data usage rules meaningful.

[14] PRESCIENT, International Conference of the PRESCIENT Project, Berlin, 27-28 November 2012, Session 3: 1. Accountability by Design for Privacy, <http://prescient-project.eu/prescient/inhalte/download/prescient2012.pdf>; Daniel J. Weitzner and others, 'Information Accountability', *Commun. ACM*, 51 (2008), 82–87 <doi:10.1145/1349026.1349043>; Denis Butin, Marcos Chicote and Daniel Le Metayer, 'Log Design for Accountability', in *2012 IEEE Symposium on Security and Privacy Workshops* (Los Alamitos, CA, USA: IEEE Computer Society, 2013), 0, 1–7; Cavoukian, Ann, *Surveillance, Then and Now*, cit.

Rather than specifically impact on certain effects and side effects of surveillance, AbD seems to be the key for the effectiveness of MHbD and TbD because they may function effectively only if checks are possible.

2.a.III. Beyond design

The remedies foreseen by the proposed methodology are meaningful and can be effective only if backed up by a broader context in which they can actually operate. For instance, the technical features of a particular technology allowing for tracking accountability for violations and misuse are meaningful only in a context that foresees sanctions for such violations.

Pertaining to such a context, among others, are societal, institutional, political, and legal settings.

Examples of measures at the legal level include mandatory, previous judicial authorisation for the deployment of surveillance technologies, strict and binding codes of conduct for surveillance operators, and a mandatory two-signature protocol to access data collected by surveillance systems[15].

Other, non-legal measures include public discussions carried out routinely before the installation of new surveillance systems and centres for facilitating communication between individuals and institutions[16].

Moreover, making different, practicable options available among which individuals can choose can strengthen the results attained by applying the proposed methodology. There should be options between not flying at all and letting one's biometric data be collected, or between having one's email exchange intercepted vs. having to renounce writing emails.

[15] Cavoukian, Ann, *Surveillance, Then and Now*, cit. For existing regulatory instrument at the European level, in particular Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data and Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 on the Processing of Personal Data and the Protection of Privacy in the Electronic Communications Sector see SurPRISE D3.2, Report on regulatory frameworks concerning privacy and the evolution of the norm of the right to privacy, March 2013, http://surprise-project.eu/wp-content/uploads/2013/06/SurPRISE_D3.2_Report-on-regulatory-frameworks-concerning-privacy-for-final-formatting_v094.pdf, last visited on 12th November 2013.

[16] Cavoukian, Ann, *Surveillance, Then and Now*, cit.

2.b. Proposed methodology – second part: Perceived effectiveness

Of the three variations of the relationship between perception and effectiveness described in section 1, the first and the second seem to be determined by factors external to the use of surveillance such as social and interpersonal relationships and the actual crime rates. Hence, it seems that a meaningful intervention for addressing perception in these cases should focus on those external factors rather than on the design of new technologies.

Therefore we will concentrate here on the third way perception and effectiveness relate to each other, i.e. on perceived effectiveness.

There are many ways people may think of surveillance as being effective in reducing crime: for instance, they may refer to the prevention of crimes being committed due to the deterrence of potential offenders as well as to the identification of offenders in the prosecution phase. Respondents' perceptions of the effectiveness of surveillance in such different meanings vary considerably, but, unfortunately, surveys available often do not clearly distinguish between them[17].

How can such perceptions be addressed already in the design phase of new technologies?

2.b.I. B.A.: Addressing background conditions rather than surfaces/II

As in the case of negative perception, also for perceived effectiveness, the proposed methodology is based on the idea that measures aimed only at addressing perceptions are insufficient. Therefore, the basic assumption of the proposed methodology can be reformulated as follows:

> *Interventions should first address the background conditions affecting perceived effectiveness rather than only focussing on perceptions.*

In the realm of perceived effectiveness, the basic assumption expresses the need to avoid measures inspired by the so-called "security theatre". This "covers measures taken, ostensibly in the name of security, whose value lies solely (or at least mostly) in their capacity to give the reassuring impression that *something is being done*, that *steps are being taken*, that *someone is on the case*—rather than in actually increasing security, considered from an objective standpoint. The role of security theatre is to increase *perceived* security, without necessarily having any positive effect in terms of *actual* security".[18]

[17] See SURVEILLE D3.2, section 3.c).

[18] PACT D1.4 Societal Impact Report, 2012, <http://www.projectpact.eu/deliverables/wp1-root-branch-review/d1.4-social-impact-report>, p. 16 and Bruce Schneier, *Beyond Fear : Thinking Sensibly about Security in an Uncertain World* (New York, NY, 2003) <https://katalog.ub.uni-freiburg.de/link?id=106960318>.

In the design of new technologies, interventions inspired by the security theatre should be avoided for two reasons. First, like in the case of negative perception, manipulative interventions would contradict and threaten basic principles of democracy and the rule of law. Second, as demonstrated by the analysis of surveys carried out in D3.2, respondents are aware of the possibility that measures are taken just to reassure people, without tackling the actual problems[19]. Therefore, besides being morally and politically problematic, approaches inspired by the security theatre do not seem to have good chances of success.

2.b.II. Two levels of intervention

After looking for the background conditions of poor perceived effectiveness, the proposed methodology requires designing technologies in order to achieve:

3. **Higher effectiveness**
4. **TbD**

The first point requires acting accordingly with the basic assumption: the first question to ask is whether the perception that technologies are poorly effective is well grounded. If this is the case, the first step consists of improving the system's effectiveness as much as it is compatible with legal, ethical and social restraints.

Only after dealing with actual effectiveness are further measures meaningful. Efficiency and effectiveness issues have been tackled extensively in the SURVEILLE Deliverables 3.3 "*Report on system effectiveness, efficiency and satisfaction assessment*", 3.4 "*Design of a research methodology for assessing the effectiveness of selected surveillance systems in delivering improved security"* and *3.5 "Cost modelling"*, so we will not discuss them here further.

As to AbD, at this stage of research it seems that no measures of AbD are needed for addressing issues of perceived effectiveness.

2.b.II.i. Transparency by design (TbD)

In order to achieve TbD for addressing perceptions of effectiveness, technologies should be designed in a way that keeps tracks of their operations. Combined with further information, this data on system operations should make it possible to document the success rate of the system.

From the collected data it should be possible to reconstruct, for instance:

5. How many cases the system analysed;

6. How many warnings the system issued;

[19] SURVEILLE D3.2, section 3.c).

7. How many warnings eventually led to successful interventions and how many did not;

8. Whenever possible, how many potential dangers or infractions the system failed to detect.

For instance, a metal detector used for luggage screening at airports should keep track of the number of items examined and of the number of items selected for further checks and such data should be combined with information from the security agencies on the number of dangerous items eventually detected and, whenever possible, the number of dangerous items that went undetected through checks. An electronic plate recognition system should keep track of the number of vehicles tracked and of the number of infractions registered; these data should be combined with the number of infractions eventually sanctioned and, whenever possible, with the number of infractions that remained undetected.

2.b.III. Beyond design

Also in the case of perceived effectiveness, design measures should be supported by a social, political, institutional and legal context that makes them fruitful.

First, as we have seen, the data collected should be integrated and elaborated in order to achieve knowledge about the actual effectiveness of the technologies considered. The data keeping track of the system functionality should be integrated with data on the number of false positives, false negatives and of the success rates. Such data should be further statistically elaborated.

Second, openness about the effectiveness of surveillance technology is needed. Both raw data and statistics should be made public. Annual reports about the effectiveness of different security technologies, including information about the strengths and limits of each technology, could also be a useful instrument for backing up TbD. Whether the further use of a technology is meaningful or not in the light of effectiveness should be a matter of public and open debate.

Third, consequences have to be drawn from the information about effectiveness. If a technology proves to be inadequate for achieving the purposes for which it was adopted, it should not be used further. Clearly, this presupposes that the purposes of the deployment of a particular technology should be clearly stated from the beginning.

Finally, statistical data and further information about the effectiveness of surveillance technologies should be communicated in a way that takes into account the most recent research on the perception of risk and the role of emotions in risk perception[20]. Existing psychological research shows a possible link between the communication of information

[20] Sabine Röser and others, 'Handbook of Risk Theory Epistemology, Decision Theory, Ethics, and Social Implications of Risk', 2012.

and the acceptance of security interventions and that acceptance increases when decisions about the deployment of security technologies are perceived as achieved through a fair process[21].

[21] Magdalena Schuler and Larissa Wolkenstein, 'Psychologie und Sicherheitstechnologie – Psychologische Auswirkungen von Sicherheitstechnologien auf den Menschen und die Einstellung von Menschen dieser Technik gegenüber', in Gander, Hans-Helmuth and Rischer, Gisela (Hrsg.), *Sicherheit und offene Gesellschaft. Herausforderungen, Methoden und Praxis einer gesellschaftspolitischen Sicherheitsforschung* (Baden-Baden, Nomos, forthcoming).

3. Shortcomings and need for further research

The research carried out for this Deliverable is pioneering work: as far as the authors know, no literature exists on how to specifically address perception issues in the design phase of new technologies.

Due to the initial character of such research, further developments on all the relevant topics, and in particular on MHbD, TbD for both negative perceptions and perceived effectiveness and AbD is needed.

The part of the research for which we could rely more on existing literature is the part on MHbD. However, as we have seen, research so far almost exclusively concentrated on PbD and related issues such as data protection. However, as we argued above, the PbD approach is unsatisfactory for our purposes because its focus is both too broad and too narrow. On the one hand PbD, in spite of its name, merges different levels of intervention, not referring only to the design phase of new technologies but also targeting the whole life cycle of complex surveillance systems. On the other hand, PbD focuses only on information privacy, whereas there is a need to consider also infractions of other fundamental rights and values, as we sought to do by introducing the notion of MHbD. Technical research in this direction, including proposals on how to design technologies in order to minimise their harm on individuals and society beyond privacy violations would be much welcome.

4. Conclusions: combined methodology guidelines

In this paper, we separately developed the methodologies for addressing negative perceptions and perceived effectiveness. However, the two methodologies have a parallel structure and several similarities. It is thus possible to combine them in the following methodology guidelines.

The first step for applying the methodology is to ask if the technology to be developed may be perceived negatively and whether it will be perceived as being effective. This preliminary inquiry may rest upon surveys (existing on similar technologies or *ad hoc*) to be carried out with the modifications suggested in D3.2, i.e. recruiting respondents to adequately represent the views of those who are most affected by surveillance, and on the basis of simulations and literature. The inquiry should aim at finding out whether the new technology, in the context for which it will be employed:

A) Potentially has the following negative perception-related effects and side effects:

1. Technologies perceived as threats themselves;
2. Security dilemma and surveillance spiral;
3. Fear of misuse (incl. function creep);
4. Fear of insufficient protection of personal data;
5. Fear of unlimited expansion and irreversibility;
6. Self-surveillance;
7. Chilling effect;
8. Conformism and loss of autonomy;
9. “Control society”;
10. Social exclusion and discrimination;
11. Social homogenisation;
12. Decline of solidarity.

and

B) May be perceived as “security theatre” or otherwise ineffective.

Once the potential (side) effects of the use of a technology in a particular context are identified and it has been ascertained that it may be perceived as poorly effective, the second step consists of identifying the actual circumstances from which such perceptions arise, according to the basic assumption of the proposed methodology.

The successive steps consist of interventions at the further levels: minimum harm, transparency and accountability by design to address negative perceptions, improvements on effectiveness and transparency by design to address perceived effectiveness.

The table below outlines the methodology guidelines for addressing both negative perceptions and perceived effectiveness.

Table summarising methodology

Domain	NEGATIVE PERCEPTIONS	PERCEIVED EFFECTIVENESS
BA	ê BACKGROUND CONDITIONS ê	
	Check for side-effects: 1. Technologies perceived as threats themselves; 2. Security dilemma and surveillance spiral; 3. Fear of misuse (incl. function creep); 4. Fear of insufficient protection of personal data; 5. Fear of unlimited expansion and irreversibility. 6. Self-surveillance; 7. Chilling effect; 8. Conformism and loss of autonomy. 9. Control society; 10. Social exclusion and discrimination; 11. Social homogenisation; 12. Decline of solidarity.	Is the use of technology effective in the particular context?
	ê	ê
1st level	MHbD	Effectiveness improvement
2nd level	TbD	TbD
3rd level	AbD	/
Beyond design	Social – institutional – political- legal measures	

References

Literature

- Balasch, Josep, Alfredo Rial, Carmela Troncoso, Christophe Geuens, Bart Preneel, and Ingrid Verbauwhede, ‘PrETP: Privacy-Preserving Electronic Toll Pricing’, in *19TH USENIX SECURITY SYMPOSIUM* (USENIX Association, 2010), pp. 63–78.

- Bier, Christoph, Pascal Birnstill, Erik Krempel, Hauke Vagts, Jürgen Beyerer, *Enhancing Privacy by Design From a Developer’s Perspective*, 1st Annual Privacy Forum 2012, Limassol, Cyprus.

- Butin, Denis, Marcos Chicote, and Daniel Le Metayer, ‘Log Design for Accountability’, in *2012 IEEE Symposium on Security and Privacy Workshops* (Los Alamitos, CA, USA: IEEE Computer Society, 2013), 0, 1–7.

- Cavoukian, Ann, 'Privacy by Design. The 7 Foundational Principles', August 2009, <http://www.privacybydesign.ca/content/uploads/2009/08/7foundationalprinciples.pdf/>,

- Cavoukian, *Ann, Surveillance, Then and Now: Securing Privacy in Public Spaces*, June 2013, http://www.privacybydesign.ca/index.php/paper/surveillance-then-and-now-securing-privacy-in-public-spaces/ [accessed 27 November 2013].

- Chaum, David, ‘Security without Identification: Transaction Systems to Make Big Brother Obsolete’, *Commun. ACM*, 28 (1985), 1030–1044.

- DeCew, Judith, ‘Privacy’, in *The Stanford Encyclopedia of Philosophy*, ed. by Edward N. Zalta, Fall 2013, 2013 <http://plato.stanford.edu/archives/fall2013/entries/privacy/> [accessed 26 November 2013]

- Diaz, Claudia, Eleni Kosta, Hannelore Dekeyser, Markulf Kohlweiss, and Girma Nigusse, ‘Privacy Preserving Electronic Petitions’, *Identity in the Information Society*, 1 (2008), 203–219.

- Gürses, Seda, Carmela Troncoso, and Claudia Diaz 'Engineering Privacy by Design', <https://www.cosic.esat.kuleuven.be/publications/article-1542.pdf/> [accessed 27 November 2013].

- Röser, Sabine, Rafaela Hillerbrand, Per Sandin, and Martin Peterson, ‘Handbook of Risk Theory Epistemology, Decision Theory, Ethics, and Social Implications of Risk’, 2012.

- Schneier, Bruce, *Beyond Fear : Thinking Sensibly about Security in an Uncertain World* (New York, NY, 2003).

- Schuler, Magdalena and Larissa Wolkenstein, 'Psychologie und Sicherheitstechnologie – Psychologische Auswirkungen von Sicherheitstechnologien auf den Menschen und die Einstellung von Menschen dieser Technik gegenüber', in Gander, Hans-Helmuth and Rischer, Gisela (Hrsg.), *Sicherheit und offene Gesellschaft. Herausforderungen, Methoden und Praxis einer gesellschaftspolitischen Sicherheitsforschung* (Baden-Baden, Nomos, forthcoming).

- Solove, Daniel J, *Understanding Privacy* (Cambridge, Mass.: Harvard University Press, 2008).

- Weitzner, Daniel J., Harold Abelson, Tim Berners-Lee, Joan Feigenbaum, James Hendler, and Gerald Jay Sussman, 'Information Accountability', *Commun. ACM*, 51 (2008), 82–87.

EU-Directives

- Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data

- Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 on the Processing of Personal Data and the Protection of Privacy in the Electronic Communications Sector

EU-Projects Deliverables and other documents

- European Forum for Urban Security: Charter for a democratic use of video surveillance, 2010, http://www.cctvcharter.eu/fileadmin/efus/CCTV_minisite_fichier/Charta/CCTV_Charter_EN.pdf and the

- PACT D1.4 Societal Impact Report, 2012, <http://www.projectpact.eu/deliverables/wp1-root-branch-review/d1.4-social-impact-report>.

- PRESCIENT, International Conference of the PRESCIENT Project, Berlin, 27-28 November 2012, Session 3: 1. Accountability by Design for Privacy, <http://prescient-project.eu/prescient/inhalte/download/prescient2012.pdf>.

- SurPRISE D3.2, "Report on regulatory frameworks concerning privacy and the evolution of the norm of the right to privacy", March 2013, <http://surprise-project.eu/wp-content/uploads/2013/06/SurPRISE_D3.2_Report-on-regulatory-frameworks-concerning-privacy-for-final-formatting_v094.pdf>.

- SURVEILLE D2.3, "*Paper by local authorities end users*", February 2013, <http://www.surveille.eu/index.php/research/publications/>.

- SURVEILLE D3.2, “Review of European level studies on perceptions of surveillance. Negative perception, effects, side effects and perceived effectiveness”, September 2013, <http://www.surveille.eu/index.php/research/publications/>.

- SURVEILLE D3.3, “Report on system effectiveness, efficiency and satisfaction assessment”, September 2013, <http://www.surveille.eu/index.php/research/publications/>.

- SURVEILLE D3.4, “Design a research methodology for assessing the effectiveness of selected surveillance systems in delivering improved security”, September 2013, <http://www.surveille.eu/index.php/research/publications/>.

- SURVEILLE D3.5, “Cost modelling”, forthcoming, http://www.surveille.eu/index.php/research/publications/.