UaESMC

Project N^{o}: **FP7-284731**

Project Acronym: **UaESMC**

Project Title: **Usable and Efficient Secure Multiparty Computation**

Instrument: **Specific Targeted Research Project**

Scheme: **Information & Communication Technologies**

Future and Emerging Technologies (FET-Open)

Deliverable D1.4
Expert Feedback on Prototype Application

Due date of deliverable: 30th April 2014

Actual submission date: 5th June 2014

SEVENTH FRAMEWORK PROGRAMME

Start date of the project: **1st February 2012** Duration: **36 months**

Organisation name of lead contractor for this deliverable: **UT**

Specific Targeted Research Project supported by the 7th Framework Programme of the EC		
Dissemination level		
PU	Public	✓
PP	Restricted to other programme participants (including Commission Services)	
RE	Restricted to a group specified by the consortium (including Commission Services)	
CO	Confidential, only for members of the consortium (including Commission Services)	

Executive Summary:
Expert Feedback on Prototype Application

This document summarizes deliverable D1.4 of project FP7-284731 (UaESMC), a Specific Targeted Research Project supported by the 7th Framework Programme of the EC within the FET-Open (Future and Emerging Technologies) scheme. Full information on this project, including the contents of this deliverable, is available online at `http://www.usable-security.eu`.

This report contains the results of the second round of interviews conducted with selected experts from various fields. The first part of the report describes the research methodology. The second part presents an analysis of five baseline conditions for the adoption of secure multiparty computation (SMC) as well as the corresponding barriers to adoption. The main baseline conditions/barriers to adoption were as follows:

- Being sufficiently/insufficiently informed
- Presence of perceived need/existence of a functional substitute
- Presence/absence of task-technology fit
- Sufficient amount/lack of resources
- Good organizational fit/avoidance of uncertainty

Section three outlines the potential contexts of use suggested by the interviewees. A preliminary assessment of SMC's potential impact on the following practices of data handling is performed:

- Anonymisation prior to giving data to research
- Anonymisation by research team
- Data filtering combined with anonymisation
- Keeping data and user separate
- Delegating data operations to the third party
- Limiting the research problem
- Providing a secure pathway to data
- Making data sharing obligatory by law

The report also includes suggestions for further R&D activities.

List of Authors

Laur Kanger (UT)
Pille Pruulmann-Vengerfeldt (UT)

Contents

Chapter 1

Overview of methodology and presentation of interviewees

1.1 Methodology

Interviewees were contacted by e-mail that contained information about the project, description of the interview process and references to background information about the project. The latter also included a link to the statistics demo prepared on the basis of feedback from the first round of interviews. The interviewees were encouraged to familiarize themselves with the basic working principle of SMC and to try out the statistics demo.

The overall aims of the second round of interviews were as follows:

- to assess the potential barriers to the adoption of SMC identified in the first round of interviews (but also using additional material, see Section 2.1);
- to gather user feedback on the statistics demo. The demo enabled the user to simulate SMC in practice by allowing to perform various statistical operations (e.g. mean, box plot, heatmap, outlier detection) using a number of variables (e.g. gender, age, income). The data itself was computer-generated. The demo is available online at `http://demo.usable-security.eu/`;
- to continue mapping the potential contexts of use along with the identification of adoption baseline conditions as well as potential barriers to adoption;
- to probe the feasibility of a couple of use case scenarios (related to business process integration) conceived by the UaESMC development team.

Semi-structured interview design was used for the interviews. Interviews were conducted by two researches, all of them using the same interview design. The questions were flexibly tailored to interviewees' field of expertise during the interview.

For the analysis, the data were reduced by coding them into groups. The analysis proceeded in three steps:

- first, the data was coded according to the main aims of the interview (baseline conditions of adoption, barriers to adoption, feedback on statistics demo, potential use contexts);
- in the second step the baseline conditions were further reduced to five categories (information, need, task-technology fit, resources, organizational fit). The process was repeated for barriers to adoption. In parallel, potential contexts of use of SMC and existing practices of data handling were located and systematized;
- in the final step the baseline conditions were matched with barriers to adoption (see Table 2.1). In parallel, a comparison of SMC's advantages and disadvantages in relation to existing practices of data handling was conducted.

Table 1.1: List of interviewees.

Interviewee	**Field of expertise**	**Organization**	**Length (in minutes)**
I1	Statistics (sociology)	University of Tartu	115
I2	Security and defence	Headquarters of the Estonian Defence Forces	89
I3	Statistics	Statistics Estonia	65
I4	Security and defence	Government Office	55
I5	Business process optimization	Telema	72

1.2 Interviewees

Interviewees were selected from various fields in Estonia. Overall, 5 interviews were conducted. Interviews were conducted over a period of February to April 2014. All interviews were conducted in a face to face setting. More details can be found in Table 1.1.

Full transcripts and audio files of the interviews are stored for future reference.

Chapter 2

Results

2.1 Adoption baseline and potential barriers

The experts interviewed in the second round were asked to assess the potential barriers to the adoption of secure multiparty computation (SMC). The barriers were compiled on the basis of two data sources: 1) first round interviews (part of the UaESMC project); 2) map of barriers to the adoption of privacy- and security-related ICT solutions (part of the FIRE project, see [Kanger 2013] for more detail). Ten barriers were identified by the researchers as potentially most salient. These were as follows:

- The business value of SMC remains unclear
- SMC is too complex to adopt
- SMC is too costly to adopt
- There are no sufficiently trustworthy partners for using SMC
- Concerns about SMC's security
- There might be legal obstructions to employing SMC
- SMC's contribution to problem-solving remains unclear
- The problem could be solved by other means
- SMC does not help to solve the underlying problem
- SMC would not enable one to access the information deemed necessary by the user

The barriers were printed out on separate cards. During the interviews the experts were asked to sort the cards in order of importance and to comment on them. The aim was to test the assumptions of the researchers and to distinguish between more and less important problems. Table 2.1 provides an overview of the preliminary findings.

As seen from the table there are five barriers reflecting a relatively common agreement (ratings similar for three or more interviewees). Of those three were deemed important (complexity of adoption, cost of adoption, solving the problems with other means) whereas two were not (security concerns, SMC's inability to solve the underlying problem).

It must be reminded here that no statistically generalizable conclusions can be drawn on the basis of five interviews. In fact, as all interviewees were from Estonia, the pool of cases may be somewhat biased when it comes to assessing at least two barriers:

Table 2.1: Experts' assessments of the potential barriers to the adoption of SMC.

Potential barrier	**I1**	**I2**	**I3**	**I4**	**I5**
Unclear business value	?	?	✓	–	Important substantial problems
Too complex to adopt	✓	✓	✓	–	
Too costly to adopt	✓	?	✓	✓	
Lack of trustworthy partners	×	×	✓	–	
Security concerns	×	×	×	–	Problems of lesser importance pertaining to insufficient communication
Legal obstructions	×	–	?	–	
SMC's contribution to problem–solving unclear	✓	×	–	–	
Problem could be solved with other means	✓	✓	✓	–	
SMC does not help to solve the underlying problem	×	×	×	–	
Necessary information would not be accessed	?	–	✓	–	

Explanation of symbols used:

✓	Important problem
×	Unimportant problem
–	No (clear) assessment provided
?	Difficult to assess (e.g. importance of the barrier may be subject to change over time)

- Security concerns – many interviewees had already heard of SMC by means of prior dissemination activities by Cybernetica and expressed trust in the technology. However, this can be at least partly attributed to the fact that Cybernetica is well-known in Estonia and has established a certain prestige – a condition which may not hold to that extent on an international level;
- Legal obstructions – as the legal frameworks differ considerably from country to country the relative lack of legal obstructions in Estonia can not be taken as a sign that the barrier would not be salient at all. Moreover, in this interview round no law experts participated. Therefore, the perceived insignificance of potential legal obstructions should be taken with a grain of salt.

In addition to providing explicit assessments the interviewees also provided additional information on the baseline conditions for and potential barriers to the adoption of SMC. These can be thought of as two sides of the same coin: if a baseline condition is absent or is too high it becomes a barrier to adoption. Based on the interviews, five different types of baseline conditions and corresponding barriers to adoption were identified. These are summarized in Table 2.2.

2.1.1 Information

Having a sufficient overview of the nature and possibilities of SMC forms a major baseline for adoption. Two aspects were singled out by the interviewees as particularly important: first, a clear understanding of how the use of technology translates into increased profits/savings in monetary terms, that is, a clear business case. Second, a convincing use precedent must be present. In principle, this would enable the potential adopters to avoid (or to reduce) the cost of first-hand trial-and-error learning.

The interviewees also stressed potential downsides of insufficient information. Naturally, if a business case seems to be lacking or if it is not clearly communicated, the users see no reason to adopt the technology. However, the mere lack of use cases may also hinder the adoption as the majority of potential users could adopt a collective "wait-and-see" attitude thereby delaying further diffusion. According to one interviewee (I1) confusion about the underlying working principle may also promote irrational fears about data leakage despite the fact that SMC is designed to avoid it. Another interviewee (I4) argued that developers from the private sector are often unaware of the actual needs of the defence sector and therefore unable to propose sufficiently attractive applications of new technologies.

Table 2.2: Baseline conditions for and barriers to adoption of SMC.

BASELINE CONDITION	**IF BASELINE CONDITION FAILS**
1. Being sufficiently informed	1. Insufficient information/communication
Perceived business case	Lack of business case
Presence of other users (possibility of learning from experience)	“Wait-and-see”
	Insufficient knowledge
	Working principle is not understood
	Fear of data leaks “despite everything”
	Insufficient knowledge about the client’s actual needs
2. Presence of perceived need	2. Functional substitute already exists
Sufficient motivation to consider adoption	Research ethics
Decreasing the complexity of work processes	Trusted third party
Negotiations, organizing	Centralized collection of data guaranteed by law
Speeding up work processes	The perceived costs of inaction do not seem high enough
Monetary costs	
3. Presence of task-technology fit	3. Technology fits ill for the task
Data must be delicate enough	The demo does not offer enough attractive features
The need for data processing must be regular enough	May be unfeasible to use for big data
	Coordination costs become very high in case there is a need to agree on more than 100 definitions
4. Sufficient amount of resources	4. Lack of resources
The enterprise must be big enough to afford such data processing	Connectivity problems between various databases
Need to have enough time and resources to experiment with the technology in order to imagine possible uses in the first place	Not enough information on the process costs of enterprises
	Intelligence data may be difficult to formalize
5. Minimal need for change: good organizational fit	5. Avoidance of uncertainty
Low entry barrier, need for a standardized application	Lack of willingness to change existing work routines
	Desire to see the initial data
	Centralization of databases
	Desire to avoid the complexity and uncertainty involved in the adoption of new technology
Sufficient usability	Uncertainty about adoption costs
Compatibility with existing data analysis software	Implementation may demand too much organizational effort
	Implementation may be too expensive
	Low usability (high learning barrier)
Management support	

Targeting “technical visionaries” – people with technical and organizational know-how (including the standard procedures of data collection, handling and analysis), interest in new technologies, imagination and courage to envision new possibilities of use – was advised as a solution (I4). However, as the interviewee himself recognized people with such characteristics are few and far between and usually quite busy. This drawback was also experienced during the second round of interviews: although I4 agreed to help to establish contact with a few specialists in the defence and security sector eventually none of them answered the call for an interview.

The issue is accentuated by the fact that at the moment few practical real-life applications of SMC exist. The importance of this baseline was repeatedly established during interviews and the dissemination seminar in the University of Tartu devoted to presenting the practical applications of SMC. As evidenced in these meetings the people were often struggling to understand the basic working principle of SMC. Considering this, following recommendations are made:

- various dissemination activities by different channels should be continued to ensure awareness in public and private sector. Among those, simple means to explain and demonstrate the basic working principle of SMC should be found;
- the communication of real-life applications as well as their results warrants special attention here as these are likely to signal prospective users the potential benefits of adoption;
- effort should be devoted to establishing contacts with “technical visionaries” of target organizations. The latter may be viewed as entry points for further contacts. Moreover, in case SMC techniques spark their interest they may start to act as intra-organizational promoters of adoption;
- more attention should be turned to the profitability calculation of SMC applications.

2.1.2 Need

A perceived need constitutes the second baseline for adoption. The motivation to consider adoption obviously plays a major role. More specifically, however, the interviewees also stressed that the adoption of SMC should translate into less coordination efforts, less monetary costs or a smoother work process – but preferably into the combination of all those. Therefore, more attention needs to be turned to the extent to which the adoption of SMC would actually enable to optimize the work process instead of simply altering it. A case in point comes from the issue of input data quality: whereas in the case of centralized databases one party can clean all the data (I1), the situation is different for SMC in which case access to and comparison of separate data inputs is excluded by definition. Therefore, means to ensure the quality of input data need to be found. This, in turn, implies reaching certain agreements on how the data is to be cleaned by each party. Whether such coordination costs outweigh the ones of centralized data cleaning may not be clear at the outset.

Another trouble from the viewpoint of SMC adoption is that often existing functional substitutes work well enough. For example, one interviewee (I1) pointed to research ethics and mutual trust in academic community while the role of neutral third parties in data collection and analysis was stressed by many. In certain domains the incumbent third party is so well established that potential re-thinking of current practices is outright rejected. For example, an enterprise involved with market research declined the interview on the grounds that its data collection and analysis is already centralized. In the public sector the situation may be further institutionalized by legal means – for example, interviewees from the Statistics Estonia (I3) pointed out that they had legal rights to obtain data from various sources. In fact, SMC was even argued to be potentially harmful for the organization as this would give various parties a good pretext for not giving out the source data. Therefore, for the Statistics Estonia the adoption of SMC may result in a decreased access to data.

Finally, often the problems SMC could solve in principle may not seem that pressing at all and hence the costs of inaction are not that apparent. As expressed by an interviewee (I4): “One can always say that we did not need it before, we do not need it now". Hence new possibilities offered by SMC are not considered and existing organizational practices continue to be enacted.

Following from this, two recommendations are made:

- attention should be turned to mapping the changes in the work process introduced by the adoption of SMC across various application contexts. The substitutions of certain types of activities for another, the gained efficiency in the overall work process and possible trade-offs should be clearly outlined for the prospective users in order to facilitate adoption;
- considering the level of maturity of SMC solutions at the moment, less attention should be turned to targeting incumbent organizations characterized by one or more of the following conditions:
  - lack of particular internal or external pressure for change;
  - perception of SMC as potentially undermining the existing organizational competencies, resources and practices;
  - a need for substantial re-thinking of an existing business model and/or working practices in order to employ SMC.

2.1.3 Task-technology fit

An important aspect of technology adoption also concerns assessing whether the technology in its current form is suitable for the task at hand (and subsequently, whether it is the technology or the task itself that requires further adjustments). In this regard it was argued that the data should be delicate enough and the need for such data processing regular in order to consider the adoption of SMC (I1).

The interviewees also pointed out some instances in which SMC may fail to deliver on its promises. One interviewee (I1) worried about the lack of features on SMC statistics demo by pointing out that in her field of expertise (sociology) much more refined techniques are usually required. However, according to her, social sciences may constitute a somewhat special case since on average these tend to employ more sophisticated techniques of statistical analysis.

Another interviewee worried that in the case of big data SMC may remain too slow and costly (I3). A specialist from the domain of business process optimization (I5) referred to high coordination costs between parties when the amount of definitions to be agreed upon (e.g. the exact meaning of terms such as "delivery date") exceeds more than 100 items. An even more extreme example comes from the first round of interviews in which one interviewee turned the potential application of SMC on its head by suggesting to use it for dividing data between servers so that a successful attack on one server would not result in any data leakage.

What is at stake here is the question of the extent to which a technology can be accommodated to the task at hand. For example, in the case of statistics demo, the addition of new functionalities was not deemed particularly difficult by the development team. The cases of big data or business process optimization may present somewhat bigger challenges, however. And in some cases other ways of problem-solving may be more effective than SMC (e.g. encryption for the case of data protection outline above). As the focus of the UaESMC project is on practical applications it may therefore be useful to construct a hierarchy of use cases on the basis of the degree of task-technology fit and to focus on cases in which the fit is relatively big. This, however, also assumes an ongoing dialogue with prospective users about the desired functionalities (e.g. regarding the statistical techniques to be implemented).

Therefore, following suggestions are made:

- when developing SMC for the purposes of academic research the prototype should include a variety of latest statistical techniques, especially if the aim is to appeal to social sciences as potential users. A systematic mapping across various domains of science is also required to establish basic needs and priorities;
- when it comes to applying SMC in the private sector, the first "signalling" applications should focus on smaller, well-definable problems in order to avoid the issue of accumulating coordination costs;

- the development of SMC techniques should currently focus on cases in which the task-technology fit is either relatively big or at least relatively easy to achieve in the short term. To achieve this, ongoing dialogue with prospective users should be maintained.

2.1.4 Resources

Even when the prospective user is sufficiently informed, has a clearly perceived need for SMC and has defined the task in a manner for which the technology would be highly suitable, no adoption may still follow because of lacking resources. As one interviewee (I3) suggested, although the cost of using SMC is difficult to assess at the moment, only large enterprises may turn out to be wealthy enough to afford it. Another argued that the public sector would need some time and resources to experiment with the technology in order to even start imagining novel use opportunities. Failing that possibility, SMC could merely remain "an interesting and fun toy" (I4) the usefulness of which is acknowledged in rhetoric but not in practice.

But the cost of use is not the only possible barrier to adoption. Important obstacles may arise from the quality of input data itself as well as the connectivity issues between various databases (I1, I2). Intelligence data is often presented in a format which is difficult to quantify (I4). Moreover, only mature industries (e.g. automobile, electronics) have sufficient amount of information about various business process costs (e.g. the cost of buying from different suppliers) which others simply lack (I5). In other words, in a number of cases the adoption of SMC would seem to set too high requirements on the quality of input data.

The need for sufficient time and other resources for experimentation can be explained by the amount of cognitive work involved along a number of dimensions. The potential adoption of a new technology involves individual and organizational learning on a number of dimensions:

1. familiarizing oneself with the working principles of SMC (understanding what it is capable of doing in general);

2. locating some potential contexts of use for SMC, some real-life problems to be solved (connecting the abstract working principle to possible specific applications, connecting local problems to abstracted use cases);

3. assessing the feasibility of using SMC for the identified problem, filtering out some possibilities (i.e. assessing the the task-technology fit);

4. assessing the resources necessary to employ SMC, including the absorptive capacity for technological change of an organization in question (i.e. resources and organizational fit, see section 2.1.5);

5. trying out SMC pilots in practice, obtaining first-hand implementation experience, considering the need for further adjustments (possibly innovating during diffusion – see [Fleck 1994] for the concepts of "innofusion" and "learning by trying").

This situation is well acknowledged in the Strategic Niche Management literature [Kemp et al. 1998, Schot and Geels 2008]. Compared to mature and well-established technologies characterized by crystallized practices, routines and institutional support, emerging technologies are better thought of as "hopeful monstrosities" [Mokyr 1990] exhibiting low actual but high expected performance. The creation of niches serves to protect emergent technologies from direct market competition until the technologies in question have matured enough. Considering the amount of learning involved applications in the public sector and/or public-private sector cooperation may offer the best starting point for pilot applications. For example, one interviewee (I4) speculated that SMC could be used for assessing the frequency and diffusion of critical incidents (e.g. cyber-attacks on banks). Defence resource planning and pooling of intelligence data were also mentioned as possible domains of application (see Chapter 3 for more suggestions).

On the basis of the foregoing discussion three suggestions are made:

- in addition to focusing on limited and well-definable problems, business process optimization solutions should be targeted at large enterprises in mature industries;

- pilot applications of SMC should prioritize the public sector or the cooperation between private and public sector as the locus of experimentation until the technology has sufficiently matured;
- the niche provided by the public sector may enable to turn more attention to overcoming or decreasing the issue of insufficient degree of formalization of input data.

2.1.5 Organizational fit

The last baseline refers to existing organizational practices, e.g. relations between employees, formal and informal work procedures, technological infrastructure, rules and regulations etc. etc. It means that every organization has a set of routines, a "way of doing things" which new technologies are likely to disrupt – to a different extent, depending on an organization. According to the evolutionary theory of economic change [Nelson and Winter 1982] organizations try to maintain their routines. If problems occur, the search for solutions starts around initial organizational and technological competencies. If the problem is not solved, the scope of the search gradually widens. Therefore, in order to facilitate the adoption of new technologies the existing organizational routines should be minimally altered (or alternatively, focus should be directed to organizations involved in wide-scale search activities).

This argument provides a context for the following suggestions. One interviewee (I3) indicated that the entry barrier to SMC applications should be low so that a prospective user would not have to spend too much time thinking on "servers, software and data preparation". Another observed that in the military sector technological changes are often accompanied by changes in management (I2). Yet another (I1) pointed out that the statistics application should have similar functionalities and look roughly the same as existing data analysis software. All these arguments point to the need to keep the learning barrier low. One way of achieving this would be to offer standardized configurations of SMC application in contrast to current ad hoc pilot applications. Whether and when this standardization can be achieved and what forms it would take are questions currently open for debate.

It is also interesting to note that this turned out to be the only baseline condition for which the question of usability (I1) was raised. Indeed in general the interviewees showed little interest towards statistics demo having considerably more to say about the potential application contexts of SMC instead. Therefore, it may be assumed instead that the question of usefulness of SMC was deemed more important than the question of usability of potential applications.

The desire to avoid uncertainty was listed as an important obstacle to technology adoption. In many cases the organizational routines can be so well embedded and the organization itself doing so good that deeper reflection on the meaningfulness of current practices never emerges. For example, one interviewee (I3) claimed that scientists often want to see the initial data but fail to offer convincing reasons as to why this is necessary. At the extreme the presence of organizational routines may result in a failure to take advantage of the distinctively novel possibilities offered by new technologies (I4). This observation ties in with the above suggestion to avoid targeting incumbent organizations for the time being (see Section 2.1.2).

Building on this, following suggestions are made:

- future efforts should be devoted to probing the feasibility of developing standardized configurations of SMC applications which, in turn, would enable to keep the entry barrier low thus facilitating further adoption;
- recent management change (at least in the public sector) may be a facilitating organizational factor for attracting interest in SMC pilot applications;
- the statistics applications of SMC should resemble existing data analysis software in terms of functionality and design;
- in general, however, considering the level of maturity of SMC solutions attention to ensuring the usefulness of SMC should take precedence over usability concerns.

Chapter 3

Potential application contexts of SMC and existing practices of data handling

The analysis in section two identified five baselines conditions of and corresponding barriers to adoption. However, the potential application contexts themselves were not specified. Also, no explicit attention was turned to existing practices of data handling and the extent to which SMC can (or cannot) complement or replace them. This section takes up these two tasks. The first part of the chapter presents an overview of various potential application contexts of SMC as imagined by the interviewees. The second section attempts to assess SMC's potential in relation to existing practices of data handling.

3.1 Potential application contexts of SMC

The interviewees were asked to imagine possible uses for SMC, that is, to localize the general working principle to a specific empirical problem. Owing to the amount of cognitive work involved (see Section 2.1.4) the results were expectedly mixed, ranging from a clearly defined problem to a general domain of application. The following list attempts to give an overview of the application contexts mentioned:

- **Private sector**
  - Market analysis (I5)
    - Especially positioning the enterprise in relation to the market average (I5)
  - Fusing the business processes of enterprises for collaborative purposes (I5)
  - Processing of e-receipt data to draw conclusions about consumption on average (I5)
  - Comparison of data in professional associations (I3)
- **Research**
  - Academic statistics (I1)
  - Analysis of sensitive personal data (medicine) (I1)
  - Providing an additional layer of security for the respondent in online polling (I1, I4)
    - Especially to include the reliability of answers during times of social upheaval (I4)
- **Public sector**
  - Working with official secrets (I3)
  - Enabling the citizen to connect the data from various official databases directly (I1, I3)
    - X-road applications (I1, I3)

- **International cooperation**
  - Statistics between states (I3, I4)
  - Comparison of intelligence data of different states (I2, I4)
    - Quantified data (e.g. size of units, accuracy and flight distance of missiles etc.) (I4)
  - Directing and coordinating unmanned planes and vehicles (I2)
- **National defence**
  - Defence resource planning (I2, I4)
  - Planning of emergency situations (I4)
  - Planning of critical services (I4)
  - Assessment of the frequency and diffusion of critical incidents (I4)
    - Security of mobile communications networks (I4)
    - Cyber-attacks on banks (I4)
  - Logistics (I2, I4)

3.2 SMC and existing practices of data handling

The data protection has already been done in multiple ways. The following section elaborates on the key issues outlined in our interviews. The last section attempts to analyse SMC's potential to impact existing practices

3.2.1 Elaboration of existing practices for protection of data in case of data analysis

Anonymisation. This could have been done prior giving data over to the researcher or by research team themselves, where by signing contracts of keeping the privacy of the participants, researchers code of ethics has been the key. Research team assigns a member amongst themselves to anonymize the data. The other option is that anonymous data is given to the researcher. This option often limits the kind of research doable, as in many cases, the data is filtered severely before handing it over to the researchers. This means that sometimes many cases are eliminated from the analysis that could have been kept in with SMC and eliminating results based on these analysis only for certain computations. The strength of this kind of approach is that the researcher has full access to the data approved to them and they can see the data and potentially correct the errors in it (or delete the erroneous cases). The critical part is heavy reliance on the trustworthiness of the individual researchers.

Keeping data and analyser separate. This practice has meant that operations on the data are conducted only by authorised personnel, meaning that privacy is again protected by contracts and research ethics, but the number of people having access to the data is even more limited. If researcher has limited access to the database, they can compose queries either based on prior knowledge or on sample anonymised mini-versions of the database. Queries are then sent for processing and only results are presented. This is in some sense very similar to SMC, but instead of the SMC algorithms, the statistician is manually processing all such queries. This can be very time-consuming (sending queries, waiting for results and improving queries can be frustrating experience) and it can be also very expensive, the costs driven often also by the monopoly position as the authorised handlers of the data can only be very limited in number. This also limits the researchers possibility to check the quality of the data, especially to find data entry mistakes (note: this can be challenging in any kind of data and access possibilities). The system is heavily reliant on trustworthiness of the individual researcher.

The military sub-case of such occasion is when routing information is not shared by third party organisations at all, but rather the secure pathway is only demonstrated case-by-case. This can be resource-consuming as the security patrols need to be spared to demonstrate the way, also this implies trust between different levels of allies as the pathway information is fully disclosed to at least some people.

Third-party solution. This is a version of the above, where the third party is trusted with the data and operations. Often such third party can also collect data on behalf of the contractors and their service is the whole package of research. This can be also time-consuming and costly, also the third-party needs to be trustworthy and uninterested in leakages.

Limiting research problem. This is a version of privacy preservation where research questions are formulated to match the real-life access, not necessarily to the needs of research or policy making. The benefit, you do what you can do, what you have access to. The drawback, you might miss out on important analysis and important potential of data collected.

3.2.2 Elaboration of existing practices for protection of data in case of data sharing

Secure is only the pathway to the data (e.g. VPN is used to access or designated space or designated computer is used for calculations), but data handling is not secure and trust is placed in the ethical consideration and contracts with the researcher. A military sub-case of such practice is when data shared during the operations is kept in separate channels/screens/networks. It still meets and privacy is lost in case of analysis, and trust in personnel needs to be very high.

Sharing of data is made legally obligatory. For instance in logistics or resource management (e.g. storage of military critical food resources in different locations over the country among competing private enterprises). In case of different non-military parties information can be requested by law and the military then acts as a trusted third party with hopes that the information will not be leaked to the competitors. The benefit of this solution is that it is a long-standing tradition. The drawback can be inefficiency of such system and reluctance of different parties to share full information. Often also severe delays will need to be built in to such system making the info less operative.

3.2.3 Summary of existing practices with evaluation of potential impact of SMC

The application of SMC entails complementing or replacing various existing practices of data handling. We have attempted to summarise the existing practices discussed by the interviewees inn a table, some of them are applicable across multiple potential application contexts.

Table 3.1: Comparison of pros and cons of existing practices with evaluation of SMC's potential to impact these practices.

EXISTING PRACTICES	**PROS**	**CONS**	**SMC's POTENTIAL TO IMPACT EXISTING PRACTICES**
Data analysis			
Anonymisation prior to giving data to researcher	Secret is kept through professional ethical standards	• Very dependent on individual ethical integrity • Can involve extra work and be time-consuming • Individuals may still be identified through auxiliary data	+ No prior anonymisation needed − Introduces computational overhead
Anonymisation by research team	Secret is kept through professional ethical standards	• Very dependent on individual ethical integrity	+ No prior anonymisation needed − Introduces computational overhead
Data is filtered as well as anonymised	Very unique cases are excluded from the analysis	• Can limit the type of analysis • Can exclude too many cases in an attempt to avoid back-tracing the individual	+ No prior anonymisation needed + No filtration needed − Introduces computational overhead
Keeping data and analyser separate	Secret is kept through professional ethical standards	• Very dependent on individual ethical integrity • Can involve extra work and be time consuming • Can be costly in terms of resources • Needs authorised personnel able to devote enough time	+ Responsiveness of SMC is higher than that provided by authorised personnel

EXISTING PRACTICES	**PROS**	**CONS**	**SMC's POTENTIAL TO IMPACT EXISTING PRACTICES**
Data operations are performed by third party	Secret is kept through professional ethical standards	• Very dependent on individual ethical integrity • Can involve extra work and be time consuming • Can be costly in terms of resources • Needs authorised personnel able to devote enough time	+ Responsiveness of SMC is higher than that provided by authorised personnel
Limiting research problem	You only work with what you have access to	• Research is done only on what can be done with existing tools, excluding potential breakthrough discoveries or data-based policies	+ More data can be made available for analysis – Individual values cannot be accessed (raw data cannot be seen)
Data sharing			
Secure pathway to data, full disclosure	While access to data is kept secure, secret is kept through professional ethical standards	• Very dependent on individual ethical integrity • Can involve extra work for those authorising access • Can be costly in terms of resources (providing special working space or maintaining secure channels)	+ Not dependent on individual ethical integrity to such high extent – Individual values cannot be accessed (raw data cannot be seen)
Data sharing can be made obligatory by law	Legislatively everyone has to share their data no matter the privacy concerns	• Can reduce willingness to cooperate • System can be slow or inefficient	+ As individual values cannot be seen, there is no incentive to withhold information

Bibliography

[Fleck 1994] Fleck, J. 1994. Learning by trying: the implementation of configurational technology. *Research Policy* 23 (6): 637-652.

[Kanger 2013] Kanger, L. 2013. D6.1 - Addressing societal concerns on legal and privacy issues in ICT-related projects. Public deliverable, The FIRE Project (call FP7-ICT-2011-8 of the ICT Work Program 2011/12). Available online at: `http://www.trustworthyictonfire.com/outcomes/public-deliverables?download=4:d6-1`

[Kemp et al. 1998] Kemp, R., Schot, J., and Hoogma, R. 1998. Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management. *Technology Analysis and Strategic Management* 10 (2): 175-196.

[Mokyr 1990] Mokyr, J. 1990. The Lever of Riches. NewYork: Oxford University Press.

[Nelson and Winter 1982] Nelson, R. R., and Winter, S. G. 1982. An Evolutionary Theory of Economic Change. Cambridge, MA & London: The Belknap Press of Harvard University Press.

[Schot and Geels 2008] Schot, J., and Geels, F. W. 2008. Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technology Analysis and Strategic Management* 20 (5): 537-554.