Towards Software-Agent Enhanced Privacy Protection

Addi Rull, Ermo Täks and Alexander Norta

Abstract The ability to control the use of personal information is part of the right to privacy. With higher digitalization than ever, the lack of control is an essential privacy issue discussed extensively. Estonia is a unique society in terms of the highest level of digital public services available for a citizen, enabled by the omnibus X-Road infrastructure and personal identification solution developed some time ago. The technology has certain security elements essential for the protection of privacy. However, there are no technical measures to achieve better control over the subsequent use of personal information once it has been obtained from a database. We suggest a task-oriented approach to be exercised in the retrieval of personal information. This can be accomplished by using agent technologies. The aim of the technology is to control access to personal information so that a public servant only obtains a citizen's information limited to the performance of her particular task. In other words, the information system, with the help of a software agent, shall supply a public servant only with the information necessary for performing a decision concerning one citizen. Such enhanced control over the use of personal information contributes to better privacy protection. The prospect addresses the prevention of the misuse of personal information as well as the enforcement of data protection laws. The chapter is an introduction to the discourse of agent technologies and law together with a conceptual example for a possible technological solution in the police work of Estonia.

A. Rull (✉)
Tallinn Law School, Tallinn University of Technology, Tallinn, Estonia
e-mail: addi.rull@ttu.ee

E. Täks · A. Norta
Faculty of Information Technology, Department of Informatics, Tallinn University of Technology, Tallinn, Estonia
e-mail: ermo.taks@ttu.ee

A. Norta
e-mail: alex.norta@gmail.com

T. Kerikmäe (ed.), *Regulating eTechnologies in the European Union*,
DOI 10.1007/978-3-319-08117-5_5

1 Introduction

Continuous increase of the use of public databases poses new challenges. The European agenda to re-use information collected into public databases is the policy aimed at utilising information for the creation of economic and social benefits of a society.[1] The expansion in the area of public services immediately raises the question whether the protection accorded to individual privacy is adequate. Limiting the amount of data to be collected and used is a principle by which the protection of privacy can be better ensured, but there is a growing need for data in order to be able to create new public services. With the amount of data growing it is equally important to focus on the quality of data stored in databases and on the way it is used by public officials when they make decisions concerning an individual. In other words, it is important to consider the relation of the information available in a database to the decision to be made by an official. If an official has access to information which is unnecessary for the decision-making or it is wrong, then this may have unwanted influence on the decision. Although there are data protection laws, the access to information and how it is being used is often difficult to control.[2] This is not because of the lack of rules, or inability to trace enquires by log files, but because there are so many databases, so many different tasks to be performed by so many officials, and no efficient way of controlling the access to and the use of data. Daniel Solove wrote about Computer Databases and Metaphors of Information Privacy in 2001 and called it the aggregation problem.[3] He used Kafka's The Trial metaphor to explain that the information about a person circulating in databases lives the life of its own which is difficult or impossible to control. This may bring about unwanted decisions affecting the life of an individual.

The problem of aggregation can be explained by several scenarios. An official may have access to information unnecessary for the decision-making, but the decision is affected by this information, and this decision is the source of information for another decision made by another official. Or wrong information about a person may be accidentally recorded in a database, which is a source of information for several other databases, and several decisions are being made based on this information and subsequently recorded in various databases. Aggregation is caused by automated, semi-automated or human actions in the decision-making process. An action as a part in the process is not a problem if it can be corrected before the decision is affected, otherwise the sequence of actions causes the aggregation.

One way to tackle the problem of aggregation is to sustain better control over the use of personal information. Alan Westin wrote in 1970 that "[p]rivacy is the claim of individuals […] to determine themselves when, how, and to what extent information

[1] E.g., Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information; see also HOMER Report 2013, Janssen and Dumortier 2003, pp. 184–201.

[2] See Männiko 2001 for a comprehensive overview of privacy and data protection laws in Estonia.

[3] Solove 2001, pp. 1413–1434.

about them is communicated to others".[4] His description of privacy is more relevant than ever, given the vast development of information technologies over past decades. The discussion in the following does not seek to define privacy or to analyse data protection laws, which is an ongoing debate and research amongst many academics.[5] Instead, the motivation of the discussion is to offer novel ideas how better control over personal information could be achieved with the help of 'agent technologies'.[6] The reason to explore technologies subsides in the conviction that there are not so many other efficient and purposeful options left to solve the problem. There is the omnibus data protection regulation at all levels: international, regional (EU) and domestic. Further harmonization may be imminent and the continuous improvement of laws remains to be the objective,[7] but it does not seem to be the solution for the enforcement problem. Every day numerous misuses of personal information which people do not know about take place. If one knows about a misuse and decides to pursue a case, then most of the time the cost of proceedings calculated in terms of time, money or stress exceeds possible moral or pecuniary relief granted. For this reason most people do not react to relatively minor acts of unnecessary uses or misuses of personal data.

The following provides an insight to the use of a database in the police work in Estonia. We describe problematic uses of personal data and discuss violations based on the case law. The underlying problem is the lack of control over the use of personal data. We suggest using agent technologies independently or in combination with the legislation. The discussion of these possibilities is on a conceptual level. Practical solutions are yet to be developed.

2 Public Registers in Estonia

There are approximately 600 public registers in Estonia.[8] Population Register is one of the main databases containing information about citizens and foreigners living in Estonia.[9] The increase of the number of enquires made to this register from

[4] Westin 1970, p. 7.

[5] The concept of privacy has been discussed by many know authors e.g., Alan F. Westin; Judith Jarvis Thomson; Richard Posner; Robert Bork; William Prosser; Ferdinand D. Schoeman; Raymond Wacks, Daniel J. Solove.

[6] Nwana 1996, pp. 1–40. A software agent is a computer program that acts for a user or other program in a relationship of agency, which derives from the Latin *agere* (to do): an agreement to act on one's behalf. Such "action on behalf of" implies the authority to decide which, if any, action is appropriate.

[7] See e.g., Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) 2012/0011 (COD), Brussels 25 Jan 2012.

[8] Administration System for the State Information System RIHA https://www.ria.ee/administration-system-of-the-state-information-system/. Accessed 22 Mar 2014.

[9] See information about Population register at https://www.siseministeerium.ee/35796/. Accessed 22 Mar 2014.

42.1 million in 2012 to 55.8 million in 2013 indicates the growing use. According to the Estonian Statistics Agency, the number of people living in Estonia is 1,311,870, which means more than 40 enquiries per person. Institutions that enquire most include the Police and Boarder Guard Board, notaries, courts and local administrations.

A person can check for enquiries by entering www.eesti.ee and making the enquiry about enquiries made to certain registers, for instance, the registers related to the Citizenship and Migration Board. The enquiry shows when the enquiry to a certain register happened, the number of the file this enquiry relates to, the name of the institution, the ID and the position of the enquirer. In some instances, a person can easily assume reasons why enquiries have been made, because this may be related to the visits or applications made by the person to public institutions.

Many registers retrieve information from the Population Register. Chapter 4 of the Population Register Act provides the composition of data recorded in the Population register.[10] Name, date of births, citizenship, marital status, postal address, or person's legal capacity, ethnic nationality and identification documents are examples of data contained in the database. The population register contains basic information about a person. Other registers contain specific information depending on the purpose of the register. For instance, Car Register contains information about registered cars in Estonia and the E-health system contains medical information about diseases diagnosed, medicine prescribed, etc.

Some registers like the Population Register, the Environmental Register, the Marital Property Register, and the Register of Economic Activities are regulated by laws. Most registers are established by regulations on the bases of delegations prescribed in laws. For instance, the statute on the maintenance of the police database, which is called POLIS,[11] is established by the Minister of Interior by regulation. This delegation is derived from the paragraph 13 subsection 1 of the Police and Boarder Guard Act.[12] These legal acts are publicly available and provide detailed information about public databases. The transparency decreases the risk of the arbitrary use of data by the state. In case the violation of the use of personal data occurs then it is possible to resort to data protection and penal laws for a remedy.

At least two databases in Estonia cannot be studied on the same bases as other public databases. There is nothing that can be learned about the database by the Estonian Internal Security Service called KRISTI. The database was publicly mentioned by the parliament member who was the member of the Security Surveillance Authorities Select Committee.[13] This is a parliamentary committee which exercises supervision over the legality of surveillance activities of the

[10] Population Register Act RT I 2000, 50, 317; RT I, 22 Nov 2013, 2, Chap. 4.

[11] Police and Boarder Guard Act RT I 2009, 26, 159; RT I, 02.07.2013, 18, § 8 (2).

[12] Ibid. § 13 (1).

[13] Security Surveillance Authorities Select Committee http://www.riigikogu.ee/index.php?id=42701&parentid=34615. Accessed 24 Mar 2014.

Security Police Board as well as the Police and Boarder Control Board. The existence of the database is certain, because one of the main functions of the Estonian Internal Security Service is to collect information for the prevention and combating activities directed against the state.[14] Most probably, this is not the only classified database in Estonia. We assume also that the Ministry of Defence or the Defence Forces have a database for military intelligence purposes.

Another database, occasionally discussed in public media, is called KAIRI.[15] The Estonian Police and Boarder Guard Board operates KAIRI that is linked to the general police database called POLIS. KAIRI is not entirely classified, but the documents regulating its use are not accessible to the general public. Nevertheless, several court cases include references to these documents. There are reasons for keeping the rules of the use of the database and its content secret, because it is being used in daily police work. KAIRI contains data collected during crime investigations and surveillance operations.

2.1 Database KAIRI

One of the problems identified in cases related to KAIRI shows how information retrieved from the database is not used for the performance of a public duty, but for a personal interest. Yet, instances discussed amongst people who have been stopped by the traffic police over the years suggest that sometimes the decision-making based on the information available in the database may be biased. Also, the quality of information collected into the database has been criticised. Similar problems may arise with other databases.

Every year, thousands of people are stopped by the traffic police, because they have been found guilty for misdemeanour. The mistakes are mostly over speeding, driving without fixing a safety belt or violations related to different traffic signs. The fines are calculated on the bases of fine units whereby one unit equals four euros. If a person is driving a car with the speed of 65 km/h in the area where the limit is 50 km/h then the excess speeding up to 20 km/h is punishable by a fine of up to 30 fine units i.e. 120 euros.[16] A police officer has other options. He may warn the driver and let him go. Cautionary fine is also a possibility, but it is applied when the over speeding is registered by speed cameras.

[14] See information available at https://www.kapo.ee/eng. Accessed 24 Mar 2014.

[15] Authors are responsible for any incorrect assumptions made in regards to the database KAIRI, because we have had no chance to study the database and the documents related to it. Information related to the use of KAIRI is a state secret. However, any possible mistakes discovered would not make our suggestions and conceptual discussion obsolete, because the ideas we propose for achieving better control over the use of personal information can be applied in multiple instances in different spheres of life in relation to other databases.

[16] Traffic Act RT I 2010, 44, 261; RT I 14.02.2014, 3, § 227 (1).

Oral warnings and cautionary fines are not recorded in the Punishment Register. If a person is punished for over speeding, then the misdemeanour is recorded in the register. Misdemeanour is substantially different from criminal offence. Paragraph 3 of the Penal Code provides that the punishment prescribed for a misdemeanour is a fine or detention and for a criminal offence it is imprisonment or pecuniary punishment.[17] The future of a person is only affected when the punishment is recorded, because it has a bearing in instances where a person is found guilty for an offence the second time. Often the past behaviour is taken into account and the second punishment is more severe. Punishments recorded in the Punishment Register can also be seen via KAIRI, probably, because the data is regularly synchronised.

A police officer uses a radar speed gun and stops a car when the over speeding is registered. He walks to the car and asks the driver to come to the back seat of the police car. A small mountable device with a screen similar to GPS car navigation device below the radio or the air conditioning unit enables the person sitting in the back seat to peak and see what is on the screen. The police officer uses the device to access KAIRI or POLIS. Firstly, the driver has to be identified. The information on the driver's identification document is compared with the information retrieved from the database. Then the police officer raises a discussion about the over speeding, asking questions in order to understand why the speed limit was not followed. Did the driver know that he was over speeding, why did he choose to drive at the speed above the limit, did he see the traffic sign, etc. The police officer has to make a decision based on the information received from the driver, taking into account the past behaviour. Often the police officer reminds the driver of the punishments recorded in the past. The record of punishments is the important source of information, because it is reliable. The truthfulness of the answers of the driver is more difficult to evaluate.

The police officer exercises the right of discretion to make the decision. The driver may be given a warning or a fine in between 12 and 120 euros[18] if the over speeding was up to 20 km/h in the area where the speed limit is 50 km/h. If the driver does not have an earlier record of misdemeanours, then he may be given a warning or a low fine. In case the driver has a record of over speeding or something else, then higher or maximum fine is an option.

A problem arises with the information accessible to the police officer. The full record of criminal offences and misdemeanours is accessible, whereas the limitation period of the execution of the decision in regards to misdemeanours is one year.[19] The Punishment Register Act provides that after a year the record of the misdemeanour has to be deleted from the register and archived.[20] Similarly, different types of criminal offences have archiving deadlines.

[17] Penal Code RT I 2001, 61, 364; RT I 14.01.2014, 10, § 3 (3), (4).

[18] Ibid. § 47 (1).

[19] Ibid. § 82 (1) 3.

[20] Punishment Register Act RT I, 21 Mar 2011, 3; RT I, 13 Dec 2013, 15 (hereinafter as Punishment Register Act) § 24 (1) 1.

Punishment Register was made publicly accessible in 2011 with exceptions related to minors and archived offences. The archiving procedure was further improved in 2013. The rule was that the calculation of one year after which a misdemeanour would have been archived elapsed and started again when the same person received another punishment within this timeframe before a year passed. The dependency of one misdemeanour to another in the calculation of one year was abolished. The Punishment Register held approximately 600,000 records of 250,000 thousand people excluding criminal offences of 22,500 people. Approximately one third of the adult population had public records of offences. Since misdemeanours older than one year are now archived, the number of people with punishments is considerably lower. It is unclear whether the whole catalogue of offences is still synchronised with KAIRI or the archive is excluded.[21] If the archive of misdemeanours cannot be studied via KAIRI, which we do not believe to be the case yet, then this only solves a part of the problem. Records of criminal offences have longer archiving deadlines, and if these records can still be studied in instances when this information is unnecessary for the decision-making, then biased decisions cannot be ruled out. Paragraph 20 of the Punishment Register Act regulates the right to obtain archived data. Clauses 3 and 7 of the subsection 1 of the paragraph 20 suggest that data can be obtained for the purpose of criminal investigation or surveillance activities.[22] KAIRI is the database used for these purposes.

The question raised is whether the police officer who retrieves the record of punishments via KAIRI disregards the information he sees about misdemeanours older than one year? It is known from the practice that police officers raise the discussion about older punishments which should not have the bearing on the decision. Even if they do not discuss the relation of past offences to the situation yet to be decided, it is impossible to know what they think while screening the history, including possible criminal offences.

Suppose a driver has records of punishments for over speeding older than one year and he is caught again. The police officer checks the record and must understand as if the driver has no valid punishments for over speeding. Shall he give the warning or a low fine, which seems appropriate, or a high fine, because in his mind the information seen suggests that the driver has the recidivist behaviour and deserves to be punished?

[21] Since authors did not have a chance to study the architecture of the police databases, then it is unclear how exactly data collected from different databases is recorded in KAIRI. It may be that data is not recorded in KAIRI, but KAIRI functions as an access filter showing data which is available in different databases. Nevertheless, authors are sure that somewhere in this chain the cache of data, which is a copy of the data originally recorded in different databases, is created. Therefore, this does not change the objective of the discussion of how it would be possible to eliminate the creation of cache and how it would be possible to achieve better control over the use of personal data regardless of the possibility that our assumptions about the system architecture may be wrong.

[22] Punishment Register Act. § 20 (1) 3, 7.

KAIRI has another problem related to the quality of data recorded there. KAIRI includes a wide variety of information about a person. A lot of it is retrieved from other registers. For example, it includes information about car registration, immovable property, family relations, offences, etc. If this information had to be enquired from different databases each time there is a need for that, then a lot of time may be lost in critical situations. KAIRI also includes information that does not exist in other databases. Surveillance information is recorded there, but also information about ties a person has with criminals or people who are suspected of criminal activities. Even gossip and unverified statements are known to be found there. The collected information in KAIRI is categorised based on how trustworthy it is. Unreliable information existed in 2011 when the Security Surveillance Authorities Select Committee raised the issue of legality and purpose of such information and ordered the police to clear the system of unreliable data. The extent of the improvement remains unknown. The question remains what exactly is the purpose of recording unreliable data? Hints are useful when the police officer is looks for leads to work on a case. However, wrong information is detrimental to the investigation of a case.

2.2 Case Law

Police officers have misused personal information in several instances. Three Supreme Court cases deal with the fact that personal information retrieved from KAIRI was not used for the performance of public duties, but it was transferred to people outside the police force who did not have the right to access the information.[23]

Directive of 14th of June 2004 by the Director General of the Police and Boarder Guard Board established the procedural rules for KAIRI. The Director General exercises the right to regulate the collection of data on the bases of the Government of the Republic Act.[24] The non-disclosure obligation of the police officer is specified in job descriptions and in § 26 of the Personal Data Protection Act, which requires persons who process personal data which become known to them in the performance of their duties to maintain the confidentiality.[25] The violation of the non-disclosure obligation is criminalised. § 157 of the Penal Code provides that the disclosure of personal information obtained in the course of professional activities is punishable by pecuniary punishment or 3 years imprisonment.[26]

The definition of the private personal data is provided in § 4 of the Data Protection Act. The Supreme Court has stated that the definition of private

[23] Estonian Supreme Court cases 3-1-1-81-08; 3-1-1-25-12; 3-1-1-56-13.

[24] Government of the Republic Act RT I 1995, 94, 1628; RT I, 27.12.2013, 33, § 73.

[25] Personal Data Protection Act RT I 2007, 24, 127; RT I, 30.12.2010, 11, § 26.

[26] Penal Code RT I 2001, 61, 364; RT I 14.01.2014, 10, § 157.

personal data should not be treated narrowly. It includes any factual information concerning the life of an individual. For instance, one's address is a personal data although it may be possible to obtain it from publicly accessible registers such as the Register of Economic Activities in Estonia.[27] Either the person has given the consent to make the address publicly available or it is a precondition for a certain activity. Economic activities require that a person has publicly available address, as it helps to achieve certainty in the business environment. Another example is the Traffic Registry. It is not possible to identify the owner of the car in Estonia, because the car registration information is not publicly available. Car registration and address information can be obtained from KAIRI, but the address may be publicly available in other databases. If address is copied from KAIRI not for the performance of the duties of the police officer, but for someone who does not have access to the database, then data protection principles are violated. The fact that the address can be obtained from other sources does not wave the non-disclosure obligation of the police officer.

One Supreme Court case tells about the assistant police officer who asked the police officer to provide him with the information about two people. The information included addresses, car registrations, records of misdemeanours and criminal offences. The assistant police officer could not have obtained this information by himself except addresses perhaps. His position did not allow him to access KAIRI, because the assistant police officer is not the member of the police. People who apply to work on voluntary bases together with the police in street patrols become assistant police officers. Rights, obligations and activities of the assistant police officer are regulated in the Assistant Police Officer Act.[28]

In another case the husband of the police officer received an e-mail with the information about his co-workers. The intention of his wife working in the police was to provide him with sensitive information about these two people so that they could be fired if necessary. Again KAIRI was used as the source of information.

The most recent case was about the police officer who used her position in the criminal investigation of theft to give information to victims who wanted to track down and to threaten thieves as a way to force them to pay. The police officer shared the information obtained from the database and the information collected during the investigation with victims.

In general, it is fair to assume that a police officer does not misuse personal data, because he is bound to the regulation prescribed to ensure the protection of privacy. The referred case law shows that the misuse of personal data is argued in the context of complex legal instruments including laws, regulations, directives and job descriptions. The police officer must have the comprehensive understanding that he is the processor of personal data who is under the obligation not to use personal information for any other purposes except for the performance of a public duty. The quality of education and the awareness of the regulation help the police

[27] Register of Economic Activities http://mtr.mkm.ee/default.aspx. Accessed 19 Mar 2014.

[28] Assistant Police Officer Act RT I, 20.12.2010, 1; RT I, 30.12.2011, 58, § 2 (1).

officer to achieve the required understanding. On the other hand, cases of misuse of personal information require the expertise of highly qualified prosecutors and attorneys to argue a case. Evidence has to be collected before that. And, most importantly, there is no case without a victim. A person who has learnt about the misuse of personal data may initiate the case or the violation may be discovered by internal audits. No one knows how many violations have taken place, which victims do not know about, nor have internal audits discovered them.

3 Purpose of Software Agent

Traces to prove the use of information in cases discussed were log files and other evidence left behind. Every enquiry leaves a footprint of a person who made it. Traceability is a feature which helps to enforce the protection of privacy, because the probability exists that a person who made an arbitrary enquiry may be discovered. The problem with such a system is that it has minimal knowledge about the intended use of the information enquired in order to decide whether the person looking for the information should be granted access to it or not. Police officers in different positions have different access rights to the information in KAIRI. Classified surveillance information is available to approximately one third of 4000–5000 officials who can use the system. The system identifies the individual, recognises the level of access and provides the information accordingly, but the subsequent use of the information is not controlled by the system. The possibility that a system would exercise full control over the use of the information is difficult to imagine in the police work. Instead, the more sophisticated approach to the task oriented use of the information should be developed.

The described misuses of personal data in the investigation and the traffic patrol work are the result of intentional or subconscious human error. The traffic patrol may be inclined to make a biased decision, because the full history of past mistakes of the driver can be studied via KAIRI. As a result, a severe punishment may be a subconscious decision as a reaction to recidivist behaviour. The current information system does not address these issues. We suggest to explore the idea of using the agent technology which could stand in between the police officer and databases in order to aid the police officer in the decision-making process, yet, at the same time to achieve the enhanced protection of privacy thro' more sophisticated way of processing personal data.

The primary task is to solve two basic issues: (1) how the agent technology can regulate the flow of information so that it only gives the police officer the right type of information required for the performance of a certain task; (2) and how such technology can enforce legal norms and react to regulatory changes. Consequently, the successful performance of these tasks will decrease the number of intentional or subconscious human errors as it will correct the mistakes resulting from the imperfect knowledge of the police officer about data protection law. Furthermore, the decrease of the number of misuses of personal data shall affect the resources spent for internal audits and court cases. Overall, the fundamental right to privacy is protected better than today.

Two tasks can be accomplished together or separately. The combined implementation of these tasks would require a systemised approach to semantics of law and computer linguistics in order to be able to develop a methodology that enables to build a link between the legal system and a particular task to be performed by a computer programme.[29]

3.1 Legal Software Agent

Agent technologies have been developed over past decades. The notion of "agent" was first used by Hewitt in 1973.[30] The technology has been used for various purposes, for example, to predict the perception of consumers before the launch of new consumer goods.[31] Agents assimilate humans, indicating their consumer preferences and behaviour, and in this way contributing to the making of sales forecasts for manufactures. Another example is the IBM solution called Watson, which beat the best player in the quiz competition in Jeopardy TV show in 2011. Intelligence demonstrated by agent technologies renders remarkable results, which in some aspects are better or comparable to the best of human beings. Artificial intelligence has become possible with agent technologies. The same cannot be stated for the automated decision making applications developed in the field of law. In 1977 Harvard Law Review published the article about legal reasoning and artificial intelligence written by L. Thorne McCarty who explained how tax issues related to corporate reorganisation can be resolved with the help of a computer programme called TAXMAN.[32] Little progress has been made ever since. Tax authorities use algorithms to check the payment of taxes. Systems have been programmed to recognise irregularities or to react to certain predetermined conditions which trigger the automated processing of administrative acts, for instance, issuing the administrative act which informs a person about the start of the tax investigation procedure. Fines sent to drivers who have been caught for over speeding by speed cameras is also an example of automated enforcement of law. Still, most processes are semi-automated, requiring human intervention in a certain stage of a process.

The problem is that computer programmes do not have the capability to adjust themselves automatically to the regulatory changes. For instance, in Estonia the V.A.T. tax was changed from 18 to 20 % in 2009 in order to raise funds for the state budget in the midst of the financial crises. The change in the tax law was announced and implemented in a short notice. The business sector could not adapt to the new regulation without the extra costs that occurred with the implementation of changes into software and the changing of price labels in stores, etc.

[29] See e.g., Nyman-Metcalf and Täks 2013, pp. 1–30.

[30] Hewitt et al. 1973, pp. 234–245.

[31] Gowda 2008, pp. 246–251.

[32] McCarty 1977, pp. 837–893.

The Chancellor of Justice argued that the state violated the *vacatio legis* principle by not giving ample time for the business sector to comply with the new regulation, because the notice given was barely a week. In the end, the state agreed to compensate the extra costs. The fact that software had to be programmed to comply with the new tax regulation reiterates the disability of software to adapt itself to the changes in law. In other words, there is no interaction between two environments. The tax law regulates taxes in the living world and a software providing certain functionality replicates this environment. A change of law in the living world does not transform a functionality of software without human intervention. A sequence of activities performed by one or more people may be needed in order to achieve compatibility between two environments:

1. identifying regulatory changes which have to be transformed into software;
2. designating a person who is qualified to plan and implement changes in software so that the compatibility with new legislation is achieved;
3. evaluating the quality of changes;
4. checking for the compliance of changes to the legal system as a whole;
5. monitoring the operation of software as to its compliance to legal changes over time;
6. intercepting if there are inclinations from the result initially planned (beginning the process from the start).

The ability of software to adjust its functions can be described in the following scenario. Police officers have different access rights to KAIRI. The system is programmed to differentiate between police officers who have different responsibilities and tasks in the police work. The data in the system has been categorised into A, B, C, D categories, and the access to data has been determined by these categories. If a police officer who works in a street patrol is promoted to the position of the investigator, then he is given access rights to more categories of data. The person responsible for changing access rights has to adjust it to the new position. The person responsible for employment contracts and other related documents such as job descriptions has to prepare and introduce new documents to the police officer. The change of access rights could be automated by integrating the access functionality of different categories of data with the level of access defined in the employment contract. The job description can be linked to the employment contract as well. If necessary, the change in the employment contract automatically changes the access right and the job description. Furthermore, data protection law can be linked to the employment contract by the definition of the processor of personal data. The police officer being the processor of personal data is bound to several obligations prescribed in law. One could speculate whether a change of the obligation in law could be automatically implemented into the employment contract, the job description and thereby also affecting the access right. This is a theoretical example how software uses information received from legal sources so that it does not have to be reprogrammed, but the changes in law are automatically transformed into certain functionality.

The model of the legal software agent is shown in Fig. 1. The agent derives its legal knowledge from the legislation. The purpose of the agent is to provide the

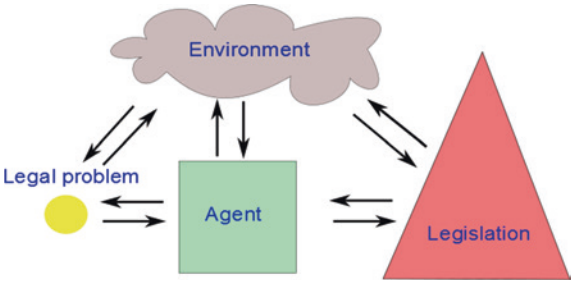


Fig. 1 Legal agent model

functionality for solving a legal issue, for instance, the limiting of access to personal information in order to achieve better privacy protection.

The environment in the model depicts the living world. Agents have growing capabilities to sense the world in a similar way as humans do. Different sensory technologies researched and developed today and in the future can establish and increase the sensory capabilities of an agent, for instance, technologies using smart dust or medical technologies which enable to detect health conditions including mental state of a human. A smart agent could be useful for multiple purposes in the police work. For example, if the agent could detect the over weight of heavy trucks, then the efficiency of the police work would grow considerably. Today the procedure of checking heavy trucks requires a crew of police officers together with the van full of equipment which has to be carried and mounted under the wheels of a vehicle each time the checking is performed. So can the problem of access to information be tackled by a software agent? The capability of the agent to evaluate the health of a police officer may include the ability to sense alcohol or drug consumption and the level of stress. For example, in case the agent diagnoses the depression of the police officer who is working in the traffic patrol and uses KAIRI for obtaining information about drivers stopped, then software agent could limit the access to information and direct the police officer for a medical check. The role of the agent in this scenario is to limit the risk of the making of biased decisions, or to prevent other misuses of personal information. If the agent has detected the alcohol consumption, then it could initiate the disciplinary procedure, which means that the police officer has to be removed from work for the duration of the procedure. The role of the agent is to block the access to databases and to inform the people responsible for the disciplinary procedure. Furthermore, a smart agent may have an ability to adjust its behaviour if the conditions of the disciplinary procedure prescribed in law change. It means that software has to be able to understand civil service law and to adjust its actions accordingly.

It is thrilling to speculate about conceptual ideas and benefits that software agents could deliver in the future. The purpose of these examples and illustrations is to explain possible practical solutions ahead of time. The discussion of the possibilities to establish the link between software agent and legislation is abandoned in the following chapters. The focus is on the basics of the information technologies that are forming the grounds on how the agent technology can regulate the

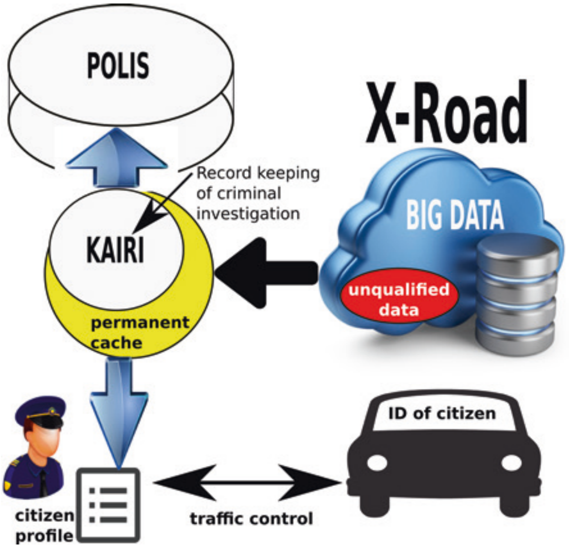


Fig. 2 System landscape supporting a police officer during a traffic control

flow of information so that it only gives the police officer the right type of information required for the performance of a certain task—the first task proposition formulated above.

4 Current Technology Surrounding KAIRI

The previous chapters sum up the legal scenario of privacy violations in Estonia that are specific of the degree of societal digitalization.[33] In Fig. 2, this larger digitalisation aspect also affects the way how traffic controls happen. Most importantly, the so-called X-Road system is the underlying system infrastructure on top of which also the police operate in Estonia. X-Road is a platform that enables secure Internet-based data exchange between the states information systems in a distributed peer-to-peer (P2P) and scalable way.[34] Thus, the X-Road allows organizations and people to securely exchange data from public and private domains.

Pertaining to the privacy-violation problem discussed in this chapter, the X-Road enables public enquiries during traffic controls of Estonian police officers. In order to use the services, the car driver must produce an identification to the police officer the latter uses for searching the personal record from state databases. The advantage of X-Road for officials work is the avoidance of the

[33] http://e-estonia.com. Accessed 27 Mar 2014.

[34] https://www.ria.ee/x-road. Accessed 27 Mar 2014.

labour-consuming processing of paper documents, large-scale data entry and data verification. Communication with other officials is also faster and more accurate, e.g., with other staff from the police force or the boarder patrol unit.

We conceptually depict the system landscape in Fig. 2. We assume the earlier mentioned case of a police officer performing a traffic control on a citizen. The latter produces her identification that comprises of the name together with the ID number. The police officer uses this data to search for citizen-related information on a mobile, wireless IT-device that connects to the database KAIRI. KAIRI is connected to a bigger database system that is part of the police's information system infrastructure called POLIS. The initial purpose of KAIRI is to manage data around criminal investigations and as Fig. 2 shows, it connects to a big data pool around the X-Road system that constitutes a federation of distributed databases.

As the police officer has to perform the citizen check quickly during the traffic job, it is not possible to perform elaborate searches in distributed database systems related to X-Road. The solution is to cache in KAIRI information from the big-data cloud to the right of Fig. 2. Depicted in a contained ellipse we show a subset of data termed *unqualified data*. The latter is a specific collection of rumour that could be important for resolving a criminal investigation at some point of time. However, as the name indicates, the quality of this data is not certified.

The KAIRI database in Fig. 2 comprises two tiers. The inner tier comprises the facts around a criminal investigation which adheres to the original purpose of the KAIRI database. The second tier serves as a cache of data taken from the big-data cloud, including unqualified data. A cache is an extra store of data so that future requests for that data are served faster. Without such a cache, the data must be first recomputed or fetched from the original storage location. Thus, the more requests are served from the cache, the faster the overall system performance while the police officer carries out the traffic control. Additionally, this architecture overcomes the bottleneck of limited IT-skills of the police officer as the KAIRI-system configuration automatically.

The problem of this pragmatic solution to use KAIRI as data cache has multiple disadvantages. First, as the caching procedure from the big-data cloud repeats periodically, the former consequently keeps growing in size. As the big data is from a heterogeneous distributed source, they lack structure and are of questionable quality. Additionally, being in the cache, keeping the data up to date is a challenge in correlation to the big-data source. Finally, a lot of the data may be very sensitive and in condensed combination could give insight about a citizen that a police officer must not be aware of.

More problems occur when the police officer looks at the automatically generated profile about a randomly stopped citizen in a traffic stop. As previous sections discuss, records about a citizen beyond a certain age must not be available to a police officer. However, the current architecture of the system depicted in Fig. 2 grant the police officer full access to all records beyond what he is permitted to see. The lack of provided privacy protection mechanisms during the automatic profile generation is problematic when the police officer must decide on possible punishment degrees during the traffic job. Earlier sections discuss this issue as the police officer decides more severe punishment in case prior violations of the law exist. Assuming profile data access that reaches beyond the intended level for this decision-making process, it is likely the police officer takes into account, for instance, the traffic violations that are older than one year.

5 Privacy-Ensuring Socio-technical Solution

The current state of KAIRI does not comprise adequate mechanisms to protect the privacy of individuals adequately and consequently, the system infrastructure requires a resolution. However, that solution must be of a socio-technical nature in that the system must adhere to a specific set of principles.

- *Responsible autonomy* addresses a shift towards teams or groups as the primary unit that also conforms to the traffic control case as policemen engage at least in pairs with a citizen and have potential reinforcement on standby. Thus, the privacy-assuring solution must pay particular attention to internal supervision and leadership at the level of the "group" and avoid rigid and inflexible silo thinking.
- *Adaptability* pertains to the way how an organisation responds to the external complexity by reducing the internal control and coordination needs by adopting the strategy of simple organisations and complex jobs. This strategy implies that groups must be relatively empowered to make their own decisions. Mapped to the traffic control case this means the police officer judges herself how the data delivered by KAIRI leads to a possible punishment degree of a law-violating citizen.
- *Whole tasks* that can be assigned on to a single, small, face-to-face group which experiences the entire cycle of operations within the compass of its capabilities and permissions. Thus, a police offer must complete an entire traffic control as a task but the sequence of activities involved changes in a flexible way on a case-by-case basis.
- *Meaningfulness of tasks* is the consequence of the earlier three principles. This task meaningfulness implies for each participant the task has total significance, dynamic closure and requires a set of skills to achieve the desired degree of autonomy. For the traffic-control scenario, it means KAIRI must provide the targeted means with the right level of utility to the police officer for completing a citizen inspection in one place at one time. In other words, in classic organisations the "wholeness" of a task is often diminished by multiple group integration and spatio-temporal disintegration. Thus, KAIRI must perform the right degree of information logistics so that privacy of the citizen remains assured.

The socio-technical solution for restoring privacy assurance rests on two factors. Firstly, the introduction of a detailed definition of role profiles for police officers that specify the permissions, competencies, access rights, and so on, in Sect. 5.1. Secondly, in Sect. 5.2 socio-technical artificial agents use these profiles for creating citizen profiles on the fly that protect a citizen's privacy.

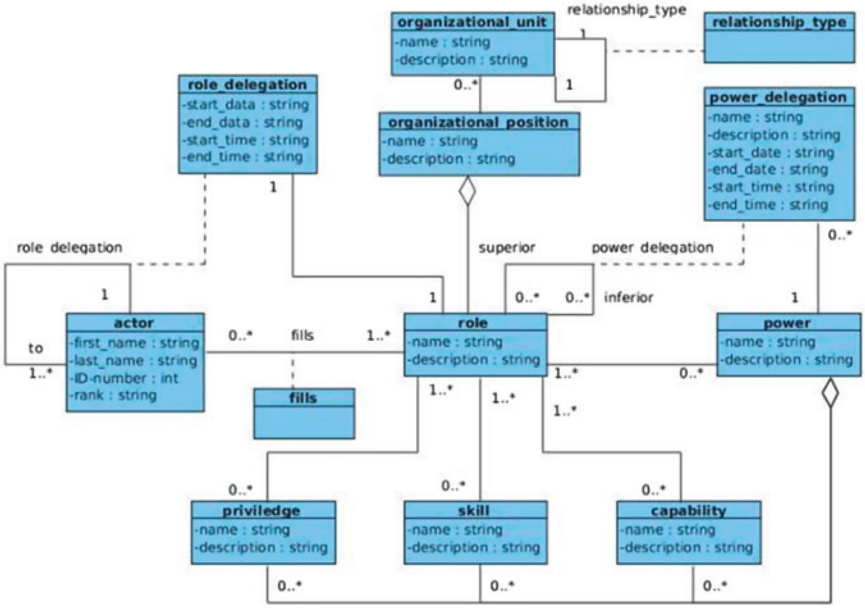


Fig. 3 Role model

5.1 Role Definition

An integral ingredient is the proper capture of details around a role affiliated with a police officer. The concept of a role in a system allows for enhanced flexibility compared to hard person assignment of properties. If the latter seizes to exist, it may cause problems of the functioning of the overall system. Instead functionality assignment to a role such as police officer into which an individual person slips, enhances the resilience of system use.

We use a class diagram for expressing the entities and relationships of the role model. A class diagram is a diagram for describing the structure of a system. The diagram in Fig. 3 shows classes as rectangular boxes and with contained attributes. The relationships between classes are logical connections and comprise several types. An association link is a straight line between classes, e.g., in Fig. 3 between classes actor and role. Numbers on both sides of the association link express the relationship cardinality, e.g., an actor can slip into one or many roles while a role either may remain unpopulated or populated by many actor instances. An association link may also link to one class only, e.g., an actor instance may delegate a role to another actor instance. If cardinalities on association links indicate a many-to-many relationship between classes, we assign a so-called association class to assure it is always clear which class instance relate to each other, e.g., role delegation. Finally, Fig. 3 also comprises a so-called aggregation association that represents a part-whole or part-of relationship. For example, an instance of power is composed of optional parts of privilege-, skill- and capability instances.

The role model depicted in Fig. 3 is a small part of a larger version that we omit due to space limitation. An individual resource has an actor as a subclass who is a concrete person. Such an actor may be directly assigned to a task such as performing a traffic control. An actor references one or many roles that can be delegated to other actors, e.g., another police offer. Furthermore, an actor has also one or many organizational positions that are related to organizational units that reflect the rank of a police officer. Such an organizational position may mean several privileges are attached that are also related to roles. For example, a police officer is privileged to directly collect money for a fine. A role is a subclass of a resource type and may be filled by several actors.[35] Besides already mentioned, several capabilities may be required to fill a role such as checking unqualified data. Furthermore, a role can give certain power that can also be delegated to other roles for a limited time, e.g., for punishing a citizen who commits a traffic violation. Power that is attached to a role is also related to capabilities and privileges.

5.2 *Socio-technical Agents*

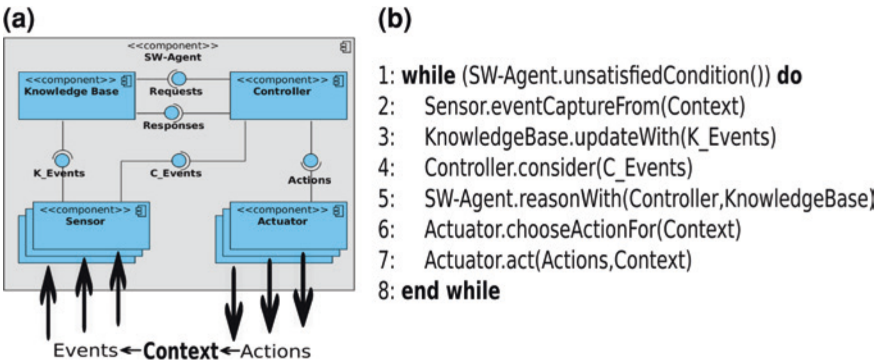
For privacy assurance during a traffic stop, socio-technical artificial agents are instrumental. In this context, an agent is a software application that supports social behaviour of a computational system. An agent is an intelligent system that perceives its environment and takes actions that maximize its chances of success. Additionally, an agent is an active entity that reasons on behalf of a police officer.

The depiction in Fig. 4a shows the reference architecture of a sociotechnical agent.[36] It comprises four components with different functions. The bottom left component labelled sensor gathers events as input that occur in the context of an agent. Those events are split inside the agent and partially the knowledge base and the controller receive. The knowledge base comprises entities and facts of the agent's context together with ontological repositories for allowing a correct interpretation of the stored data. The second recipient of sensor-processed events the controller receives. The latter component uses in addition the knowledge base for algorithmic processing to perform pseudo anthropomorphic reasoning that copy humans in a machine-learning way. The latter is a branch of artificial intelligence and focuses on the construction and study of systems that learn from data.

- *Belief* in a human-agent sense is a state of mind in which an individual holds an unproven proposition or assumption of something to be true.
- *Responsibility* in a legal sense is the mental capacity to decide if a person can be held accountable for a crime.

[35] See e.g., Norta 2007.

[36] Sterling and Taveter 2009.



```
while (SW-Agent.unsatisfiedCondition()) do
    Sensor.eventCaptureFrom(Context)
    KnowledgeBase.updateWith(K_Events)
    Controller.consider(C_Events)
    SW-Agent.reasonWith(Controller,KnowledgeBase]
    Actuator.chooseActionFor(Context)
    Actuator.act(Actions,Context)
end while
```

Fig. 4 Conceptual agent architecture in (**a**) and a pseudo-code algorithm for agents in (**b**)

- *Expectation* is a belief centered on the future, with a particular probability to be realistic.
- *Capability* is the ability to perform or achieve certain actions or outcomes through a set of controllable and measurable faculties, features, functions, processes, or services.
- *Goal* is a desired result a person or a system envisions, plans and commits to achieve an individual or socially desired end-point in accordance with a plan and within a deadline.
- *Desire* is a sense of longing for an agent, or object, or hope for an outcome.
- *Intention* is an agent’s specific purpose in performing an action, series of actions, or targeted goal.

The pseudo-code algorithm in Fig. 4b shows the abstract structure of this machine-learning algorithm in the controller component of a socio-technical agent. Accordingly, the main encompassing control-flow element is a while-loop that performs as long as the agent is unfulfilled. Inside the while-loop, the agent senses events from the environment and uses that input for updating the knowledge base if needed. These events also serve for the reasoning in the controller in a way that the agent’s machine learning algorithm displays the pseudo anthropomorphic properties in an artificial-intelligence sense as discussed above. Consequently, the socio-technical agent projects events through the actuator component onto its contextual environment. The latter reacts to that projection and the loops starts again from the beginning unless a satisfaction occurs of the condition-statement in the while-loop.

6 Resolution Suggestion for Privacy Protection

The proposed approach from the previous section we map on a TO-BE architecture that Fig. 5 depicts. The changes in comparison to the AS-IS architecture in Fig. 2 are as follows. The most important change the introduction of

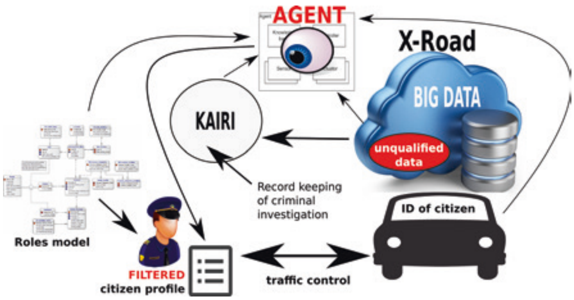


Fig. 5 Privacy-protection resolution proposal using socio-technical agents

a socio-technical agent that a traffic-control case instantiates i.e. each control instance has a dedicated agent instance too which terminates its lifecycle once the control-case ends. The police officer has a proper affiliation with a data-model instance the agent uses on inception. The KAIRI database shrinks back to its original purpose of criminal-investigation record keeping. The problematic cache is now not necessary any longer as the agent creates tailor-made citizen profiles based on the data-model instance assigned to a police officer.

With the cache falling off from KAIRI, the data-exchange protocol shrinks between KAIRI and the big-data cloud affiliated with the Estonian X-Road system. Instead, the agent now has dedicated access to the distributed databases for accumulating profile data. The agent assures not only that data access and processing happens at feasible speed but the resulting citizen profile comprises only the data a respective police officer has permission to see. Note that KAIRI is still part of the larger POLIS system, however, in Fig. 5 we omit a repeated depiction. Finally, the agent also assures a fast relaying and committing of newly created data the police officer creates during the lifecycle of a traffic control, e.g., the citizen drives under the influence of alcohol.

7 Conclusion

With the intense degree of digitalisation of Estonia by using the unique X-Road system for administrational purposes in public but also private domains, novel problems occur with respect to privacy of personal data. We show in this chapter that the pace of technological innovation makes it a challenge for the legal situation to keep up. As the listed cases in the introduction show, several cases of privacy violations have gone to the Estonian Supreme Court. Particularly prevalent among the privacy violations are situations related to police activities such as during regular traffic controls.

While there are general privacy protection laws in place, the high degree of digitalization makes it very challenging for police officers during a traffic control to respect the privacy laws. The situation is simply too complex for the following reasons. In order to adhere to the complex privacy protection regulations, the police officer must make a continuation expert assessment during a traffic control situation as if s/he would be trained in a comparable way as a professional lawyer. Secondly, the required data for performing a traffic control with a citizen is distributed in many different databases. Thus, the police officer must also act in a comparable way to an ICT-expert and know how to quickly access all data in the distributed technological infrastructure of X-Road. These two factors show the complexity of the situation that results in data privacy violations even without mal intent.

Not only is the police officer overwhelmed with the imperative of respecting citizen data privacy, also the system designers of X-Road play a role in causing this problem to occur. While the initial designing and implementation of the X-Road is initially driven by purely technically considerations such as security and dependability of interoperability, privacy assurance has a socio-technical dimension that requires taking into account the nature of human action during system design. The currently existing system architecture forces the maintenance of a local data cache around a database called KAIRI that merely has the original purpose of maintaining criminal-investigation records. However, as it is the objective to quickly process personal data during a traffic control, it is necessary to copy from the highly distributed X-Road databases permanently into the KAIRI cache to allow for the quick permanent access. The problem is that complete datasets become available that a police officer should not be able to see. The cached data is impossible to keep in an updated state compared to the source in the distributed X-Road databases. Additionally, the cache itself continuously keeps growing in size, making it ever more complex to maintain dataset consistency.

As a remedy to this novel level of complexity that public workers experience in highly digitalized Estonia, we propose two specific solutions that not only restore data privacy on the fly but also correct the original architecture flaw in KAIRI system design. Firstly, we give a role-focused model to capture additional facts that describe the profile of police officers. The model captures facts about the police organization, recording privileges, skills, capabilities. In a flexible way persons may assume specific roles which grant them certain powers such as during a traffic control situation. The delegation of specific powers and roles are possible, making the model flexible to contextual changes.

The second remedy is the use of software agents that take into account the facts recorded in the role model to create on the fly tailor made citizen profiles with exactly the facts an official is permitted to see based on the assumed role without any violation of data privacy. The software agent comes into existence when a traffic control commences and terminates when this control comes to an end. The software agent has privileged access to the distributed X-Road databases so that the creation of personal data happens in real time so that a citizen can be controlled quickly. The side effect of introducing such a software agent is that the need disappears for a cache attached to the KAIRI system.

For future work two directions exist. First, the introduction of software agents only partially resolves all problems related to a traffic inspection. The complexities involved around a traffic control can be reduced even more when the administrational process is automated too. The police officer can then instantiate an administrational-process template and follow pre-defined steps for traffic controls that software agents support by delivering on the fly targeted data flow from the distributed X-Road database systems. Adopting administrational processes also allows for a design that enforces adherence to public law and regulations.

Another angle of future work focuses on the safeguarded introduction of intelligent software agents. The artificial-intelligence community recognizes the dangers of allowing unchecked introductions of AI systems into society.[37] Specifically in a highly digitalized society such as Estonia, the introduction of badly designed intelligent software agents carries the potential for considerable destruction in the X-Road system and the affiliated distributed databases.

References

Gowda, R. S. (2008). Role of software agents in e-commerce. *International Journal of Computational Engineering, 3*(3), 246–251.

Hewitt, C. et al. (1973). *A universal modular ACTOR formalism for artificial intelligence*. In Proceedings of the 3rd International Joint Conference on Artificial Intelligence (pp. 234–245). San Francisco: Morgan Kaufmann Publishers Inc.

HOMER Report (2013). *Socio-economic impact study*, March 2013, http://homerproject.eu/publications-documents. Accessed 1 April 2014.

Janssen, K., & Dumortier, J. (2003). Towards a European framework for the re-use of public sector information: A long and winding road. *International Journal of Law and Information Technology, Oxford University Press, 11*(2), 184–201.

Männiko, M. (2001). Õigus Privaatusele ja Andmekaitsele. Juura.

McCarty, L. T. (1977). Reflections on TAXMAN: An experiment in artificial intelligence and legal reasoning. *Harvard Law Review, 90*(5), 837–893.

Norta, A. (2007). *Exploring dynamic inter-organizational business process collaboration*. PhD thesis, Technology University Eindhoven, Department of Information Systems.

Nwana, H. S. (1996). Software agents: An overview. *Knowledge Engineering Review, 11*(3), 1–40.

Nyman-Metcalf, K., Täks, E. (2013). Simplifying the law—can ICT help us? *International Journal of Law and Information Technology*, 1–30.

Solove, D. J. (2001). Privacy and power: Computer databases and metaphors for information privacy. *Stanford Law Review, 53*, 1413–1434.

Sterling, L., Taveter, K. (2009). *The art of agent-oriented modeling*. Cambridge: MIT Press.

Westin, A. F. (1970) *Privacy and freedom*. The Boadley Head, 7.

Yampolskiy, R. V. (2012) AI-complete CAPTCHAs as zero knowledge proofs of access to an artificial intelligent system. *ISRN Artificial Intelligence*.

[37] Yampolskiy 2012.