THE STATUS OF PRIVACY ENHANCING TECHNOLOGIES

(PET) online and offline

John Borking

Associate Commissioner of the Dutch Data Protection Authority

P.O. box 93374, 2509 AJ The Hague,

The Netherlands

jbo@cbpweb.nl

Abstract: This paper explains Privacy-Enhancing Technologies (PETs), their anchoring in the EU Data Protection Directives and other data-protection systems, and how they might contribute to the lawful processing of personal data. We look back since the Dutch and Ontario Privacy Commissioners defined PET in 1995 to see what has changed and what has been achieved. The EU project PISA (privacy incorporated software agent) will be used as an example of future PET.

Keywords: Privacy; EU Data Protection Directives, Privacy-enhancing Technologies, (PETs), Privacy-Incorporated Software Agent (PISA), Identity Protector.

1. INTRODUCTION

The introduction of legislation for the protection of personal data, in accordance with EC Directive 95/46, has implications for the way information systems are set up in organizations of all kinds. The data controller – the party that determines the object and means of data processing – is responsible for making sure that the legislation is complied with. The controller is potentially liable for any unlawful processing of data. The processor – the party that processes data on the controller's behalf – is separately liable for shortcomings in the way the process is organized within his organization.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35696-9_19](https://doi.org/10.1007/978-0-387-35696-9_19)

E. Nardelli et al. (eds.), *Certification and Security in E-Services*

© IFIP International Federation for Information Processing 2003

To ensure that personal data is processed appropriately, it is necessary to give processing policy an important position within the management cycle. This will also have the effect of providing adequate support for the privacy rights of the private individual or consumer. Privacy protection will normally take the form of an additional system of measures and procedures. Existing measures and procedures for the control, protection and processing of personal data may need to be reviewed and tested against the objectives of the national legislation.

When the Dutch Data Protection Act (WBP) was passing through the Senate of the Dutch Parliament, the Minister of Justice specified that technical means should be used for the protection of privacy (WBP, section 13). Privacy-enhancing technology (PET), as it is known, can be very useful for ensuring that personal data is handled correctly and with due consideration, and that privacy principles are respected. Motion 31, which was put forward by an MP called Nicolai and accepted by the Lower House, called on the government to make use of PET in its own information systems. Budgetary provision was also made for the government to act as a PET launching customer.

PET has since become an important element of the practical and theoretical privacy protection repertoire. This contribution examines the role that PET can play in the protection of privacy.

2. INFORMATIONAL PRIVACY, OBJECTIVES OF THE LAW

Privacy⁹ defined by Westin as: "the claim of individuals...to determine for themselves when, how and to what extent information about them is communicated to others", is a fundamental human right, defined in Article 8 of the 1950 European Convention of Human rights. It is one of the most important human right issues of our evolving information age¹⁰. Informational privacy has two distinct characteristics:

1. The right to be left alone;
2. The right to decide oneself what to reveal about oneself.

So, although it is a situation that is wanted by an individual, it primarily comprises a set of rules of conduct between the individual person and the person's environment with respect to the processing of personal data like PII in the Internet environment. A major step towards privacy protection in Europe was the adoption of Convention 108 of the Council of Europe.

⁹ A. Westin, *Privacy and Freedom*, New York 1967 p.7

¹⁰ D. Banisar, *Privacy & Human Rights*, Washington/London 2000

Today, informational privacy protection for individuals is articulated through different European Union Directives like 95/46/EC (Data Protection Directive hereafter: DPD), Directive on Telecommunications 97/66/EC¹¹, 1999 Digital Signature Directive 99/93/EC and non-EU legislation¹². This kind of legislation defines a set of rights concerning personal data accruing to individuals irrespective of sector of application, and creates obligations concerning the processing of data by third parties. The DPD has two objectives:

1. Creating a high level of protection of personal data, and
2. Enabling the free movement of data within the EU.

Furthermore the DPD has two major functions:

1. Empowering the individual to manage his or her own personal data and protecting these data within the limits as defined in the Directive;
2. Creating a protective legal regime when personal data are processed¹³.

3. BASIC PRIVACY PROTECTION REQUIREMENTS

The DPD sets out a number of basic requirements for the lawful processing and acceptable use of personal data. There are nine privacy principles incorporated into the DPD¹⁴. The principles are collations of articles from the DPD that are frequently applied together. These may be summarized as follows:

1. *Intention and Notification* - The processing of personal data must be reported in advance to the Data Protection Authority or a privacy officer (where applicable), unless the processing system in question has been exempted from notification.
2. *Transparency* - The person involved must aware of who is processing his personal data and for what purpose.

¹¹ Recently the 2002 privacy telecom directive has been adopted: 2002/58/EC

¹² See the 2000 Canadian Personal Information Protection and Electronic Documents Act

¹³ Summary, Conclusions and recommendations of the TNO report STB-02-13a Privacy Enhancing technologies and Information Systems of the Central Administration The Hague 28 February 2002 commissioned by the Ministry of Interior (BZK), handed out 23rd of May 2002 in The Hague, after the symposium Privacy by Design organised by the Dutch Data Protection Authority (CBP)

¹⁴ J.P. Leerentveld RA RE, G. W. van Blarckom RE, WBP Raamwerk Privacy Audit. Samenwerkingsverband Audit Aanpak, The Hague 2000. The articles of the Directive 95/46/EC have been 'translated' into nine principles. The Directive itself is too abstract and too complex to directly generate technical specifications. Furthermore the abstract articles need interpretation and interpretation may lead to design risks.

3. *Finality principle* - Personal data may only be collected for specific, explicit and legitimate purposes and not further processed in a way incompatible with those purposes.
4. *Legitimate Ground for Processing* - The processing of personal data must be based on a foundation referred to in national legislation, such as consent, contract, legal obligation, justified interest and such like. For special data, such as health, stricter limits prevail.
5. *Quality* - The personal data must be correct, accurate, sufficient, to the point and not excessive in relation to the purpose in question.
6. *Data subject's rights* - The data subjects involved have the right to have access and to correct their data as well as the right to object.
7. *Security* - Providing appropriate security for personal data held within IT-systems is one of the cornerstones of the DPD. Measures of technical and organizational nature suitable and proportional to the sensitivity of the personal data and the nature of possible risks have to be taken to avoid potential harm should the PII be misused or disclosed in an unauthorized matter.
8. *Processing by a processor* - If processing is outsourced to a processor, it must be ensured that he will observe the instructions of the controller.
9. *Transfer of personal data outside the EU* - In principle, the transfer of personal data to a country outside the EU is permitted only if that country offers adequate protection.

The first seven privacy principles are always applicable while processing personal data. The principles 8 and 9 are dependent upon the existence of a processor and whether any processing occurs outside the EU.

4. PERSONAL DATA

In legal terms, personal data means any piece of information regarding an identified or identifiable natural person¹⁵. Whether we can talk of 'personal data' depends on a number of elements of which, within the scope of this document, 'identification' is the only significant element. According to Article 2 of the DPD, a natural person can be identified 'directly or indirectly'. Direct identification requires basic details collected in PII. PII is name, address, a personal number, a widely known pseudo-identity, a biometric characteristic such as a fingerprint, etc. Indirect identification

¹⁵ See Article 2 of EU Directive 95/46/EC; 'personal data' are defined in for instance Article 1 of the UK Data Protection Act 1998 and there is a lot of confusion about identification and verification.

requires other unique characteristics or attributes or a combination of both, to provide for sufficiently identifying information¹⁶.

Non-identification is assumed if the amount and the nature of the indirectly identifying data are such that identification of the individual is only possible with the application of disproportionate effort¹⁷. Whether we can talk of disproportionate effort depends, on the one hand, on the nature of the data and the size of the population; and on the other hand, the resources of time and money one is willing to spend in order to be able to identify the person¹⁸. Internet identifiers such as an IP address, browsing activities of a user, session login data and the listing of websites visited by an Internet user are classified as personal data.

5. THE PRIVACY RISK ANALYSIS

The general approach for risk analysis and subsequent requirements determination etc. is derived from a comparable domain: the risk assessment for information security in British Standards 7799, the Code of Practice for the Risk Analysis and Management Method, Information Security Handbook of the Central Computers and Telecommunications Agency (CCTA).

The presented approach highlights the way privacy consolidation and subsequently privacy protection can be handled, but by the same token, after due adaptation, the same line of thinking can be applied to other topics where vital elements can be violated, like (physical) safety, security, etc.

The privacy consolidation approach starts from the privacy regulations, which mandates the Privacy Consolidation Ordination, or: the obligation to implement such measures in the system that the privacy regulations are met to the satisfaction of the regulators.

At the first layer of analysis, there are two branches. The “assets” represent the objects or subjects that bear personal data and hence are vulnerable to privacy threats. The identification of such assets is necessary, because their valuation is a factor in risk assessment. In parallel to this the threats are identified. These threats can be linked to an asset but may also be a direct consequence of the regulations. Later, they can be linked to specific assets.

¹⁶ See J.J. Borking, C.D. Raab, *Laws, Pets and Other Technologies For Privacy Protection - Journal of Information, Law and Technology (JILT)* January 2001

¹⁷ See Recital 26 of the EU Directive 95/46/EC.

¹⁸ We can only note in passing here that the concepts of ‘identification’ and ‘identity’ are essentially contested in theory and often arbitrary and ambiguous in practice. See C.D. Raab, ‘Identity Checks - and Balances’, in E. Bort and R. Keat (eds.), *The Boundaries of Understanding: Essays in Honour of Malcolm Anderson*, Edinburgh: International Social Sciences Institute, 1999, pp. 87-95.

The second layer of results establishes the two factors relevant for the risk assessment.

Figure 1. Model Risk Assessment

The first factor has two inputs. The first input is the valuation of the assets: how important is it that the assets are protected possibly expressed in an ordinal list to produce a sequence, in some monetary unit to make it financially tangible or in some other unit; the purpose is to rank the assets in value. The second input is the result of the assessment of the severity of the threats: what is the gravity of the consequences of a privacy violation. The

second factor represents the likelihood of occurrence, expressed as the likelihood that it happens in a certain time frame. This is determined by expert judgment or, preferably, by statistical analysis.

The third layer is the result of the risk assessment. The approximate formula is that a risk is the product of the valuation of the consequences of the threat times the likelihood of occurrence of such threat. This should at least give some insight in the priority of countermeasures (or control risk), where and how to invest effort and money.

The fourth layer is the definition of requirements that are deemed adequate to meet the safety regulations. However, because the countermeasures may, out of themselves, also create secondary threats, eventually the whole privacy protection approach has to be repeated, until no further threats are expected.

The last layer is the implementation of the requirements for the countermeasures, which is the practical result of the activities. In actual practice, it can be that privacy loopholes are still present, in which case a new privacy protection ordination arises, but this is outside the scope of the current considerations.

In privacy threat analysis the focus will be on the right hand side, representing threat identification and assessment of severity of consequences of such threats.

6. FIVE-PRONGED APPROACH TO THE PRIVACY THREAT ANALYSIS

A methodology is based on a '*way of thinking*' about a certain area of investigation. Then the '*way of thinking*' is strengthened by the '*way of working*' that is expressed in the methodology itself. The final part is the '*way of writing*' that describes in any desired level of formality how the results have to be put on paper¹⁹.

For the analysis of threats to privacy a five-pronged approach is chosen (of which the privacy regulations, as adopted in a certain state (i.e.: EU or Canada) is one of them) which can be depicted in the following figure.

That five different perspectives on the issue of personal data are applied, finds its rationale in the observation that one single line of thought may fail to illuminate all the aspects of privacy that are relevant for the ultimate implementation. Five different perspectives, and hence five different lines of thinking, is more likely to reveal everything that is to know about the

¹⁹ Examples of this are abundant in the arena of software system development. A case in point is object-oriented development. There is object oriented thinking object oriented design methodologies and description languages like UML

privacy threats. That this may result in the same threat identified more than once isn't a problem.

The five perspectives chosen are:

- Privacy regulations, as defined in a certain country or country union: these regulations inherently list a number of privacy threats, if these regulations are not adhered to;
- Purpose of the system, which creates its own threats: because the user (private person) wants to achieve something, that person creates privacy threats;
- Solution adopted, which may or may not create threats of its own;
- Technology used: because of the way a certain system is implemented, certain threats may emanate which are not necessarily consequences of the intended purpose. Meanwhile, the technology will harbour some of the privacy enhancement measures;
- Situation in which the ultimate system will be used: which, although not necessarily creating threats of its own, may or may not aggravate (or alleviate) previously identified threats and hence may incur more demanding technological measures. This part is especially needed when a commercial off the shelf (COTS) product is going to be used in an unforeseen situation; the previous four types can be followed whether or not the system is a COTS or dedicated to a certain problem and environment.

Figure 2. Five-pronged approach to Privacy Threat Analysis

6.1 Risk analysis Branch 1: threat identification and appraisal of likelihood of occurrence

Following the five-pronged approach, the sequence of steps in the Threat Identification branch, the right-hand ramification of the risk analysis diagram, can be elaborated in the following diagram

Figure 3. Risk Analysis Branch 1

As can be derived from the figure, the subsequent steps result in separate threat survey documents. The sequence of steps also shows the feature of following the system design lifecycle fairly closely, which has the obvious advantage of easy incorporation.

6.2 Risk analysis Branch 2: asset identification and appraisal of consequences of threats

The investigation of the left hand branch of the risk analysis diagram can be depicted in the following figure.

Figure 4. Risk Analysis Branch 2

6.3 Risk analysis Branch 3: risk identification, prioritization, privacy requirement determination, and implementation

In the previous figure also the list of risks is indicated. This list can be prioritized, because the higher the risk, the more stringent the privacy enhancing measures to counter the threat.

This results in requirements for the intended system and this, in turn, will be implemented in the ultimate system.

Thus the objective of such risk analysis is to provide insight in what is the best way to improve the privacy protection within the information system, which privacy protection measures to apply and with what rigor.

6.4 Basic Threats

Derived from the privacy principles and DPD the following threats can be discerned:

- Secret possession of (control over) personal data files: the data subject and the authorities are unaware of the existence of the personal data and the control the controller of these data has;
- Secret processing of personal data: processing out of sight or knowledge of the data subject;
- Out of bounds processing by controller: processing of personal data that is not within the bounds stipulated in the personal data constraints or can be expected to be outside the scope and intention of the collection;
- Out of law processing: processing of personal data that is illegal, forbidden by national law (or is not explicitly allowed if it can be expected to be of dubious nature);
- Personal data deterioration: the personal data is in contradiction with the current situation, either caused by external changes or by incorrect or incomplete insertion, collection or insertion;
- Irresponsiveness to discontent: the controller does not respond, or incorrectly, incompletely or unduly late, to requests for correction or other implications to the personal data or the personal data constraints of a data subject; the controller thwarts communication; also: there is no authority with reprehension, correction, sanction or other influence on the controller to sustain the data subject's legal rights;
- Out of bounds processing by processor: the processor does not follow the personal data constraints as provided by the controller or violates the rules;
- Out of jurisdiction processing: the personal data are transferred to a controller which has no legal obligation to obey the personal data

constraints or where legal obligations about privacy are less stringent than in the data subject's privacy regime;

- Personal data and personal data constraints violation: the controller and processor disobey the obligation to follow the personal data constraints concerning disclosure, retention, termination and safeguarding of correctness, including the obligation to take precautions against loss or mutilation of the personal data or the personal data constraints.

7. TRADITIONAL SECURITY MEASURES NOT SUFFICIENT

The requirements referred to in the DPD must be implemented efficiently in the organization in order to give proper support to the citizen's right to privacy with respect to personal data. It is therefore important to devise a proper system of general processing measures and procedures that should be present in order to protect company processes and in connection with specific protective measures for the processing of personal data. The restrictions that the organization of data systems can impose on the possibility that their users can comply with privacy legislation are evident. One simple example is where a system contains an inescapable 'date of birth' field, while analysis of the company's processes shows that recording the birth date of all persons included in the system is excessive. System design can just as easily ensure that users correctly observe the law. As a rule, privacy protection will constitute a supplementary system of measures and procedures in addition to the usual processing and security measures, but it should be assigned a significant place in management processes in order to implement and maintain a balanced processing policy for personal data.

When an organization is asked what it has done to protect privacy, it is apt to emphasize the personal data security measures it has in place. Although the use of safeguards to prevent unauthorized access to personal data is an important aspect of privacy protection, it is not sufficient in its own right. This is because such safeguards rarely involve the encryption of stored data; consequently, effective protection depends entirely on the security measures being correctly implemented and functioning properly.

It is therefore preferable to take technical measures that protect the individual's privacy at the point of data collection. Such measures may do away with the need to generate or record any personal data at all. Alternatively, they may minimize or even obviate the need to use or store identification data.

Given the basic legal requirements for privacy protection, it will be apparent that, if technical provisions are to be deemed adequate, they must go beyond the implementation of traditional security measures.

8. THE PET REPORT

ICT offers solutions in the shape of privacy protection for users, consumers and citizens. The application of ICT to protect privacy has become widely known under the name Privacy-Enhancing Technologies (PET or PETs)²⁰. PETs have been defined as a coherent system of ICT measures that protects privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data; all without losing the functionality of the data system²¹.

In keeping with data protection principles and their practical application to particular personal data systems, research and development towards the establishment of PETs as a leading set of instruments for privacy protection led to the publication, in August 1995, of the report *Privacy Enhancing Technologies - the Path to Anonymity*, written in association with TNO/FEL (the Dutch national research center) and the Information and Privacy Commission of Ontario, Canada²². There has been six years of experience in the Registratiekamer²³ (The Dutch Data Protection Authority) with this, and research continues towards its realization in practice. The report's researchers posed two central questions:

- What conditions must be kept in mind when engineering an information system in order to guarantee that the system be used effectively and efficiently without revealing the user's identity?
- What types of information and communication technology can contribute towards achieving this goal?²⁴

The main issue that was also posed was whether identity is necessary for all processing steps within a data system²⁵. The report shows that in many

²⁰ See revised edition By R. Hes and J. Borking, *Privacy-Enhancing Technologies: The Path to Anonymity*, The Hague: Registratiekamer 1998

²¹ See J. Borking, 'Der Identity Protector', *Datenschutz und Datensicherheit*, 11, 1996, pp. 654-658.

²² See the revised edition by R. Hes and J. Borking, *Privacy Enhancing Technologies: The Path to Anonymity*, The Hague: Registratiekamer, 1998.

²³ See J. Borking, 'Der Identity Protector', op. cit.; J. Borking, 'Einsatz datenschutzfreundlicher Technologien in der Praxis', *Datenschutz und Datensicherheit*, 11, 1998, pp. 636-640. See also the articles in the Schwerpunktheft 'Digitales geld', *Datenschutz und Datensicherheit*, 7, 1997.

²⁴ R. Hes and J. Borking, op.cit., p. 13.

²⁵ Ibid., pp 19-22.

cases the identity of the user, consumer or citizen is irrelevant. In some cases, however, identity *is* relevant for legal reasons, for instance to pay for certain services, or to open a bank account.

To implement matters technically, a system element called the 'identity protector'²⁶ is used within the data system to convert the identity of the person involved (the person whose data are being processed - the 'data subject') into one or more pseudo-identities²⁷. The placement of the identity protector provides for at least two different domains within the data system; one domain where the identity of the person involved is known or accessible (the identity domain) and at least one domain where this is not the case (the pseudo-identity domain). The aim of the pseudo-identity domain is to make sure the person involved cannot be traced on the basis of previously obtained personal data, and vice-versa, to make sure the personal data cannot be found on the basis of the obtained identity²⁸.

One technical mechanism for protecting the identity of data subjects is the use of an identity protector element within an information system (see figure 5). The identity protector converts the identity of the subject into one or more pseudo identities.

Figure 5. The IDENTITY Protector

²⁶ Ibid. p. 23.

²⁷ Pseudonymity as well as anonymity are key concepts in the design of information systems and privacy-protecting tools, and have received considerable attention in recent years. See R. Clarke, 'Identification, Anonymity and Pseudonymity in Consumer Transactions: A Vital Systems Design and Public Policy Issue', in *Conference on 'Smart Cards: The Issues'*, Sydney, 18 October 1996, available at <<http://www.anu.edu.au/people/Roger.Clarke/DV/AnonPsPol.html>>.

²⁸ E. France, op. cit., p. 215: 'The notion with privacy enhancing technologies is that ... the individual's true identity can be protected when it is not necessary to use it.'

Installation of an identity protector (see figure 5) as system element of an I.S. entails dividing the information system into at least two domains: one or more in which the subject's identity is known or accessible (the identity domain(s)) and one or more in which the subject's identity is not known or accessible (the pseudo-identity domain(s)). The identity protector can be incorporated at any point in the information system (see figure 6).

Figure 6. The Working of the IDENTITY Protector

The pseudo-identity domain serves partly to ensure that the data subject cannot be traced from previously obtained personal data, and partly to ensure that personal data cannot be consulted on the basis of a known identity.

The identity protector in a data system can take several forms, for example:

- A separate function implemented in the data system;
- A separate data system supervised by the individual (for instance, the smart card for biometrics identification);
- A data system supervised by a party entrusted by a service provider and consumer ('Trusted Third Party' (TTP))²⁹.

The use of an identity protector thus makes it possible to intervene preventively within the data system to hide the identity of the person

²⁹ TTP services are for: a) authenticity and integrity and b) for confidentiality. See J. Versmissen, *Keys of Trust: TTP Services and Privacy – The Judicial Preconditions Explored*, A&V Study No. 22, The Hague, 2001.

involved and to control subject identifiability within an information system with a view to preventing privacy problems.

One technology that is almost indispensable for ensuring reliability in an open electronic environment is cryptography. An increasingly popular option in this field is public-key cryptography. This technique can be used in two ways. One can have a public encryption key, which anyone can use to write an encrypted message that can be read only by the person with the corresponding private decryption key. Or one can have a public decryption key, so that users can authenticate the source of an encrypted message, which can only have come from the owner of the corresponding private encryption key. The second arrangement is known as a digital signature system.

Public-key cryptography is only effective if one has a reliable method of linking the key to the identity or another attribute of the owner. To make this possible, a public-key infrastructure (PKI) is required. Within the PKI, a trusted third party (TTP) provides the necessary link, making use of its own electronic signature. A digital certificate is a digitally signed electronic document issued by a TTP, linking the public key to attributes of the certificate holder.

The use of MIX nodes in telecom and other networks is a promising way of protecting data in transit between sender and recipient. By installing a series of nodes and using certain encryption/decryption techniques, it is possible to modify and regroup data in transit so as to make it almost impossible for an unauthorized party to tell whether a message is 'coming' or 'going', or therefore to analyze the data.

As the range of PET technologies available for network use increases, there is more and more scope for protecting the identity of the user and the provider, and for preventing the identification of a network, server, query or the like.

Thus an Identity Protector can be used in conjunction with various identification technologies, such as digital signatures, digital certification and MIX nodes³⁰.

³⁰ A MIX node is a processor that takes as input a certain number of messages that it modifies and outputs in a random order. The messages are modified and reordered in such a way that it is nearly impossible to correlate a message that comes in with a message that goes out. The MIX nodes can be used to prevent traffic analysis. See O. Berthold, A. Pfitzmann and R. Standtke, 'The Disadvantages of Free MIX Routes and How to Overcome Them', in *Proceedings of the Workshop on Design Issues in Anonymity and Unobservability*, Berkeley (CA), 2000, pp. 27-42. See also P. Syverson, G. Tsudik *et al.*, 'Towards an Analysis of Onion Routing Security', in *ibid.*, pp.83-100; O. Berthold, H. Federrath and S. Kopsell, 'WebMIXes: A System for Anonymous and Unobservable Internet Access', in *ibid.*, pp.101-115.

9. THE LEGAL CONTEXT FOR PET DEVELOPMENT

The EU Directive 95/46 provided the sharpest stimulus for the revision of data protection legislation in the Member States from the late 1990s on. The incorporation of PETs into strategies for privacy receives some encouragement from Article 17 of the Directive, which requires data controllers to implement ‘appropriate technical and organizational measures’ to protect personal data, especially in network transmissions. Recital 46, which augments the meaning of Article 17, highlights the requirement that these measures should be taken ‘both at the time of the design of the processing system and at the time of the processing itself’, thus indicating that security cannot simply be bolted onto data systems, but must be built into them. This provision mainly concerns data security, but it is generally intended as a safeguard against other forms of unlawful processing. This has been transposed into Dutch law as Article 13 of the WBP:

‘The person responsible shall ensure suitable technical and organizational measures to protect personal data against loss or any form of unlawful processing. Taking into account the technical status and enforcement expenses, these measures shall guarantee suitable protection given the risks that accompany the processing and the nature of the data to be protected. The measures should also avoid unnecessary collection and the further processing of personal data’.

During considerations in the Lower House, MPs Scheltema-de Nie and Wagenaar added the last sentence of Article 13 by means of amendment 22³¹. Discussing the Dutch Privacy Protection Act in the Senate, the Minister of Justice said that:

‘Current IT capabilities to abuse personal data necessitate a search for supplementary possibilities to make sure personal data are treated properly and accurately. Consider partial or complete ‘anonymizing’, for instance, by eliminating from personal data their identifying characteristics, or protecting them against use by certain applications/users, or by limiting their use to certain purposes. In this thinking, amendment 22 of the Lower House to Article 13 of the bill added that the prescribed security measures must also

³¹ By way of comparison, the United Kingdom’s Data Protection Act 1998 transposes the Directive by stating, in the seventh Principle in Schedule 1, that ‘[a] appropriate technical and organisational measures shall be taken against unauthorised or unlawful processing of personal data and against accidental loss or destruction of, or damage to, personal data’. In the German legislation, the stipulation has been adopted in § 3 Abs. 4 of the Teledienstedatenschutzgesetz and in § 12 Abs. 5 of the Mediendienste-Staatsvertrag. For more information, see: L. Gundermann, ‘Das teledienstedatenschutzgesetz – ein virtuelles Gesetz’, in H. Bäumler (ed.), *E-Privacy: Datenschutz im Internet*, Braunschweig: Vieweg, 2000, pp. 58-68.

focus on the prevention of unnecessary collection and further processing of personal data. This will provide a legal foundation for the application of PETs. Such rules respond to the restrictions of the developing information technology³².

Article 13 thus outlines that the person responsible for the processing of personal data takes suitable technical measures to protect personal data³³. Wherever technical measures are insufficient or unfeasible, organizational measures can be taken, or organizational measures can enhance the technical measures in a coherent package. Whenever there is a choice between organizational and technical facilities as part of a balanced processing policy, the Data Protection Board always prefers the latter. Technical measures are usually more efficient, as it is more difficult to escape their effects³⁴. While the definition of 'organizational measures' is left open, it is important to note that it draws attention to managerial and other human systems through which technical devices are put into effect, and therefore opens up a path towards shaping the regimes of staff accountability and responsibility within business firms or government agencies.

A further legislative provision is that the Lower House accepted Motion 31 of member Nicolai, in which the government is enjoined to apply such technologies to its own data systems. The explanatory memorandum of the Ministry of Home Affairs for 2001 indicates that 12 million Euros will be necessary each year to enforce the 'Contract with the Future' memorandum to give the government's role real meaning as a booster and user of technological innovation, including PETs³⁵. In its letter of 13 January 1999, the Data Protection Authority (Registatiekamer) pointed out to the Lower House that:

'This means that the person responsible (controller) will have to take suitable measures against collecting, recording and saving personal data in violation of the conditions set elsewhere in the WBP. In particular this means that collecting and processing personal data without sufficient basis as referred to in Article 8 of the WBP will have to be prevented. Article 13 of the WBP makes the person responsible translate the legal standards of the WBP into the actual processing of personal data and also take this into account when designing and further developing data systems'.

³² See Memo of Reply First Chamber regarding the WBP, op.cit. p. 16

³³ PETs offer not only the possibility to take suitable technical measures, but also a structural measure for the right application of the WBP.

³⁴ See Registatiekamer, *Advies Beveiliging Persoonsgegevens*, Rijswijk 1994, p. 11.

³⁵ See Parliamentary Documents 27400 VII, No. 2. During the debate on bills 26643 and 26387, the Minister for ICT within the Administration said that the budget had been raised to Dutch Florins 25 million per year enabling the administration to become a launching customer of PET-enabled systems; see Parliamentary Document 26387, No. 15, Parliamentary Year 2000-2001.

Article 13 WBP states that the measures taken shall be sufficient, taking into account what is technically possible at the time and the cost of implementation, to ensure a level of security that is adequate in relation to the risks associated with processing and the nature of the data to be protected. The measures shall be designed in part to prevent unnecessary collection and further processing.

It is therefore up to the data controller to put in place technical measures to protect the data to be processed against loss and against any form of unlawful processing. The measures taken should also serve to prevent the data from being collected unnecessarily and from being processed further.

The measures should be assessed by reference to:

- What is technically feasible
- The cost
- The risks associated with the processing and with the nature and volume of data concerned.

Where technical measures are not feasible or are not sufficient to provide the necessary protection, organizational measures may be taken instead, or in the context of a coherent technical and organizational package.

10. PET STRATEGIES

In applying PETs, the controller can choose two strategies: either focusing on preventing or reducing identification³⁶; or focusing on preventing unlawful processing of personal data, in accordance with the WBP. A combination of both is also possible.

A controller will often find it helpful to use PET in conjunction with organizational measures.

10.1 Preventing identification and the disproportionate effort criterion

In relation to the first strategy, it is important to determine whether personal data is involved and to consider the question of ‘identifiability’. The WBP defines personal data as any information relating to an identified or identifiable natural person. According to article 2 of DPD, a natural person may be directly or indirectly identifiable. A person may be directly identifiable from his/her name, address, personal reference number, generally known pseudo-identity or biometric characteristic (e.g. fingerprint). Indirect identifiability depends on (a combination of) other

³⁶ F. Rötzer, ‘Das recht auf Anonimität’, in H. Bäuml (ed.), *op. cit.* pp. 27-34.

unique characteristics or attributes, from which it is possible to extract sufficient information for identification. If, for example, one has a postcode and a house number, it is possible to identify the occupant of a house using a CD telephone directory.

PET can be used to anonymize direct identification data within an information system. If, in addition, all links are removed between such data and any indirect identification data, the data on the system is no longer personal data. Under such circumstances, the DPD ceases to apply, since there is no personal data requiring protection.

Indeed, a system may be considered not to contain personal data if the nature and amount of indirect identification data are such that disproportionate effort would be required to identify an individual from the data in question. Similarly, if the co-operation of a third party over whom the controller has no authority, and for whom the controller is not responsible, is required to identify an individual from the data held, the data is not considered personal. Whether the effort required for identification is disproportionate depends partly on the nature of the data and the size of the population. However, the time and money people are prepared to invest in identification is also relevant in this context.

10.2 Preventing or securing against unlawful processing of personal data

PET can be used to prevent personal data from being processed in various unlawful ways. Its application can, for example, stop such data from being unnecessarily collected, recorded, stored, distributed or disclosed (within or between organisations), and matched or shared, or associated with (linked to) other data.

By using PET in connection with the processing of identification data, a controller can set up an information system with an identity domain and a pseudo identity domain, so that fewer or no personal data is processed (e.g. in the context of collection or recording), or so that the processing of such data is minimized. The controller can also ensure that access to data and the ability to obtain non-anonymized data is dependent on the user's rights within the information system or, depending on the protocols within the data system, provision of or access to anonymized data are or are not allowed for various users. A user engaged in medical research, for example, would have access only to data from which identification was not possible. On the other hand, identification data would be available in a hospital on the basis of functional authorization and the relationship between the care provider and the patient.

One of the most important principles of privacy protection is that no more data should be collected or processed than is strictly necessary for the specified purpose. If an assessment indicates that the use of PET would reduce the amount of personal data being processed, thus enabling the controller to comply with this principle, such use is effectively a legal requirement. Furthermore, PET can be employed to block certain forms of processing, thereby helping to ensure that data is used only for a specified purpose. PET is also valuable in the context of information security. The use of PET can therefore be necessary to comply with Section 13 of the WBP, which – according to the explanatory information – applies to all activities involved in the processing of data.

In its letter of 13 January 1999 to the Lower House, the Regstatiekamer indicated that the criteria set out in Section 13 of the WBP should be used to determine whether it is reasonable to expect a controller to use PET. Hence, the easier it is to employ PET, the less reasonable it is for a controller not to do so.

10.3 Example: a hospital information system with PET

The PET Hospital Information System now in use at dozens of hospitals is divided into various domains. In some of these domains, data subjects are identified; in others they are anonymous. Furthermore, only authorized users are able to make data in pseudo-identity domains identifiable. In June 1997, an international software developer successfully installed an identity protector in the hospital information system.

The process of implementing the PET solution was as follows. Following a privacy audit conducted by the Data Protection Authority, the supplier of the hospital information system was asked to develop practical measures based on the theoretical identity protector definition in order to hide the true identity of the patient and the physician as well as the related information in the database of the information system (including the electronic patient record system).

Data relating to patients has been divided into three sets. The first set consists of direct identification items, such as the patient's name, address, date of birth, insurance details, etc (the identity domain). The second set is made up of all diagnostic and treatment data (the pseudo-identity domain).

In both domains, a patient number identifies the patient. A number encryption protocol is applied, however, and different numbers are used in each domain. No link can therefore be made between an item of information in the one domain, and an item in the other. This has the effect of making it impossible for anyone to obtain an overview of a patient's data without going through the authorized channels.

In the identity domain, a system of serial numbers is used to identify patients. The patient numbers used in the pseudo-identity domain are obtained by encrypting the serial numbers. The encryption protocol allows the original patient numbers to be obtained by decryption, with encryption and decryption performed by the application software. The encryption keys used by the protocol are made available – preferably by an independent and trusted third party – only once the identity of the application user has been confirmed.

A third domain has been created for statistical and medical research, containing aggregated data only.

Pseudo identities are used here by means of the identity protector to ensure that the person's identity cannot be established from related data in the database, and that the related data cannot be revealed once a person is identified. In the Client/Server architecture of this system, the required relations in the database have been removed and all information is encrypted. The identity protector manages access to the required information. A typical dialogue in this PET enriched database is as follows:

- Log in with name of physician;
- Check in table 'physician';
- Return sequence primary key of physician;
- Encrypt to pseudo identity ('pid') of physician;
- Search table 'care relation';
- Return sequence primary key of patient(s);
- Search table 'patient';
- Select patient;
- Encrypt to 'pid' of patient;
- Search table 'appointment' with 'pid' of physician and 'pid' of patient;
- Return appointment(s);
- Etc.

Nobody without a functional authorization in this PET system can process any data in it. Even if privacy protection is not at stake, using PETs provides for significant data security³⁷. The effect of PETs on the performance of the database has been tested in the hospital information system described above. The performance in efficacy and response time was not affected by the application of PETs functionalities. The cost for building these functionalities within was 1 percent more than without, due to the fact that PETs functionalities have been incorporated into the design (data model) from the beginning. However, the application of PETs in old, existing data systems is not always feasible. For example, opening up

³⁷ A number of known and realised PETs can be seen as measures that also effect general data security besides privacy protection. The application of PETs thus leads to 'enhanced corporate security'.

existing data systems to introduce an identity protector can be very expensive. In addition, the owner of the old data system often lacks the courage and will to carry out such operations as the 'spaghetti' often cannot be disentangled due to the many releases and patches. Thus the major opportunities for PETs are therefore in the design and implementation of new data systems.

10.4 Other privacy-supporting technologies

In situations where it is not possible to use the PET strategies described above, a mix of other technologies can be employed to protect privacy more effectively. On the basis of the basic privacy requirements referred to earlier, one may identify a number of preconditions for the processing of personal data, such as:

- Transparency
- Quality
- Respect of subjects' rights
- Security.

The following examples illustrate how technology can be used to strengthen privacy:

- Transparency can be improved by the use of P3P (a technology for assessing web-site privacy policy) if the default setting is appropriate. The default should be set so that not all data is automatically publicized.
- A statistical-linguistic analysis of an address system can optimize data accuracy and thus lead to the enhancement of data quality.
- The data subject's rights can be protected more effectively by using a system of feedback and checks. By adopting such an approach, the controller can ensure that the subject can be informed at any time what personal data has been committed to the information system. The subject can then ask to see the data, and request its supplementation, amendment or deletion.
- Logging – the maintenance of an electronic logbook recording procedures carried out on the information system – is an excellent protection measure. The origin of all collected and recorded data can be automatically logged, as can all operations involving the retrieval, consultation, amendment or internal/external distribution of data. Only the system controller should be able to delete logs, and all such deletions should themselves be logged and justifiable by the data controller.
- An electronic data-processing access logbook should also be maintained. Automatic access control can then be employed in connection with the screening, consultation, amendment, deletion or destruction of data.

- Another option is automatic data deletion. Retention periods can be specified in the software and data deleted automatically on expiry of the relevant period.
- It is also possible to employ technical measures to prevent unlawful processing (i.e. processing contrary to the WBP) in the context of processing by an external processor or transfer to a non-EU country via the Internet.

10.5 Combined technologies

A technology that enables the controller to comply with only one of the DPD's basic privacy requirements is not in itself sufficient for the realization of optimal privacy protection. For example, a statistical-linguistic analysis of an address system may provide a basis for enhancing data accuracy, but greater accuracy on its own does not guarantee privacy.

To create a satisfactory privacy environment, it is often necessary to stack technical measures within an information system. So, for example, statistical-linguistic analysis may be combined with distributed storage, the use of protocols on origin, application and distribution, and a logging system.

The implementation of PET is not simply a technical issue; it is also normative. Before putting PET into information systems, it is necessary to establish what requirements the DPD makes regarding such systems. Technologists and lawyers need to get together and translate standards into technical system requirements. Alternatively, it is necessary to conduct a PET scan or privacy audit to determine whether system requirements and applications are consistent with the DPD. Otherwise, an adequate level of protection cannot be said to exist.

If neither PET nor other technical measures are sufficient to provide adequate protection, organizational measures will be required. Although existing ICT applications are theoretically capable of ensuring that any information system meets at least one of the basic privacy requirements, it is sometimes unreasonable to require implementation, since the cost is out of proportion with the interests at stake. Under such circumstances, organizational measures should be taken to protect the privacy of the information system user, or of the consumer or citizen whose data is to be processed.

11. NEW DEVELOPMENTS IN PET AND P3P

Since 1997 the Internet has been exploding every year the number of people online was more than doubling, but more and more users of the Internet started to realize that the Internet isn't a safe haven. Polls from Forrester and IDC pointed out the due to a lack of trust the e-commerce business and e-government didn't grow as fast as expected. Privacy violations were one of the reasons why growth stagnated. PETs were developed in order to provide protection for the visitors to cyberspace³⁸. The problem of identity theft became an increasing bigger problem in USA and Canada. Privacy threats augmented within Internet and many e-commerce companies developed privacy invasive technologies to track behavior of Internet visitors. Since 1997 a lot of code has been written to protect privacy in Internet. Anonymous remailers with PGP encryption for email were one of the first products. Mix master remailers and onion routing were launched in order to protect the identity of the sender. To securely receive email the use of a 'newnym' style server was introduced³⁹ and 'Crowds' was born as the adagium was "Anonymity Loves Company". Zero-Knowledge introduced commercially the Mix network Freedom⁴⁰ in which PET design principles are used based on anonymity, unobservability and Identity protectors. Work is in progress to hide Protection of Personal Identifiable Information (PII) (indirect identifiable personal data) for unscrupulous Internet marketers. Also tools to bar and counter attack cookies are in the make⁴¹. Under development also are private credential technologies to allow the user prove things about him or her without revealing extra personal information⁴². An important development is P3P developed by W3C.

P3P⁴³ is a tool for easy communicating about your own privacy preferences in a standard machine-readable format. The vocabulary of P3P

³⁸ I. Goldberg, Privacy-enhancing technologies for the Internet, II: five years later, in Pre-proceedings of the 2nd workshop on privacy Enhancing technologies, San Francisco, April 14-15, 2002 p.4

³⁹ Goldberg, op.cit. p.5

⁴⁰ The Freedom Network Architecture, Zero-Knowledge Systems, Inc. 1998, <http://www.freedom.net>

⁴¹ Junkbusters Corporation. Internet Junkbuster Headlines, <http://internet.junkbuster.com/ijb.html>

⁴² S.A. Brands, Rethinking Public Key infrastructure and Digital Certificates-Building in Privacy. Cambridge (MA)/London 2000

⁴³ See <http://www.w3.org/P#P/> and see P3P and Privacy – Center for Democracy & Technology / IPC Ontario – see <<http://www.cdt.org/privacy/pet/p3pprivacy.shtml>>. However, there are doubts about the efficacy of P3P; see J. Catlett, 'Open Letter to P3P Developers & Replies', in *CFP2000: Challenging the Assumptions, Proceedings of the Tenth Conference on Computers, Freedom & Privacy*, New York: Association for Computing Machinery, 2000, pp. 155-164 (also available at

shows who is collecting data, what data is collected, for what purpose data will be used, what are the opt-in and opt-out alternatives for some data uses, who are the data recipients, to what information does the data controller provide access, what is the data retention period, how will disputes about the policy be resolved and where is the human readable privacy policy. Such a policy (in order to achieve a well-balanced processing policy for personal data and to implement and maintain this properly) must occupy an important place in the management cycle. A privacy policy should be based upon established privacy principles, such as the nine privacy principles described earlier in this paper and regulatory requirements specific to the region(s) the company is operating in.

Every company should have a privacy policy posted on the Web to address the needs of its privacy sensitive customers. It should reflect the organization's privacy policy and practices. The privacy statement is a critical mechanism for soliciting and winning consumer trust. The privacy policy should address the privacy principles the company has decided to adhere to, what information is collected, how information may be used and/or shared, how personal information is protected, and a central point for questions and dispute resolution. The above is not an obligation under the DPD, but from a point of view of transparency it is strongly recommended. However many companies haven't yet posted their privacy policies on their web pages and thus you could only fear for your privacy.

In this context it should be recalled that the Article 29 DPD Working Party issued an opinion concerning P3P already some years ago⁴⁴, and the P3P consortium didn't take on board that concerns expressed by the members of the Internet Task Force. In the Internet Task Force report⁴⁵ the Working Party underlined the fact that a fair number of issues still needed to be addressed. It also emphasized that, even if these issues were to be solved, the following limitations of the P3P system should be taken into account:

<<http://www.junkbusters.com/ht/en/standards.html>>). Technological systems-design solutions are discussed and encouraged at length in L. Lessig, *Code and Other Laws of Cyberspace*, New York: Basic Books, 1999; but see the critique in P. Schwartz, 'Beyond Lessig's *Code* for Internet Privacy: Cyberspace Filters, Privacy Control and Fair Information Practices', *Wisconsin Law Review*, vol. 2000, No.4, pp. 743-88. Another range of consumer privacy products has been developed by the Canadian-based Zero-Knowledge Systems, see <<http://www.zeroknowledge.com>>.

⁴⁴ Opinion 1/98: Platform for Privacy Preferences (P3P) and the Open Profiling Standard (OPS), adopted on 16 June 1998.

⁴⁵ Working Document Privacy on the Internet - An integrated EU Approach to On-line Data Protection- adopted on 21st November 2000.

- P3P cannot protect the privacy of users in countries with insufficient privacy laws: it does not have the ability to create public policy, nor can it demand that its specifications be followed in the marketplace.
- P3P cannot ensure that companies follow privacy policies. In fact, P3P cannot guarantee that the site is doing what it claims to do. The sanctions for failure to comply with a declaration of intent can only be set by law or through membership of a self-regulatory body.

At the EU Joint Research Centre (JRC) P3P Demonstration Project evaluation meeting in Ispra (Italy) on 27 May 2002 main group of issues was discussed directly related to the P3P specification itself as a means for privacy protection.

An important issue that was particularly discussed was the fact that P3P allows transactions with a level of privacy protection below the basic criteria as interpreted from the DPD. If P3P were to be used in a configuration, which allowed a user to consent to the waiving of his/her right to access his/her data, it would be against the Directive.

12. PISA, AN EXAMPLE OF PET RESEARCH

PISA is a EU funded project that deals with intelligent software agents (ISA) that protect privacy on Internet. An Intelligent software agent (ISA) has been described⁴⁶ as software and/or hardware that are capable of acting autonomously in order to accomplish a task on behalf of its user in a complex network environment⁴⁷. Thus it can act on its own without interference of the user. The ISA may have properties and attributes like mobility, deliberative behavior, and interaction with other ISAs, possibilities for identification and learning. In the PISA project the ISAs are split into a generic personal agent, specific task agents and service agents. This is reflected in the MAS⁴⁸ architecture. In order to perform properly ISAs need Personal Identifiable Information (PII, personal data) and privacy preferences (a profile of personal preferences and aversions for goods and services) and non-PII. ISAs thus capsule lots of personal information about their users or employees of their organizations. Furthermore software agents

⁴⁶ J.J. Borking, M. van Eck, P. Siepel, *Intelligent Software Agents and Privacy - A&V # 13-The Hague 1999* p. 6,9-10

⁴⁷ Many definitions of agent technology exist. Living Systems AG summarises all the important characteristics of ISA as: "software objects that proactively operate on behalf of their human masters in pursuing delegated goals"

⁴⁸ MAS: Multiagent System. In MAS is embedded decision trees that can learn from sample data and constraint satisfaction programming based on a set of constraint rules. Logic engine frameworks can extend a set of 'action selection strategies'. For design approaches see: J.J. Borking et al, *supra* footnote 1, p.14-16

send and receive (messages with) information (and requests, etc.)⁴⁹. The ISA gathers relevant task information, and personal data of both its user and others. The ISA will reveal personal data of the user according to his or her privacy protection level preferences. An agent can receive messages and perceives its environment (input) and sends messages and acts in its environment (output). The objective of the PISA project is to prove that the informational privacy⁵⁰ of the user is protected in all kinds of processes by incorporating privacy-protecting features into an ISA⁵¹.

In order to create maximum privacy protection within the PISA project personal data is divided into three categories:

Level 1: Contact Information. This information enables the identification of the data subject without the assistance of third parties. Level 1 examples are: Name and Address information, telephone number, and email address either using a real identity or a pseudo-identity.

Level 3: Special categories of personal data as defined in Directive 96/46/EC Article 8 paragraph 1. Level 3 personal data are only to be processed under the conditions specified in Article 8 paragraph 2 through 7.

Level 2: All others items of personal data.

PET means will protect Level 1 PII by encapsulating these data within the ISA directly at the time of collection. Only the recipient who is entitled to receive level 1 PII will have access to the means to unwrap the PET encapsulated level 1 personal data. The PISA project uses PET to protect all personal data with traditional PET components like identity protectors and the creation of anonymity and pseudo-identity domains. When all PII are modified such that the data subject is no longer identifiable the DPD will not be applicable any more.

13. PRIVACY KNOWLEDGE ENGINEERING (PYKE) AND DEPRM

In the summer of 1999 the Technology Assessment Group of the Dutch data Protection authority brainstormed about Lessig's opinion "Code is Law"⁵². Could this adagium also be reversed: Law is Code? If so, then privacy

⁴⁹ cf. speech acts

⁵⁰ Besides informational privacy, one can discern bodily, territorial and communicational privacy. Informational Privacy is a term that is sometimes used to identify the privacy issues that are related to the processing of data within the ICT domain.

⁵¹ See PISA contract no IST-2000-26038

⁵² L.Lessig, *Code and Other laws of Cyberspace*, Basic Books New York 1999

law could be translated in machine-readable code by which automatically the law could be built into information systems and while processing personal data privacy protection would be applied. This was the birth of the third branch of PET, in 2002 dubbed as PYKE, privacy knowledge engineering.

Problems arisen out of changes in the law are well known in conventional data processing: changes in tax law for example, must be announced well in advance of coming into effect so as to allow time for the considerable task of altering programs which have to apply these laws in payroll and other applications. However, the privacy principles as described under heading 2 are well accepted over a period of more than 20 years and therefore considered to be stable enough to be implemented in information systems,

In order to build the nine privacy principles⁵³ into the ISA a method has been developed to realize privacy knowledge engineering. The method has been named Design Embedded Privacy Risk Management (DEPRM)⁵⁴. The basic proposition of DEPRM is that in order to assure a system design against privacy risks as discovered in the privacy risk analysis, the privacy rules must have a strong influence on the design forcing the system designer to obligatory represent the data protection rights and duties into the system. As the legislation is often complex and laced with many exceptions, the solution is found in working with a simplification of the data protection legislation and using relatively stable privacy principles while retaining the capacity to integrate more aspects of the law, as they would be applied when circumstances require so. For example: medical personal data are considered to be privacy sensitive data and the law requires therefore a strict regime while processing these data⁵⁵. The legal instantiation procedure in DEPRM works as follows:

First privacy principles as for example known from the Convention 108 of the Council of Europe and the 1980 Organization for Economic Co-operation and Development (OECD) guidelines⁵⁶ are determined. Then a simplification of the law is realized through the linking together ('chaining') of selected articles of the DPD that belong to the chosen privacy principles. Take for example the principle of transparency that defines that everyone must be informed about what is done with his/her personal data and that the

⁵³ The privacy principles are well accepted over a period of more than 20 years and therefore considered to be stable enough to be implemented in information systems.

⁵⁴ S. Kenny, J. Borking, *The Value of Privacy Engineering*, in *The Journal of Information, Law and Technology (JILT)*, 2002 (1), <http://elj.warwick.ac.uk/jilt/02-1/kenny.html>. Kenny presented the concept of DEPRM for the first time at the second CEN/ISSS data privacy IPSE (Initiative on Privacy Standardisation in Europe) open Workshop in Paris on 27 September 2001. See Final IPSE Report IPSE-SG #11, Doc.no 7 – 13 February 2002

⁵⁵ See article 8 of the DPD, which defines the special categories of data, of which personal health data is one of them.

⁵⁶ 'Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data'

data subject must also be informed of this as the data subject is mostly not aware of the multitude of parties that take part in the processing nor in the data flows. This principle is a concatenation of the articles 10 a, b, c, 11 section 1 a, b, c, 11 section 2, 13 section 1 a, c, d, e, f, g, 13 section 2⁵⁷. The principle of finality and purpose limitation (personal data should be collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes) can be found in articles 6 section 1 b, e, and article 6 section 2.

Thus such representation clusters the DPD articles into logical components generic enough to encapsulate the general meaning of the articles from which they derive. Some articles are not 'translated' into principles and are not implemented into the MAS design but play a role for explanation and interpretation like articles 2 a till h (definitions), 3 section 1, 3 section 2, 4, section 1 a, b, c, 4 section 2, 5, 9, 15 section 1, etc. and influence the architecture. Some articles are not relevant for the MAS representation like the requirements for the supervisory authorities, the legal remedies etc.

Having the privacy principles formulated the next step is splitting the principles into the elements (pieces of text that are structured subsets requiring orthogonality (completely independent factors in the mathematical sense)⁵⁸ and regularity. Thus the principles are deconstructed into a set of elements that relate to the articles they are derived from and emphasize the context and what has to be achieved. Then each element is given a work item identifier. For example: 1. Principle: Transparency; 1.1. Data subject (DS) is aware of transparency opportunities; 1.1.1. Personally Identifiable Information (PII level 1 till 3)) collected from DS; 1.1.1.1. Prior to DS PII capture: DS informed of: controller identity (ID) and Purpose Specification (PS) 1.1.1.2. Prior to DS PII capture: DS informed of: controller ID/PS and re article 10 c whether PII contains special categories of data and whether PII have to be treated accordingly to the sensitivity of these data Further more elements have to be formulated around information duties, PII distribution, explicit consent, controller obligations (for example 2.2. A controller must process according to his PS), about finality (DS presented with controller PS & a proposed retention period (RP), prior to being presented with a opt in consent decision for having PII processed, except where PS is statistical etc.) All principles are treated as stated above. This is the final level of granularity that can be reached to subdivide the privacy principles before formulating into technical specifications. The next step in

⁵⁷ The article 29 DPD Working Party states in Privacy on the Net WP Working Document 37 how transparency should be applied in on-line systems. For Recommendations for online collection of data see www.europa.eu.int/comm/privacy

⁵⁸ Webster's Third New International Dictionary 1986 p.1594

terms of logic is to find the ontology⁵⁹ a mechanism to achieve shared understanding, neutralizing the problem of two databases using different identifiers for what is the same concept, such as postal code Ontology and logic are different aspects of a system design, there are mutual dependencies, but they are separate complementary models (a summary of the kinds of entities and specifically the kind of abstract entities that are to be admitted to a language system leading to related concepts and terms for achieving that all ISAs act in the same way while transferring policies) and data model for a principle as a representation of the relevant work item identifiers⁶⁰. These ontologies with taxonomies and a set of inference rules providing the relational aspects of the taxonomies lead to a simplified conceptual model of the principle and may then be implemented into ISA and used as a backbone in conversations between agents about privacy preferences, the matching of privacy policies and exchanging PIIs. Thus the sender agent and the receiver agent need to understand the same privacy ontology in which a 'privacy policy' is defined. The agent needs this to be able to automatically give the same interpretation to the information in messages they exchange in their communication. The ontology also needs to define how a privacy policy is structured, i.e. what information can be in it and how this is composed into the privacy policy. The ontology also needs to define how the system of privacy preferences is structured, because these preferences also need to be sent from one task agent to another and thus has to interpret by a receiving agent. To realistically reflect the simplified legal model the logic has to be described in a format understood by the application, the agent platform. The ontology is now enriched with a knowledge base to enable interpretation of the queries between the agents, like: if PII is labeled with high constraints then the ISA request from the interacting agent the Agent Practices Statement (APS), privacy policy under which it operates⁶¹. Transfer rules are embedded in ISA to decide when PII can be transfer to other agents. The

⁵⁹ Webster op. cit p.1577 See for further information: Newell, A. (1982). The knowledge level. *Artificial Intelligence*, 18(1), p.87-127. Levesque, H. J. (1984); *Foundations of a functional approach to knowledge representation*. *Artificial Intelligence*, 23, p. 155-212. Genesereth, M. R., & Nilsson, N. J. (1987). *Logical Foundations of Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann Publishers, Sowa, John F *Knowledge representation: logical, philosophical, and computational foundations*, Pacific Grove Brooks/Cole, 2000. <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>

T. R. Gruber. A translation approach to portable ontologies. *Knowledge Acquisition*, 5(2): p.199-220, 1993

http://ksl-web.stanford.edu/KSL_Abstracts/KSL-92-71.html; <http://www.w3.org/2001/sw/>
<http://www.w3.org/TR/webont-req/>;

<http://www.semanticweb.org/knowmarkup.html#ontologies>;

http://protege.stanford.edu/publications/ontology_development/ontology101.html

⁶⁰ Kenny, op.cit. p 17.

⁶¹ Kenny, op.cit. p18-19

general transfer rule is: IF APS-1 MATCHES privacy-preference-2 AND APS-2 MATCHES privacy-preference-1 AND PII level 2 -1 MATCHES PII level 2 -2 THEN ALLOW disclosure/exchange PII level 1 -1. Many other items are taken into consideration. Security characteristics have to be implemented. What is the identity of the other agent, who is he representing? Can the ISA trust the other agent? Controller identity, purpose specification credentials have to be exchanged and interfaces have to be designed to prevent detection of the ISA before the analysis of the situation shows a safe environment.

14. STRUCTURE OF PISA

In order to protect the ISA against attacks, security risks and privacy intrusion the ISA has an anonymity shell and built-in privacy protection related functions, Privacy-Enhancing Technologies (PETs) for pseudo identities, built-in legal know-how, and mechanisms to enforce legal aspects⁶². The ISA collects, processes and disseminates personal data and other transaction relevant information. PISA offers two privacy limit options: an ISA operation under anonymity or maximum attainable pseudonymity or under the DPD. During the operation of the PISA the process is executed on PII Level 2 items only. That means in complete anonymity. Once a contact is to be made with the organization the search had been set up for, the PET protected (encrypted) PII Level 1 are disclosed to the organization. The organization is now in a position to negotiate with the TTP to decrypt the contact information. Thus the whole process is within the boundaries of the DPD.

⁶² See model of Kathy Cartrysse, deliverable WP 3 D 10

Figure 7. Structure of Privacy Incorporated Software Agent (PISA)

The Agent Practices Statements (APS), privacy policies and privacy statements the ISA encounters in the networks in which it 'travels' within the legal realm of the EU should offer a minimum and non-negotiable level of protection for all individuals falling within its application⁶³. In our global Internet the ISA will be confronted with all kinds of legal regimes that are not on the same level as that of the EU and ISA should handle this with flexibility according to its user wishes. It is envisioned that in PISA Demonstrator the default standard will be the EU level. If ISA encounters lower privacy protection levels than it will report to the user and request for an explicit consent to transfer personal data. This is applicable if the data is to be disclosed to an environment without EU compliant privacy protection.

As P3P can't be adapted immediately within ISA environments, within the PISA-project the legal instantiation tool DEPRM will be used to reflect the privacy preferences. This new tool compares like P3P privacy policies

⁶³ See document adopted by the Working Party concerning P3P – Opinion 1/98: Platform for Privacy Preferences (P3P) and the Open Profiling Standards (OPS) 16 June 1998, http://europa.eu.int/comm/internal_market/en/dataprot/wpdocs/wp1 len.pdf

with ISA privacy preferences. It will alert and advise the ISA what to do. The privacy policy built into the ISA has to act according to the level of privacy preferences of the user. The highest possible protection would mean that all nine principles would be applicable and more. The lowest possible choice of protection would mean that all nine principles would be applicable at the lowest level. For example, the minimum position for transparency would be constant logging of the disclosure of highly sensitive personal data by the ISA and his whereabouts.

Ideally if personal data are transferred to other ISAs or controllers, than these data must have a marker or tool that prevents unwanted processing like secondary use incompatible with the primary purposes for which data are processed. For example if the privacy preferences state the other party may use the personal data for a specific purpose and a specific retention time, those data should be automatically deleted accordingly⁶⁴. No reliable technological solution exists yet that gives the sender of data personal control over data once data has been transferred from a sender agent to a receiving agent, albeit that ideas are developed that facilitate controlled distribution of digital personal information. As long as the privacy preferences of the ISA are not matched by encountering ISAs or websites or security hubs the ISA will function anonymously or under a pseudonym or no contact will be made at all because of privacy or security risks.

15. DISCLOSURE OF PII – DIFFERENT LEGAL REGIMES

Disclosure of personal data (PII level 1, 2 and 3) by a sending agent to a receiving, task or generic, agents should only be permitted when the privacy policy of the recipient of the personal data fits the privacy preferences of the sending task agent. Therefore the sending agent should ensure that the receiving agent adheres to the minimum provisions of the DPD as stated in the APS of the receiving agent. With regard to onward transfers to regions and countries that are subject to different legislation the ISA has to verify whether the privacy protection regime is compliant with the EU DPD. Personal data traffic between EU and USA are ruled by the Safe Harbor arrangements that has to be closed by each specific company and the EU in order to permit onward transfer of personal data⁶⁵. Otherwise such transfer of

⁶⁴. See S. Kenny, L. Korba, H. Reddy, Privacy Rights Management for Digital Rights Management to be published in 2002

⁶⁵ Safe Harbour Principles, US Department of Commerce 21 July 2000

data is forbidden. Some countries, like Canada⁶⁶, are considered under article 25 paragraph 2 of the DPD as providing an adequate level of protection for personal data transferred from the EU to recipients in Canada. If the privacy preferences are not matched by an ISA then either the exchange of non-PII doesn't occur or other possibilities as mentioned in article 26 1 DPD, or a special contract for protecting the personal data has to be negotiated⁶⁷. Due to the fact that the privacy preferences always are on the level of the DPD any transfer outside the EU to other agents matching these privacy preferences must be qualified in legal terms as adequate irrespective of the fact that the EU Commission hasn't qualified yet the legal regime in which these agents operate as adequate.

As the controller will be accountable, the user (the data subject) and controller will be compelled to use ISAs that can deal with privacy. The situation is comparable with buying a car. Nobody wants a car that violates existing car legislation. No manufacturer will build cars that are not compliant with the existing legislation. Therefore the developer will be pushed into a design that protects privacy.

16. CONCLUSION

Developments in information and communications technology are constantly broadening the scope for the collection, storage, processing and distribution of personal data online and offline. This constitutes a growing threat to the privacy of consumers and members of the public. However, other developments in the same field offer ways of protecting the rights and interests of users and those to whom data relates.

PET is an excellent and promising means of satisfying the basic privacy requirement that all data processing should have a legal basis. Naturally, further consideration and research remain necessary, and continuing efforts will have to be made to promote the use of PET in information systems – as, for example, through the EU-subsidized PISA Project. This project – PISA stands for privacy incorporated software agent – is concerned with the development of a guaranteed-privacy environment for Internet users.

In addition, PET test criteria have to be developed like in PETTEP. Furthermore privacy audits and special PET scans are required to determine whether PET-equipped systems actually comply with the DPD. Certification

⁶⁶ EU Commission Decision 2002/2/EC of 20 December 2001 concerning the Canadian Personal Information Protection and Electronic Documents Act

⁶⁷ There is a EU Commission Decision and standard contract clauses for such transfers available. See http://europa.eu.int/comm/internal_market/en/dataprot/modelcontracts/index.htm

in the context of a privacy audit can be helpful in this regard, and can serve to reassure consumers and members of the public that their privacy is adequately protected by an information system.

Privacy Commissioners can stimulate the deployment by actively promoting PET in information systems and PET research. The law alone can't protect privacy. Lawyers and technologists like in PYKE should proactively try to solve privacy problems in lieu of reactively waiting for complaints to be dealt with when harm has already be done.

REFERENCES

Privacy-Enhancing Technologies (PET)

H van Rossum, H Gardeniers, JJ Borking, A Cavoukian, J Brans, N Muttupulle, N Magistrale, Privacy-Enhancing Technologies: The Path to Anonymity, The Hague 1995.

R Hes and JJ Borking, Privacy-enhancing technologies: the path to anonymity, revised edition, Background studies and surveys 11, Registration Office, The Hague 2000.

O Berthold, A Pfizman, R Standke, 'The disadvantages of the free MIX router and how to overcome them' in: Proceedings of the Workshop on Design Issues in Anonymity and Unobservability, Berkeley (CA), pp 27-42.

WBP (general)

LG Sauwerwein and JJ Linnemann, Handleiding voor verwerkers van persoonsgegevens. Wet bescherming persoonsgegevens, Ministry of Justice, The Hague January 2001.

Privacy in bedrijf, AWMN, FME-CWM, VNO-NCW, The Hague 2000.

Audit Project

Wet bescherming persoonsgegevens Raamwerk Privacy Audit, Audit Approach Consortium, The Hague 2001.

Wet bescherming persoonsgegevens WBP zelfevaluatie, Audit Approach Consortium, The Hague 2001.

The protection of personal data

GW van Blarckom and JJ Borking, Beveiliging van Persoonsgegevens, Background studies and surveys 23, Registration Office, The Hague 2001.

Trusted third parties

J Versmissen, Sleutels van Vertrouwen, TTP-dienstverlening en privacy – de juridische randvoorwaarden verkend, Background studies and surveys 22, Registration Office, The Hague 2001.

Privacy incorporated software agent (PISA)

JJ Borking, 'Privacy Incorporated Software Agent (PISA): A Proposal for Building a Privacy Guardian for the Electronic Age', in: H Federath, Designing Privacy Enhancing Technologies, Springer Verlag, Berlin 2001, pp 130-140.