International Journal of Medical Informatics (2005) **74**, 257—265

ELSEVIER

International Journal of Medical Informatics

www.intl.elsevierhealth.com/journals/ijmi

Privacy protection for clinical and genomic data The use of privacy-enhancing techniques in medicine

B. Claerhout[a], **G.J.E. DeMoor**[b,*]

[a] *CUSTODIX nv, Merelbeke, Belgium*
[b] *Department of Medical Informatics and Statistics, University Hospital Gent, Building 3, 5th floor, De Pintelaan 185, 9000 Gent, Belgium*

Received 7 November 2003; accepted 11 March 2004

KEYWORDS
Genomic medicine;
Confidentiality;
Privacy-enhancing techniques;
Trusted third parties;
Pseudonymisation

Summary Privacy includes the right of individuals and organisations to determine for themselves when, how and to what extent information about them is communicated to others. The growing need of managing large amounts of both clinical and genetic data raises important legal and ethical challenges. This paper introduces some of the privacy-protection problems related to classical and genomic medicine, and highlights the relevance of trusted third parties and of privacy-enhancing techniques (PETs) in the context of data collection, e.g., for research. Practical approaches based on two pseudonymisation models, for both batch data collection and interactive data storage, are presented.

The actual application of the described techniques today proves the possible benefits for medicine that innovative privacy-enhancing techniques can provide. Technical PET solutions can unlock valuable data sources, otherwise not available.

1. Introduction

In today's information society, vast amounts of personal data are collected, stored and processed. Much of this data has a sensitive nature (e.g. medical data), and although generally used for the benefit of the community, it can be easily abused by malicious people.

Incidents are frequently reported in the public media, but concern about the proper treatment of sensitive data in general grows only gradually. However, people tend to become more apprehensive when their personal healthcare-related data are at stake, mainly because they can easily imagine motives for abuse and assess its impact. An obvious case in point is that at some point in their life practically everyone is confronted with loan and insurance applications. Recent incidents such as the one in which an outsourced transcriber threatened to disclose all medical records she had been pro-

* Corresponding author.
E-mail address: georges.demoor@ugent.be (G.J.E. DeMoor).

1386-5056/$ — see front matter © 2004 Published by Elsevier Ireland Ltd.
doi:10.1016/j.ijmedinf.2004.03.008

cessing for a US hospital [1] clearly illustrate that the threat to privacy is genuine. Public authorities are also sharply aware of these repercussions, and they are putting considerable effort into privacy protection legislation [2,3]. Today, it cannot be denied that privacy protection directly impacts personal well-being as well as society as a whole. Indeed, some go as far as to believe that failure to protect privacy might lead to our ruin [4]. Privacy is in fact recognised as a fundamental human right.

Genomic medicine is still in its infancy, but it is already evident that medicine, genomics and information and communication technologies (ICT) will continue to develop in some sort of symbiotic evolution [5,6], making unprecedented amounts of sensitive information available.

Genomic medicine encompasses predictive and diagnostic genetic testing. It can also use the information that derives from this testing to select or to fashion the best drug and therapeutic regimen for a patient, i.e. one that maximizes efficacy and minimizes side effects: pharmacogenetics is one of the avenues which will lead toward individualized health care and health maintenance.

Both genetic testing and pharmacogenetics give rise to concerns about the proper collection, storage and use of individually identifiable genetic information [7]. As the practice of genomic medicine develops, researchers and healthcare providers may want to store genetic profiles to determine treatment modalities as the need arises. The existence of such genetic databases will even increase the risk that unauthorized persons will obtain access.

Clinicians and researchers will therefore need to safeguard the confidentiality of such sensitive patient information.

Institutional review boards (IRBs) already pay careful attention to the requirement of obtaining the informed consent from subjects [8]. However, assessing the impact of the information enclosed in genomic data is very complex (see further), there is thus a real danger that informed consent is rather an ill-informed consent. Research ethics and security guidelines demand research units to divert more and more resources and time to privacy and identity protection, but burdensome requirements governing the transmission of medical and genetic information could unnecessarily discourage research. Well-intentioned privacy laws should not clash with the legitimate use of information when clearly to the public's benefit.

Protecting human rights (e.g. privacy) while maximizing research productivity is one of the coming challenges. A first step towards this goal is the research and implementation of technical solutions to the privacy problem. Privacy-enhancing techniques or technologies (PETs) provide means to unlock invaluable data sources for the benefit of society without endangering individual privacy.

This paper focuses on the possible use of privacy-enhancing techniques in the context of research and statistics for health care.

2. The nature of genetic data

In order to assess the difficulty of privacy protection of genetic data, it is important to analyse the nature of this data. Genetic data have the following specific characteristics:

- Genetic data not only concerns individuals, but also their relatives, thus people who have not been tested directly;
- Medical data deal with past and current health statuses of persons, whereas genetic testing can also give indications about future health or disease conditions;
- An individual person's genotype is almost unique and stable, hence it can become the source of an increasing amount of information;
- The full extend of the information included in the demonic data is not known yet;
- Personal genetic profiles can directly be derived from tissue samples.

A widely discussed problem is that, unlike other data from, e.g. clinical health records, genetic information is rarely about one single individual. A person's consent to release his or her genetic information constitutes a de facto release of information about other individuals, i.e. his or her relatives. In the case of genomic medicine, there is a complex interaction between individual rights and collective requirements.

Any collection of blood samples linked to identifiable persons can have an enormous impact on privacy; any material containing DNA is a potentially attractive source that can be mined for improper purposes.

Considering the risk of stigmatisation of particular subpopulations, the predictive and diagnostic testing for susceptibilities to disorders also remains problematic. This is even being complicated by the fact that some patients suspected of having a genetic disorder (e.g. Alzheimer) may lack the capacity to give their informed consent for a genetic test [9].

Given the potentially long latency period before symptoms develop, discrimination is another threat (e.g. insurers might use the results of diagnostic and

predictive testing to calculate health risks and set premiums).

The question will be whether the perceived short- and long-term benefits exceed the risks of ''improper access and use'' and what technical privacy measures can be taken to reduce such risks. The fact that physicians who will prescribe drugs without genetic testing could even face the risk of malpractice liability, complicates the search for the balance between privacy protection and clinical utility.

A couple of basic approaches to safeguarding confidentiality have been identified in the past. The first approach focuses on the creators and maintainers of the information, prohibiting them from disclosing the information to inappropriate parties. An alternative approach focuses on the use of privacy-enhancing techniques (PETs), which is technology to safeguard privacy. PETs eliminate or minimize the collection of personally identifiable information [10].

3. Privacy-enhancing techniques

There are many situations in which privacy can be an issue, accordingly PET research covers many different areas, including:

- Anonymous communication (anonymous remailers, anonymous surfing, etc.),
- Anonymous transactions,
- Anonymous publication and storage,
- Anonymous credentials,
- Anonymity in files and databases.

This paper focuses at medical applications, in which privacy issues are raised by the information content of the stored data, hence only the latter techniques are discussed. Privacy-enhancing techniques for privacy protection within databases help to protect the privacy of a subject of a database record (i.e. a person or organisation listed in the database). Simply put, these PETs allow to store relevant and useful information in a way that no one can ever find out, who the information is actually about. Examples of these techniques are (non-exhaustive list):

- ''Hard" de-identification by the owner of the data;
- Various types of anonymisation and/or pseudonymisation;
- Privacy risk assessment techniques;
- Controlled database alteration (modification, swapping or dilution of data);
- Data flow segmentation;
- Privacy-enhancing intelligent software agents for databases.

Today, PET technology has already proven its usefulness for privacy protection in health-related marketing and research data collection [11]. Focus lays on pseudonymisation techniques, and complementary PETs enhancing their effectiveness. Because of the proven track record of pseudonymisation, standardisation bodies like the CEN/TC251 committee (the standardisation committee for healthcare) are starting to put effort into including it into their work packages.

Currently, one of the CEN/TC251 work items is AURTAF (anonymity user requirements for trusted anonymisation facilities), which is still in a draft stage [12].

The first part of this work mainly focuses on terminology, because the terminology of the relatively new PET technology, it is not yet entirely understood nor fully adopted. Confusion and misunderstandings are therefore common when discussing the subject, a unified terminology could avoid this. For example, the AURTAF proposal does not define the term pseudonymisation, but considers the technique a specific form of anonymisation. Other literature, however, refers to anonymisation, only as the technique in which all nominative information is simply removed (''hard de-identification"). Mainly French experts, pioneers in the application of pseudonymisation technology, have an extensive terminology for all kinds of anonymisation/pseudonymisation forms. It is, however, not expected that these terms will become universally adopted as their translated forms may raise some questions and be counter-intuitive to non-native French speakers.

The AURTAF draft further proposes an approach for the analysis of the anonymisation needs.

4. Pseudonymisation techniques

Pseudonymisation refers to privacy-enhancing techniques and methods used to replace the true (nominative) identities of individuals or organizations in databases by pseudo-identities (pseudo-IDs) that cannot be linked directly to their corresponding nominative identities [13].

When data is being pseudonymised, identifiers and ''payload data" (non-identifying data) are separated. The pseudonymisation process translates the given identifiers into a pseudo-ID by using secure, dynamic and preferably irreversible cryptographic techniques (the identifier transformation process should not be performed with translation

tables). For an observer, the resulting pseudo-IDs are thus represented by complete random selections of characters.

This transformation can be implemented differently according to the project requirements. Pseudonymisation can:

- always map a given identifier with the same pseudo-ID;
- map a given identifier with a different pseudo-ID;
- time-dependant (e.g. always varying or changing over specified time intervals);
- location-dependant (e.g. changing when the data comes from different places);
- content-dependant (e.g. changing according to the content);
- etc.

Pseudonymisation is used in data collection scenarios where large amounts of data from different sources are gathered for statistical processing and data mining (e.g. research studies). In contrast with horizontal types of data exchange (e.g. for direct care), vertical communication scenarios (e.g. in the context of disease management studies and other research) do not require identities as such: here pseudonymisation can help find solutions.

It is a powerful and flexible tool for privacy protection in databases, which is able to reconcile the two following conflicting requirements: the adequate protection of individuals and organizations with respect to their identity and privacy, and the possibility of linking data associated with the same data subject (through the pseudo-IDs) irrespective of the collection time and place.

Because of this flexibility, however, correct use of pseudonymisation technology is not as straightforward as often suggested. Careless use of pseudonymisation technology could lead to a false feeling of privacy protection. The danger mainly lies within the separation of identifiers and payload. One should make sure that payload data does not contain any fields that could lead to indirect re-identification, i.e. re-identification based on (information) content, not on identifiers.

The key to good privacy protection through pseudonymisation is thus careful privacy assessment. Privacy gauging or privacy risk assessment is measuring the risk that a subject in a ''privacy-protected" database (in this case with pseudonymisation) can be re-identified without cooperation of that subject or against his/her will. This consists in measuring the likelihood that a data subject could be re-identified using the information that is available (hidden) in the database. The lower this re-identification risk, the better the privacy of the subject listed in that database is protected. Conducting a privacy analysis is a difficult task. At this point in time, no single measure for database privacy is fully satisfying and this matter is still a hot topic in scientific communities. However, extensive research, mainly conducted by statisticians (area of statistical databases, etc.) and computer scientists (both data miners and security experts) is making significant progress.

By using privacy risk assessment techniques, pseudonymisation performance can be guaranteed. Data collection models are used to estimate the risk level for re-identification by attackers (a priori risk assessment). How the data should be separated (identifiers versus payload), filtered (removal of information) and transformed (transforming payload information in order to make it less identifying) is subsequently determined on the basis of these results. This means that in fact that one of the uses of privacy risk assessment techniques is to determine correct configuration of PETs.

Many more aspects of the pseudonymisation process are closely linked and key to ensuring optimum privacy protection, as for example, the location of the identifier and payload processing, the number of steps in which the pseudonymisation is performed, the use of trusted third parties (TTPs). The latter is an important aspect, because use of a trusted third party for performing the pseudonymisation process offers some clear advantages:

1. As the communicating parties (data sources and collectors) not always trust each other, trust can be established indirectly by use of a third, independent party. The parties are then bound by a code of conduct, as specified in a privacy and security policy agreement they agree on with the TTP.
2. Use of a TTP offers the only reliable protection against several types of attacks on the pseudonymisation process.
3. Complementary privacy measures (PETs) and data-processing features can easily be implemented, e.g. controlled reversibility without endangering privacy.

Organisations willing to act as pseudonymisation TTP need to satisfy some important requirements like, e.g. they should be strictly independent; be able to guarantee security and trustworthiness of their methods (openness), software modules, platforms and infrastructure; be able to provide professional expertise related to the domain where the pseudonymisation is being performed; implement monitoring and quality assurance services and programs; etc.

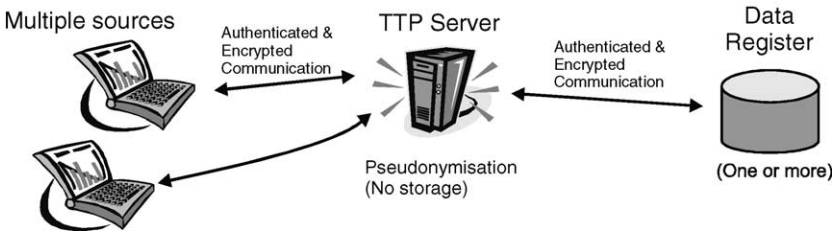

Fig. 1 Communicating entities.

5. Pseudonymisation implementations

The pseudonymisation as described above provides privacy protection for data collection for research and market studies. Two logical entities involved in handling the data are:

1. The data suppliers or 'sources';
2. The data collectors, one or several 'data registers' where the pseudonymised data are stored.

Data suppliers typically have access to nominative data (e.g. treating doctors), the data collectors should only have access to anonymous data. Pseudonymisation schemes requiring the highest level of protection use a trusted third party to deal with this.

Typically there are two different situations in which data is collected. They require different pseudonymisation solutions with different features. Both models and techniques explained below have already been tested and implemented in several different contexts, e.g. in phase 4 clinical trials and for processing drug prescriptions.

5.1. Batch data collection

A first possible scenario is the use of pseudonymisation in batch data collection. The three interacting entities are shown in Fig. 1. In contrast to traditional data collection, the sources (e.g. electronic medical record systems) do not necessarily interact directly with the database and vice versa. Communication is routed through a pseudonymisation server (TTP server), where the pseudonymisation and the processing of relevant data take place, as required.

Data is gathered and packed at the sources, typically in local databases for on-site use. An example could be a local patient database which is managed at a hospital. The data is transmitted on a regular basis to the register through the TTP server where it is pseudonymised.

As described in the previous paragraphs, data extracted from the local databases is split into identities and (screened) payload data according to rules determined during the privacy risk assessment stage. Identifiers are pre-pseudonymised at the source, i.e. a first transformation into pre-pseudo-IDs is performed. The payload data (assessment data) is filtered for indirect identifying data and transformed if necessary, to avoid re-identification of the anonymous data. Finally, the pre-pseudo-IDs are encrypted using a public-key scheme for decryption by the TTP server exclusively. The payload data are public-key encrypted to the register, so that only the register can read the data. Both are then transmitted to the TTP over secure links (authenticated and encrypted).

Full trustworthiness and integrity of the service is thus guaranteed not only by means of policy but also on a technical level. First, because the TTP never actually processes real identities (there is a pre-pseudonymisation stage). Second, because although payload information passes through the TTP server, the latter can neither interpret nor modify the assessment data. This data is encrypted for decryption by the final destination (data register) only.

Still, it must be understood that although the pre-pseudonymised information leaving the source no longer contains any real identities, this does not always guarantee absolute privacy. As the pre-pseudonymisation software is available at many sources (locations), an intruder might find a way to map identities with their corresponding pseudo-identities (a 'dictionary attack') by entering known identities and creating a translation table. Such an attack can be prevented by use of tamper-proof pseudonymisation devices; these are however not yet deployed in real data collection scenarios (see Fig. 2.).

By performing a second transformation in a centrally controlled location, i.e. in the TTP server, optimum security can be offered against such malicious attacks. But as already mentioned, there are more advantages to the use of an intermediary party. As the TTP server dynamically controls the pseudonymisation process, additional privacy-protecting functionality can be added, e.g. monitoring of incoming identities against such attacks,

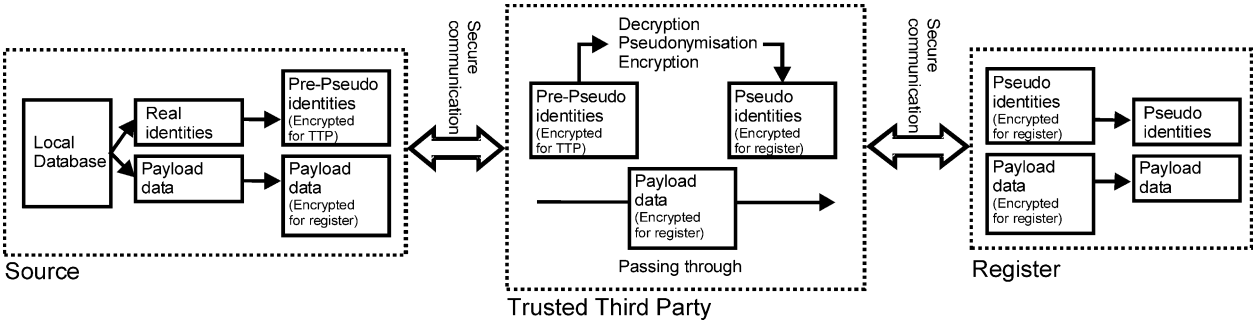

Fig. 2 Dataflow (identity data vs. payload data).

complex mappings of identifies, data flow segmentation, data source anonymisation, etc.

After this second stage at the TTP, in which the pre-pseudonymised identifiers are transformed into the final pseudo-IDs (using cryptographic algorithms), both the payload data and the pseudo-IDs are transferred to the register via secure communication. At the register, the data can then be stored and processed without raising any privacy concerns.

5.2. Interactive data storage

This batch collection model is, however, not always the most ideal solution. In some applications, there is no explicit need for local storage, or the sources are even reluctant to install such a system (for technical reasons). These are application service provider (ASP) like models in which one would rely on the services of a central data warehouse.

Although the data could be extracted from different central data warehouses for pseudonymisation in a batch model fashion, a more elegant solution exists. Besides, such a set-up would require the storage providers to handle sensitive (nominative) data, which can be a problem as such.

Privacy-enhanced interactive data storage allows to integrate the requirements of nominative data access (i.e. like in a local private database) with sharing de-identified data within a single database. This means that all participants in a data collection program can use the central database infrastructure provided by the data collector. These data sources can access their data in a nominative way through a pseudonymisation server, which means that they can truly use the central resource instead of a local application. The data collectors themselves, however, are only able to see an anonymised database. The access to the nominative data is not protected by classical security measures alone, so there is no risk of data leakage, as only de-identified data would be exposed.

The PET technology used ensures that only people having specific personal information on a patient can access the nominative data (typically treating doctors). The privacy-protection engine (PPE) at the TTP creates a transparent gateway from the 'nominative' world to the 'pseudonymous' world, and this is illustrated in Fig.3.

The simplified dataflow described below will clarify the basic concepts behind this scheme. The data at the sources is separated into three types:

1. Nominative identifiers for calculation of a pseudo-ID, e.g. name, place and date of birth of patient, treating doctor, etc.
2. Administrative identifying information, e.g. address, telephone number, etc.
3. Research data, the 'other' data that can be made publicly available for research.

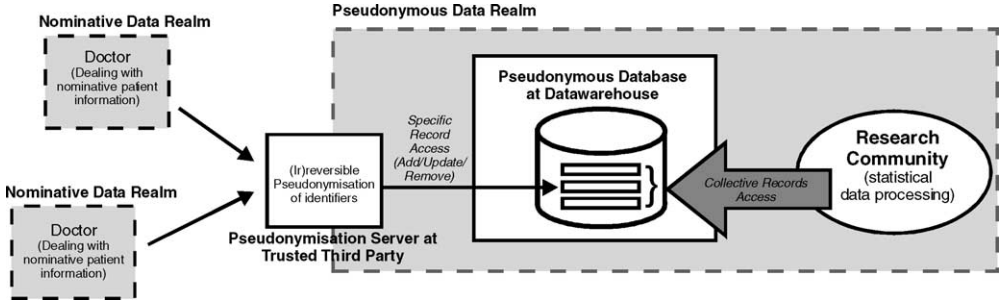

Fig. 3 Interactive privacy-enhanced data storage model.

Source and TTP server (two-step process) translate the nominative identifiers into pseudo-IDs. Note that this process can be irreversible, as someone willing to access a nominative record should always be able to provide the necessary nominative identifiers (e.g. a doctor accessing a patient record). The nominative identifiers are a limited amount of data that could be provided in each database session when dealing with a specific record or be stored locally for convenience. The latter does not undermine the central data storage concept, since its loss presents no problem, and since the (isolated) nominative identifiers contain no privacy-threatening information anyhow.

The administrative-identifying information poses a problem for privacy and storage. A user in a centralised system cannot be expected to store this administrative information on his local computer (for one, it should be accessible from different places). This data should therefore be stored in the central database server, where it should be accessible by authorised persons only. Note that limiting database access (relying on security only) is not an option, as in that case the database is not truly pseudonymised and records can be re-identified when data leaks or the database is ''hacked". A solution to this problem consists in encrypting the administrative data before entering it into the database. When one derives the secret key for encryption partially from the nominative identifiers, this 'secure vault' can only be unlocked by authorised persons (in the nominative data realm).

Finally there is a the 'research data', data which can be listed in plain text in the database and which can be accessed freely by researchers, without any privacy risks.

Basically one should see this configuration as accessing a centralised database, but without the privacy problems associated to centralising nominative data.

The above description is only a brief introduction to the interactive mode of operation of the pseudonymisation service and should certainly not be read as an exhaustive explanation. The interactive mode can range from simple configurations (without storage of administrative data, by use of basic web forms) to the most complex real-time database access schemes with a dataflow as illustrated in the figure.

A final remark is the fact that database interaction is not a one-way path, there is also a dataflow for responses. This is, however, a straightforward process, which is not elaborated in this paragraph.

6. Genomic data

Although the privacy-enhancing techniques mentioned in this paper were not originally designed to handle genetic data, it is expected that they can be used for it, or at least adapted as to suit the needs for protecting genetic data. The specific nature of the genomic data presents however a real challenge for privacy-protection technology and research.

Mainly privacy-gauging techniques can help to estimate risks that can be expected when collecting genetic data. By indexing the requirements for genomic databases, and matching this onto re-identification models, the privacy risks arising from such data collection can be better understood.

There are three important categories of such databases (for sake of clarity, only human data is considered here):

1. DNA databases (e.g. Genbank, EMBL, DDBJ);
2. Databases for medical and pharmacological research;
3. Databases used in medical treatment.

The first type poses only a minor threat. These databases mainly collect short published sequences from many different subjects. Sequences from the same entity cannot be linked and there is no reference to information which could help to re-identify the sample subject. These databases are mainly of interest to biologists and contain data submitted by individual researchers. These databases are quite similar regarding their contents and update one another periodically. However, there is a growing interest of studying longer sequences or sequences from the same subject.

When genomic data for medical and pharmacological research are collected, the re-identification risk increases considerably. In such studies, one benefits from data collection on an individual basis, typically for a longer period. When the data are linked with clinical data (e.g. naturalistic data from electronic health records) the amount of information collected on a single person is significantly increased.

The last type of database is handled by people involved in the direct care, and thus having a confidentiality relationship with the data subjects. The level of identifying information disclosed to persons not directly involved in this process (e.g. lab assistants) should be minimised when possible. This is mainly a security policy issue.

Despite a strong objection, the idea of a national DNA database is gaining support in Iceland and other countries like Estonia and USA [14]. The UK has already embarked on a voluntary BioBank that will

be used to study the interaction of genes, environment and diseases. As BioBank UK has not yet outlined a plan that will adequately protect data, the risk of privacy associated with accumulating more information on the same subject naturally rises. In addition, the database will be made available to pharmaceutical companies to develop new drugs and treatments.

When studying the re-identification methodology that could be used with this type of data collection, one specific issue attracts attention: the link between genotype and phenotype. The genotype is considered to be the information which is stored in an anonymous database. If a record can be re-identified, then sensitive information about a subject can be revealed by deriving phenotype from the re-identified genotype information. On the other hand, when deriving phenotype from anonymous genotype information, it can be used to cross-link with other nominative databases (e.g. diagnostic data) in order to re-identify the anonymous record [15]. The relation between genotype and phenotype is of great importance to the medical community, which means that it is subject to extensive research. One cannot understand the basis of a disease without understanding the function of the related gene, the proteins for which they code and the resulting complex interactions [16].

The increasing collection and use of genomic data represents therefore a great challenge for privacy protection. Solutions to this problem will not only have to consist of privacy policies (non technical means — both modifications of existing ones and completely new ones) and (modified) existing PET technology, but also of new PETs specific for the handling of genetic data. The need for efficient integration of both policies and PETs is, e.g. illustrated in [17].

7. Conclusions

Privacy includes the right of individuals and organisations to determine for themselves when, how and to what extent information about them is communicated to others. Various privacy-enhancing technologies exist that can be used for the correct treatment of sensitive data in medicine. It was shown that advanced pseudonymisation techniques can provide optimal privacy protection of individuals while still allowing the grouping of data collected over different time periods (cf. longitudinal studies) and from different sites (cf. multi-centre studies).

The increasing collection, storage and processing of genomic data, for genetic testing and pharmacogenetics, however, gives rise to new concerns. These new dangers are mainly caused by the very specific nature of genetic data, which is quite different from classical clinical data: DNA contains more information, e.g. about someone's probable medical future, about relatives, etc. Using privacy-gauging techniques, risk can now be analysed and a combination of solutions can be researched.

Two scenarios were illustrated in which human rights in the realm of privacy and optimising research potential and other statistical activities are reconciled. The privacy-enhancing techniques explained are currently deployed for medical research, which proves that the use of pseudonymisation and other innovative privacy-enhancing techniques can unlock valuable data sources, otherwise legally not available.

Acknowledgements

This work was supported by the 6th R&D Framework Programme of the European Commission via the project entitled ''Structuring European Biomedical Informatics & Support Individualized Healthcare'', the InfoBioMed-Network of Excellence (IST 2002 507585). The information presented does not necessarily reflect the policy of the CEU.

References

[1] D. Lazarus, A tough lesson on medical privacy: Pakistani transcriber threatens UCSF over back pay, San Francisco Chronicle Wednesday, October 22, 2003.
[2] Directive 95/46/EC of the European Parliament and the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
[3] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications).
[4] M. Caloyannides, Society cannot Function without Privacy, IEEE Security and Privacy, vol. 1, No. 3, May—June 2003.
[5] K.W. Goodman, Ethics, genomics, and information retrieval, Comput. Biol. Med. 26 (3) (1996) 223—229.
[6] F. Martin-Sanchez, Integrating genomics into health information systems, In: Meth. Inf. Med. 41 (2002) 25—30.
[7] R.S. Fedder, To know or not to know: legal perspectives on genetic privacy and disclosure of an individual's genetic profile, J. Legal Med. 21 (2000) 557—592.
[8] M.J. Mehlman, The effect of genomics on health services management: ethical and legal perspectives, Frontiers of Health Services Management 17 (37) (2001) 17—26.

[9] L.M. McConnell, B.A. Koenig, H.T. Greely, T.A. Raffin, Genetic testing and Alzheimer disease: recommendations of the Stanford program in genomics, ethics, and society, Gen. Testing 3 (1) (1993) 3—12.
[10] D.J. Solove, M. Rotenberg, Information Privacy Law, Aspen Publishers, New York, 2003.
[11] F. De Meyer, B. Claerhout, G.J.E. De Moor, The PRIDEH project: taking up privacy protection services in e-health, in: Proceedings MIC 2002 ''Health Continuum and Data Exchange", IOS Press, 2002, pp. 171—177.
[12] First draft of AURTAF, Anonymity User Requirements for Trusted Anonymisation Facilities. CEN/TC 251/WG III N 02-018 (2002-07-17).
[13] G.J.E. De Moor, B. Claerhout, F. De Meyer, Privacy enhancing techniques: the key to secure communication and management of clinical and genomic data, Meth. Inf. Med. 42 (2003) 148—153.
[14] J. Kaiser, Population database boom, from Iceland to the US, Science 298 (5596) (2002) 1158—1161.
[15] B. Malin, L. Sweeney, Determining the identifiability of DNA database entries, in: Proceedings of the 2000 AMIA Annual Fall Symposium, November 4—8, 2000, Hanley & Belfus, Inc., Los Angeles, CA, 2000, pp. 575—579.
[16] R. Arngrimsson, et al., A genome-wide scan reveals a maternal susceptibility locus for pre-eclampsia on chromosome 2p13, Hum. Mol. Gen. 8 (9) (1999) 1799—1805.
[17] P. Bohannon, M. Jakobsson, S. Srikwan, Cryptographic approaches to privacy in forensic DNA databases, in: Proceedings of Public Key Cryptography, Springer, 2000, pp. 373—390.