# A Linkable Identity Privacy Algorithm for HealthGrid

Ning ZHANG[a,1], Alan RECTOR[a], Iain BUCHAN[a], Qi SHI[b], Depak KALRA[d], Jeremy ROGERS[a], Carole GOBLE[a], Steve WALKER[c], David INGRAM[d], Peter SINGLETON[e]

[a]*School of Computer Science, University of Manchester, UK*
[b]*School of Computing and Math Sciences, Liverpool John Moores University, UK*
[c]*UK Biobank, Manchester Incubator Building, 48 Grafton St, Manchester*
[d]*Centre for Health Informatics and Multiprofessional Education (CHIME), University College London, UK*
[e]*Judge Institute, University of Cambridge, UK*

**Abstract.** The issues of confidentiality and privacy have become increasingly important as Grid technology is being adopted in public sectors such as healthcare. This paper discusses the importance of protecting the confidentiality and privacy of patient health/medical records, and the challenges exhibited in enforcing this protection in a Grid environment. It proposes a novel algorithm to allow traceable/linkable identity privacy in dealing with de-identified medical records. Using the algorithm, de-identified health records associated to the same patient but generated by different healthcare providers are given different pseudonyms. However, these pseudonymised records of the same patient can still be linked by a trusted entity such as the NHS trust or HealthGrid manager. The paper has also recommended a security architecture that integrates the proposed algorithm with other data security measures needed to achieve the desired security and privacy in the HealthGrid context.

**Keywords.** HealthGrid, security analysis, security architecture, privacy and accountability.

## 1. Introduction

A HealthGrid allows the gathering and sharing of many medical, health and clinical records/databanks maintained by disparate hospitals, health organisations, and drug companies. This large-scale sharing of medical records via network connections has the potential to bring us numerous benefits. It would enable real-time and remote access to large quantities of medical and clinical data regardless of the original healthcare setting in which they were acquired, and regardless of where and when the access is performed. This will, in turn, allow us to improve clinical decisions and diagnoses and to provide better patient care. HealthGrid aggregates longitudinal healthcare data

[1] Corresponding Author: School of Computer Science, the University of Manchester, Oxford Road, Manchester, M13 9PL, UK; E-mail: nzhang@cs.man.ac.uk.

giving a more complete history of patients no matter where the care was provided, allowing real-time monitoring of trial results and research outcomes as well as early detection of disease and health problems. HealthGrid would be expected to reduce costs and improve healthcare efficiency. It should also permit a wide range of clinical and bio-science research to be performed more easily, particularly where large populations of patients are being reviewed or studied.

While this HealthGrid vision enables us to provide better healthcare, it also brings security and privacy risks and challenges. Individual medical records contain intimate and sensitive personal details, such as employment, lifestyle, diseases, disabilities, medication and healthcare history, even including family details. Moreover, records can include information about mental health or psychological stability, etc, which patients may wish to limit disclosure. Large-scale collections of medical records from multiple sources managed in disparate administrative domains, but accessed and used by a diverse range of medical professionals and clinical researchers, create enormous risks for the inappropriate disclosure of private or personalised information.

The adverse effects of inappropriate disclosure can be very damaging, and can potentially influence a patient's ability to obtain employment, medical insurance, etc. Failure to protect the confidentiality and privacy of medical records can damage the trust between patients and health professionals causing patients to withhold sensitive information from their care providers in the future. This in turn will affect the quality and safety of health care. It will also affect the clinical data underpinning research, as incomplete or inaccurate data may contaminate the knowledge base for health research outcomes [1]. Therefore, the ability to enforce adequate confidentiality and privacy control in HealthGrids is both an ethical issue affecting patient care and a matter directly affecting the outcome of medical and clinical research.

In addition, protecting patients' privacy is a legal responsibility. The importance of protecting citizens' privacy has been recognised by governments in many countries. For example, the US, Europe, and some of the Asian countries have introduced, or in the process to introduce, laws and regulations to safeguard personalised information for protecting citizens' privacy. In Europe, a cornerstone EU Directive [6] has led to the passing of Data Protection laws in all European countries, such as the UK [7]. The US HIPAA (the Health Insurance Portability and Accountability Act of 1996) healthcare law is another example of such legislation.

This Data Protection legislation enshrines the rights of citizens to control the movement and processing of their personal data, of which health care records are a particularly sensitive example. Health records need to be created and managed with the consent of each data subject (i.e. each patient), and only subsequently used in pursuit of the purposes for which that consent was obtained. Whilst there are specific clauses to permit some disclosure of the data in the vital interests of the patient or of society, the use of identified patient data for research ought to be formally permitted through explicit consent.

In both Europe and the US there is specific Data Protection exemption for data that is anonymous: the data subject cannot be identified from the anonymised data directly and the data cannot be linked to any other available data that identifies the subject. There are many situations where explicit consent cannot be gained for research use of health data: for example, if existing clinical databases are to be mined for novel research questions, it is often not possible to go back to former data subjects to obtain new consent.

The creation of anonymised repositories of longitudinal health data by removing highly identifying data fields from electronic health records is a logical approach to enabling future research, but is in practice very difficult to achieve.

Confidentiality and privacy control in Grid environments is a complex and challenging task. It requires the use of effective legal, administrative as well as technical measures. In this paper, we focus on technical considerations in the protection of patient record confidentiality and identity privacy. The paper discusses security and privacy requirements in this context, and examines countermeasures to satisfy these requirements. In detail, Section 2 gives the terms and notation used in the paper. Section 3 outlines confidentiality and privacy requirements in HealthGrids. Section 4 proposes a novel algorithm for achieving traceable identity privacy. Section 5 gives a security architecture enforcing the identity and data privacy requirements. Finally, Section 6 concludes the paper.

## 2. Notation and Terms

For the sake of clarity, the following summarises the terms, notation, and pretexts used in the remaining part of this paper:

- The term **HealthGrid** is used as a generic name referring to a data repository of medical, health or clinical records generated at multiple points of care. It is linked to administrative and research databanks.
- **De-identification** (sometimes also called *pseudonymisation* or *anonymised* data) refers to the process by which all privacy sensitive information that can be used to identify the real identity of a patient is removed from the patient's records. Examples of such sensitive information are name, address, dates for admission, discharge, birth and death, NHS number, social security number, etc [4].
- A **patient** refers to a subject of care who has a medical record in the HealthGrid.
- A **health record** refers to all the data related to a patient stored in the HealthGrid. It may include genetic data, medical records, samples, consent forms, and some other items. This can be denoted as **HealthRecord** = {**Genetic Data, Medical Record, Samples**, **Consent Forms**, **Other Items**}. Different items of a patient's HealthRecord may be generated or supplied by different care providers. We assume, in this paper, that patients' HealthRecords stored in the HealthGrid have already been de-identified.
- A **care provider** refers to an entity or an organisation that provides healthcare to a subject/patient. Care providers are also the suppliers of HealthRecords (or items of HealthRecords) to the HealthGrid. They are also the users of the HealthGrid. Examples of care providers are GPs, NHS trust hospitals, private healthcare organisations, and pharmacists.
- **Users** of the HealthGrid can be care providers, health and/or medical research professionals/organisations, or drug companies. A care provider can only link a patient's health record or an item of the health record to the real identity of the patient (i.e. patient re-identification) if and only if the care provider is the originator of the record/item. Any other users should not be able to link multiple records of a same patient but generated by different care providers or link a patient's record(s) back to the real identity of the patient.

## 3. Confidentiality and Privacy Requirements

The following discusses key confidentiality and privacy requirements in the context of HealthGrids and the challenges in satisfying these requirements.

### De-identification (or Pseudonymisation)

In order to protect a patient's privacy, his/her health record needs to be de-identified before being passed outside the premises of the care provider [3] or outside a dedicated regional or national healthcare network, and used for research or other purposes. This is a complex challenge for many reasons:

1. Health data often contains rich and quite personal descriptions of an individual's health or social circumstances, which are not always recorded in sections that can be predicted and removed;
2 Some patterns of health and health care are unique or easily recognised as relating to an individual, such as unusual family histories or the management of rare conditions;
3 Some kinds of personal data are in themselves both strongly identifying and of too great a research importance to be removed: date of birth, occupation, x-ray images and genomic data are examples of this;
4 There is a need to join up contributions to a longitudinal anonymised record from multiple providing healthcare sites, over time: this cannot be done if there is no common identifier to link new contributions to an existing anonymised record.

This paper focuses on item 4 in this list. Other research projects such as CLEF [8] are exploring the other challenges listed.

Assuming that patient records can be de-identified (i.e. personally disclosive information is removed), we need a pseudonym to identify the patient and his/her de-identified record for continued healthcare. One suggestion is to use a universal patient identifier (UPI) [3].

Merely using a UPI for identifying a patient and his/her record(s) has security and privacy weaknesses. Firstly, this solution is weak in coping with patient mobility. Nowadays, medical personnel are more specialised, and patients are increasingly mobile. For reasons such as seeking specialist help or private healthcare, a single patient may have several health records generated by multiple care providers. Using a single identifier to index these multiple records allows easy linkage of the records, and this direct linkage is the weakest point for privacy attacks. Any compromise of a patient's UPI would expose all the medical details of this patient. In addition, the fact of using a single UPI for one's multiple medical records may actually help perpetrators to crack the patient's identity.

In Section 4 of this paper, we propose an alternative approach to patient identification, which uses a cryptography-based method for identity privacy, record linkage, and identity re-identification without compromise privacy.

### HealthGrid user authentication and authorisation

Enforcing controlled access and operations on health records in HealthGrid is rather complex due to the diversified nature of Grid users. Strictly speaking, patient records should only be used for *intended healthcare purposes*. However, for various obvious reasons, many other organisational or individual entities, such as medical support

personnel, medical researchers, insurance companies, etc., have certain legitimate needs to access certain information from a patient's health record [2]. Conditions under which these entities are permitted to access *the needed information* in a timely manner so that their professional responsibilities can be fulfilled while at the same time security and privacy are not compromised are complex and should be better understood. The general access policies that actually determine *who should be allowed to retrieve what* from patient records need to be defined by health services, usually at a national level, and supplemented at a fine grained level by policies defined by individual patients and healthcare enterprises.

The authentication and authorisation mechanisms used should address the above-mentioned complexity. The Fame-Permis project[2] is currently in the process of developing an integrated authentication and authorisation framework to cope with Grid authentication and authorisation needs [5]. Different authentication tokens used with different authentication protocols provide different levels of assurance (LoA) in identifying a Grid user. The project is trying to link a user's access privileges to the authentication token used by the user along with other attributes such as his/her role, time of access, etc.

**HealthGrid user accountability and non-repudiation**

User authentication and authorisation mechanisms should be linked to an auditing facility so that information such as *who has accessed (or supplied) which record* and at *what time* is recorded and enjoys integrity protection. In this way, neither a data consumer nor a data supplier is able to later falsely deny that a specific action on the HealthGrid has taken place. This security requirement not only ensures data usage accountability, but also is essential for resource management and service charge if necessary.

**Patient re-identification**

Once medical records are collected, pseudonymised, aggregated and analysed/studied by researchers, there is sometimes a need to feed the research outcome back to the patients involved for continued healthcare, for which the identity of a patient needs to be re-identified. This process can be repetitive and last throughout the lifetime of a patient.

**Patient consents**

A health Grid system should be able to enforce patients' consents (or a patient's individual privacy preference such as *I do not want entity x to see my data item j*) in its access control facility and do so in an automatic manner.

## 4. A Linkable Identity Privacy Algorithm (LIPA)

### *4.1. The Key Challenge*

As discussed in Section 3 above, from time to time, fresh de-identified health records may need to be uploaded into the HealthGrid by care providers, and patient re-

[2] http://www.fame-permis.ac.uk.

identification can only be achieved by the original record supplier, or record generator. Simply using a UPI, or a NHS number, for indexing a patient's HealthRecord in the HealthGrid is not sufficient to achieve such linkable identity privacy. Therefore, we need a viable solution that can address the following question. *How to assign a unique pseudonym to a de-identified patient record provided by a specific care provider such that the de-identified records associated to the same patient but supplied by different care providers (therefore with different pseudonyms) can still be linked together in the HealthGrid?* In other words, a patient will have different pseudonyms with different care providers, so by merely looking at these pseudonyms, one (a HealthGrid user or a care provider) can not associate together all the records of the same patient supplied by different care providers. However, the HealthGrid should be able to link all the records of the same patient provided by different care providers without knowing the real identity of the patient. LIPA is such an algorithm designed to accomplish this objective. This is what we call the *linkable* (or *traceable*) *identity privacy* algorithm.

*4.2. The Algorithm*

Before presenting our proposed solution, we first outline the structure of identifying patients or patients' records used in our model. In the UK, in addition to NHS Numbers (NHSNo), there may be other identifiers (OtherID) used by pharmaceutical companies or research labs to index their respective patients or trials. Every person in this country, who is not necessarily a patient, has an NHS number. We here introduce another identifier, HealthGridID, to index each patient record or each set of records associated to the same patient, including the patient's highly sensitive record items such as genetic data, in the HealthGrid.

Figure 1 below gives a diagrammatic view of these identifiers in relation to the data items.

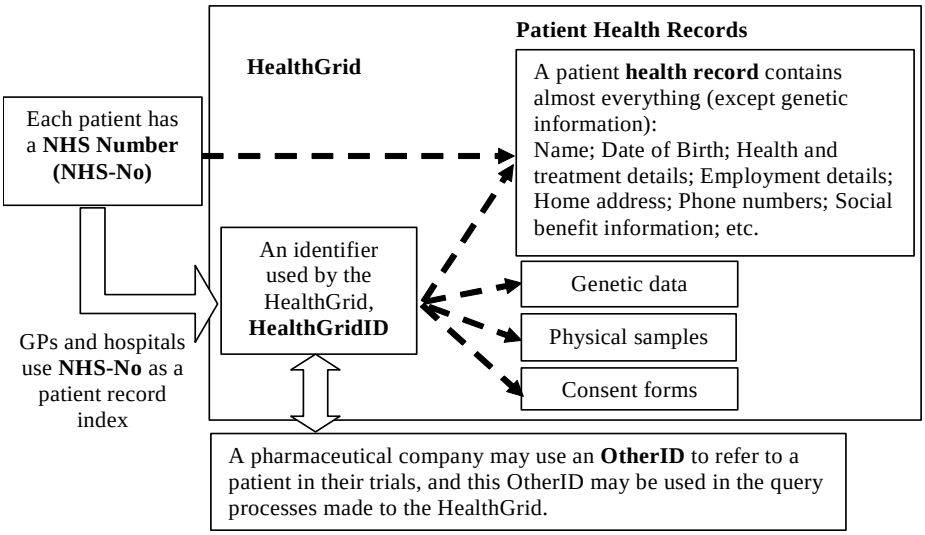


**Figure 1.** Identification structure

As mentioned earlier, preserving identity privacy is usually achieved through the use of a pseudonym for each patient. The simplest and most obvious solution is to re-encode a patient's NHS Number into a fixed pseudonym that is used throughout different practices and hospitals. In this way, the data record of a patient can easily be identified, traced and upgraded by different care providers. This approach, though straightforward in achieving the traceable property, does not provide an adequate level of identity privacy and identity-to-data unlinkability. This is because a care provider could easily link different medical conditions to the same patient by making repeated queries using the patient's pseudonym.

A better solution is to give different pseudonyms of the same patient to different care providers. However, a question here is how to facilitate the linkage of a patient's records provided by different care providers. For example, when two care providers, Hospital_A and Hospital_B, both enter data related to patient *Alice* into the HealthGrid using two different pseudonyms, how could the HealthGrid know that these two pseudonyms actually point to the same patient *Alice* and make sure that two sets of input are associated in the databank?

To enable this data linkage, we need to map different pseudonyms of the same patient to his/her HealthGridID before any data query processing. There are two possible approaches to such mapping. One is to use a mapping table that is securely protected by the HealthGrid. This mapping table is potentially a security and performance bottleneck. Should this table be compromised, the security of the entire system will be compromised. Moreover, the size of the database at the HealthGrid may grow as time passes, and if the mapping table needs to be searched for each query, then the performance of the database access will be severely affected. This means that the mapping table is not a cost-effective approach.

The second approach is to use a cryptographic algorithm to establish the mapping from the different pseudonyms of a single patient to his/her HealthGridID. This method can avoid the weaknesses of the first approach mentioned above. Our LIPA solution has been designed using the second approach to achieve the following objectives:

- *Identity privacy*: Without permission, it is computationally difficult for any user of the HealthGrid (except for the original record supplier and the pseudonym issuer) to be able to link a patient's data/record/samples stored in the HealthGrid database to the real identity of the patient.
- *Unlinkable pseudonym*: Multiple pseudonyms of the same patient can not be linked without collusion between care providers.
- *Patient mobility*: When a patient changes a care provider, his/her data record in the HealthGrid can still be linked and upgraded without compromising the real identity of the patient.

As shown in Figure 2, LIPA consists of two functional modules: the *NHSNo*-to-*Pseudonym* (N2P) Conversion Module and the *Pseudonym*-to-*HealthGridID* (P2H) Conversion Module.

#### *4.2.1. NHSNo-to-Pseudonym (N2P) Conversion Module*

The N2P Conversion module can either be run by a trustworthy HealthGrid manager or by an independent Trusted Third Party (TTP) such as the NHS Trust. Its role is to re-encode a patient's NHS number into a pseudonym that is the function of the following

parameters: the patient's HealthGridID, a random number (*Rand*), a hash value of a master secret (*MasterSecret* known only by the HealthGrid manager or the TTP), timestamp (when the pseudonym is issued), and the care provider's name and address. Other factors such as the pseudonym validity period can also be taken as part of the hash input if other conditions are imposed on the use of the pseudonym. The conversion process assumes that the encoding party (i.e. the HealthGrid manager or the TTP) maintains a mapping table of *NHSNo* vs. *HealthGridID*. More formally, a pseudonym $Pseudonym_{ij}$ for patient $i$ looked after by care provider $j$ ($CP_j$) can be expressed as the following:

$$Pseudonym_{ij} = f(HealthGridID_i, Rand_{ij}, H(MasterSecret||CPname_j||CPaddress_j||Timestamp_{ij})) \quad (1)$$

Here, $f(w)$ is the conversion function defined in Figure 3. $H(z)$ is a secure hash function, such as SHA-1 [9], with the following properties: (a) for any $z$, it is easy to compute $H(z)$; (b) given $z$, it is hard to find $z'$ ($\neq z$) such that $H(z) = H(z')$; and (c) given $H(z)$, it is hard to compute $z$. "$y||z$" is the concatenation of data items, $x$ and $y$.

It is worth noting that we can get rid of the *NHSNo* vs. *HealthGridID* mapping table, and instead use another secret for the translation between NHSNo and HealthGridID. The reason for keeping this mapping table in this model is that some existing medical/health databanks already have this table in place.

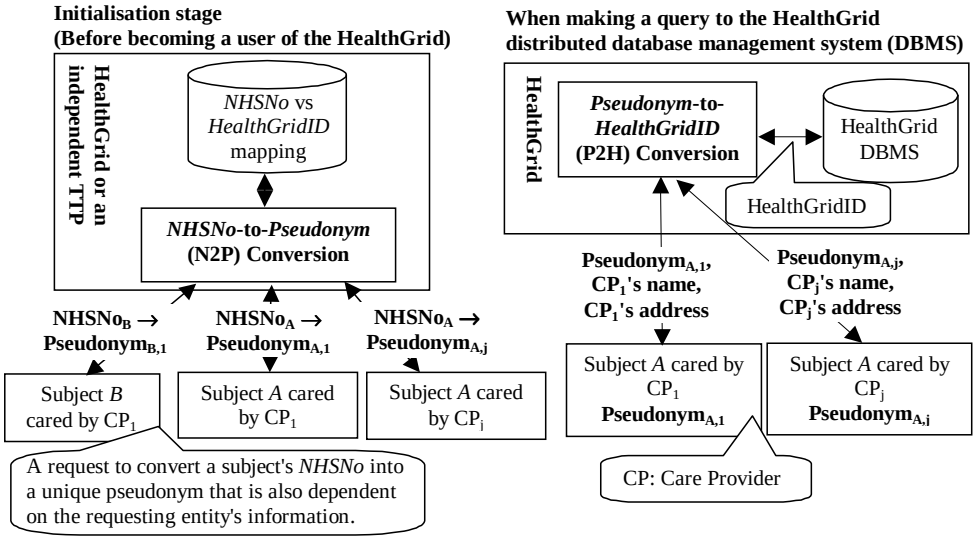


**Figure 2.** LIPA modules

Function (1) shows that the pseudonym is not only dependent on the patient's HealthGridID (linked to NHSNo), but also the information related to the care provider with which the patient is being cared for, and a random number associated to the patient as well as the care provider. The use of care provider related information as input parameters to the function ensures that the same patient cared by different care providers will have different pseudonyms. In this way, the linkage between a patient's pseudo identity and his/her HealthGridID (or NHSNo) is effectively broken. Additionally, a secret number, the *MasterSecret*, is used in the generation of

pseudonyms to enhance the security of the whole process. Any entity without the possession of this MasterSecret will not be able to generate a valid pseudonym for a patient. The possibility of pseudonym clashes is further reduced by the use of the random number, *Rand*, which should be discarded without disclosure after the pseudonym is generated.

*4.2.2. Pseudonym-to-HealthGridID (P2H) Conversion Module*

When a care provider accesses, or uploads a patient's data into the HealthGrid database, he/she needs to use the pseudonym of the patient together with his/her (i.e. the care provider's) name and address. The query is first directed to the P2H Conversion Module that converts the pseudonym back to the patient's unique HealthGridID, and then the query is indexed by the HealthGridID and sent to the Database Management System (DBMS). In other words, to update a patient's record in the HealthGrid, a care provider must possess a valid pseudonym for the patient.

Upon the receipt of a query request, the P2H Conversion Module calculates $HealthGridID_i$ using the pseudonym $Pseudonym_{ij}$ submitted and the requester's name and address, together with the secret, *MasterSecret*, that is only known to the HealthGrid or the TTP, as shown in Function (2) below.

$$HealthGridID_i = f^1(Pseudonym_{ij}, H(MasterSecret||CPname_j||CPaddress_j||Timestamp_{ij})) \quad (2)$$

Once the $HealthGridID_i$ is recovered, the P2H module uses it to index the data items of the patient in the HealthGrid repository.

Figure 3 defines the LIPA algorithm described above. To summarise, our proposed solution has the following merits and properties:

Let:

$$y(x) = (a + b * x) \bmod n$$

where $n$ should be a large prime number, and mod is the modulo operator.

We can set:

$a = HealthGridID_i$, (Patient $i$'s identifier in HealthGrid)
$b = Rand_{ij}$, (a random number dependent on patient $i$ and $CP_j$)
$Timestamp_{ij}$ = this pseudonym's issuing date & time
$h_1 = H(MasterScret||CPname_j||CPaddress_j||Timestamp_{ij})$
$h_2 = H(MasterScret||CPname_j||CPaddress_j+1||Timestamp_{ij})$

**N2P Conversion:**
Given $HealthGridID$, $Rand$, and $h_1$ and $h_2$, the N2P Conversion Module computes $y(h_1)$ and $y(h_2)$, and defines $\{y(h_1)||y(h_2)||Timestamp\}$ as a pseudonym for the subject with the $HealthGridID$.

**P2H Conversion:**
Given $\{y(h_1)||y(h_2)||Timestamp\}$ as well as the care provider's $CPname_{ij}$ and $CPaddress_{ij}$, the P2H Conversion Module computes $h_1$ and $h_2$, and then works out $HealthGridID_i$ by solving the following equations:

$$\begin{cases} (HealthGridID_i + Rand_{ij} * h_1) \bmod n = y(h_1) \\ (HealthGridID_i + Rand_{ij} * h_2) \bmod n = y(h_2) \end{cases}$$

**Figure 3.** Linkable Identity Privacy Algorithm (LIPA)

- The same patient will have different pseudonyms with different care providers.

- Different pseudonyms of the same patient can be recovered back to his/her unique and secret HealthGridID that is then used to index all the data record(s) related to the patient in the HealthGrid repository.
- The patient's HealthGrid internal identifier, i.e. HealthGridID, is secured through the use of a modular arithmetic function and a master secret.
- The most innovative part of the LIPA solution is that the method can link multiple pseudonyms of the same patient back to his/her unique HealthGridID through the use of this master secret, instead of using a mapping table that, we believe, creates performance and security bottlenecks in the system. Furthermore, the only item that the HealthGrid needs to memorise or securely store is the master secret, so almost no extra storage is required for the implementation of the proposed method.

## 5. A Security Architecture for HealthGrid

This section explains how the LIPA solution may be integrated into the HealthGrid to achieve identity and data privacy. We first address the issue of data privacy. There are two data level privacy requirements related to the HealthGrid:

DP1 *HealthGrid data level privacy*: Assume that a patient's data record in the HealthGrid database consists of four data items/objects listed in the order of descending sensitivity:

**HealthRecord for patient x** = {**Genetic-Data$_x$, Medical-Record$_x$, Samples$_x$, Other-Items**}.

Access control is performed based upon different data sensitivity levels. In other words, different data items will require different access privileges that are mapped to the roles of an access entity. For example, Genetic-Data$_x$, which is the most sensitive, can only be accessed by an entity with a role such as a "genetic specialist". On the other hand, this subject's Medical-Record$_x$ may be accessible to a wider user base such as those with the role "Nurses".

DP2 *Individual privacy preference*: Data items with the same level of sensitivity form a single superset that may be accessed collectively. For example, a query with the role of "Consultants for Genetic Disease" may request for Genetic-Data of 100 patients. This operation should be further subject to individual privacy preferences. For example, a patient who failed to sign a consent form should not be included in this 100 patient superset.

To fulfil these two data privacy requirements, we need two additional modules, *Data Filter* and *Patient Consent Manager*, to work together with those shown in Figure 3. This results in the security architecture illustrated in Figure 4.

The Data Filter module enforces privacy requirement DP1, making sure only an authorised entity with a specific role could access to a specific data item(s) associated to a patient record. Once the entity is securely identified and authenticated, the P2H module translates the pseudonym into the subject's HealthGridID, and then the Data Filter module takes over. The Data Filter module enforces *what operation(s) on which data item(s)* this entity is allowed to perform on the data record indexed by the HealthGridID with respect to the privacy policy detailed in the HealthGrid Data Privacy Policy Repository. If the queried items match with the policy, then the query

processing proceeds. Otherwise, the Data Filter deletes the unauthorised items from the query before passing it to the Data Query Processing module.

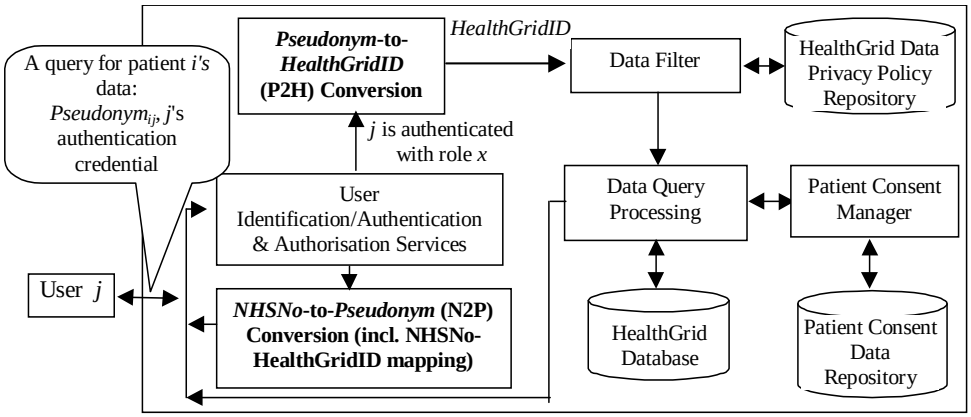


**Figure 4.** A Security Architecture with LIPA

The Patient Consent Manager module, together with the Patient Consent Data Repository, fulfils data privacy requirement DP2. There is a real need for data items with the same privacy sensitive level to be batch processed or accessed. The Patient Consent Manager ensures that all the data accesses are in conformance to individual patients' privacy preferences detailed in the consent forms.

The User Identification/Authentication & Authorisation Services module is responsible for identifying/verifying a claimed user and granting varying levels of access privileges to different user groups at the service level. For example, the HealthGrid may only open the NHSNo-to-Pseudonym conversion service to data suppliers, not to data consumers.

## 6. Conclusions

In this paper, we have analysed confidentiality and privacy requirements in the context of HealthGrids. A novel algorithm for achieving traceable identity privacy has been proposed, and a security architecture embracing the traceable identity privacy algorithm is recommended. More work is needed in order to address all the privacy requirements identified in the paper, and the effects of these countermeasures on operational requirements need to be investigated. We will pursue these tasks in our future research.

## 7. Acknowledgements

We gratefully acknowledge the constructive comments given by the anonymous referees.

**References:**


[1] Cushman, R., "Information and Medical Ethics: Protecting Patient Privacy", IEEE Technology and Society Magazine, Fall 1996, pp 36-39.

[2] Ting, T. C., "Privacy and Confidentiality in Healthcare Delivery Information System", Proceedings of 12th IEEE Symposium on Computer-Based Medical Systems, 1999, pp. 2-4.

[3] Tyler, J. L. "The Healthcare Information Technology Context: a Framework for Viewing Legal Aspects of Telemedicine and Teleradiology, Proceedings of the 34th Annual Hawaii International Conference on System Sciences, 2001, pp. 1-10.

[4] Vice, J., Privacy – De-Identification of Protected Health Information, http://irb.chw.org/Policies/HIPAA%20-%20De-Identification%20of%20Protected%20Health%20Information.pdf, accessed 14 Nov 2004.

[5] Zhang, N., Chin, J., Rector, A., Goble, C., and Yao, L., "Towards an Authentication Middleware to Support Ubiquitous Web Access", the Proceedings of the 28th Annual International Computer Software and Application Conference (COMPSAC 2004), Volume: 2, IEEE Computer Society, Hong Kong, September 28 – 30, 2004, pp. 36-38.

[6] Directive 95/46/EC of the European Parliament and of the Council of Europe 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, Official Journal of the European Communities, Number L281/31, 23 November 1995.

[7] Data Protection Act 1998, The Stationery Office Limited, London; 1998; ISBN 0 10 542998 8.

[8] Kalra D, Singleton P, Ingram D, Milan J, MacKay J, Detmer D, and Rector A, "Security and confidentiality approach for the Clinical E-Science Framework (CLEF)", Methods of Information in Medicine (in press).

[9] Schneier, B., Applied Cryptography, John Wiley & Sons, 1996.