A Conceptual Framework for Information Privacy and Security in Collaborative Environments

Geoff Skinner, and Elizabeth Chang,

Curtin University of Technology, Perth, WA, AUSTRALIA

Summary
There are numerous information privacy approaches based on the four major models of privacy protection. That is, Comprehensive Privacy Laws, Sectoral Privacy Laws, Privacy Self-Regulation, and Technologies of Privacy. These solutions, used individually or without proper system privacy design considerations, have not been very effective. This is because there has been little in the way of instruction on how developers and designers are supposed to use these privacy tools. In this paper we address the problem by providing a privacy solution for integration into information systems called Shield Privacy. The Shield Privacy solution provides an effective system wide approach to privacy protection. It integrates relevant components from the various privacy models. We have implemented our Shield Privacy in a collaborative environment application. In this paper we also describe the prototype and discuss its advantages and areas of future work.

Key words:
Shield Privacy, Information Privacy, Data Security, Hippocratic Policies, Personally Identifiable Information (PII).

Introduction

It seems that where ever you go on the Internet today every body wants to know your name. This is usually along with a host of other personal details [1]. It's a scenario that paints a bleak future for information privacy. This is due in large part to the fact that numerous services are being moved online. These services are collecting vast amounts of personal information. The need for excessive and increasing collection habits is cause for concern. This practice needs to be questioned and stopped as it represents serious threats to personal privacy. Most of the time entities are not given a reasonable spectrum of choices for what information you provide in order to use the services. It is normally a case of fill in all of the required form fields, or do not use the service at all. When you have no choice but to use the service you are placed in an uncompromising position. It is a situation where personal privacy is the added and often hidden cost.

Information Systems need to have better privacy protection procedures and mechanisms integrated into their design and implementation. Numerous new laws and regulations are continuously being introduced that are increasingly restricting personal information collection and use. This is in addition to consumer outcry which is also becoming louder and drawing more attention [2]. It is causing organizations to question and review their personal information collection and handling practices. However, organizations are finding that they do not have the proper tools to allow them to correctly manage and enforce privacy [3]. Better privacy protection tools are required to manage personal information and determine what personal information really needs to or can be collected. Most importantly the methodologies and guidelines for implementing and integrating them into information systems are required. This is because system developers and operators have had little guidance on how to implement and comply with privacy guidelines and rules [4]. Further, there have been few analytical or systematic attempts to understand the relationship between privacy and technology [5]. Therefore, Information Systems need a comprehensive systems wide approach to information privacy.

In this paper we show our solution for a holistic information privacy design and implementation. We have termed our solution the Shield Privacy. This is due to its symbolism in providing a number of protection layers for personal data privacy and personally identifiable information (PII). Shield Privacy contributes a total privacy protection methodology during system design, development and implementation. Its straightforward design makes it easy for developers and operators to integrate it into Information System implementations. It is founded on incorporating the relevant principles from each of the four models of privacy protection [6]. Shield Privacy utilizes our approaches to technologies of privacy, allows for configuration for compliance with comprehensive and sectoral laws, and provides a transparent system privacy design to allow for self-regulation. This is in addition to the ability for the attainment and maintenance of privacy certification and seals.

The structure of the rest of the paper is as follows. Related research areas and existing solutions are discussed in Section 2. This section details current solutions and proposals that are similar to our own work or of direct relevance or influence. The details of our proposal and innovative ideas are presented in Section 3. This section also provides the general architecture for Shield Privacy.

Manuscript received February 25, 2006.
Manuscript revised February 28, 2006.

The realization conditions and assumptions we have used for our privacy protecting environment are outlined in Section 4. A brief summary is provided in Section 5. This is followed by references used throughout the paper.

2. Related Work and Background Material

Current privacy laws are too concerned with data protection rather than protection of the person. Sectoral Laws have proved incomplete and inadequate. Self-regulation has not provided the stringent privacy protection required and is often negatively effected by organizational financial constraints. Many of the technologies of privacy have fallen short in their protection of privacy and often only address specific issues, rather than providing complete system wide privacy protection. The solutions we are concerned with and are proposing deal with Information Privacy. From an attempted definition of a particular dimension of privacy one can loosely categorize the solutions aimed at each of them. Privacy in general is very subjective and means different things to different people. Common among all interpretations is the perspective that it is a human right but is context and environmentally dependent. Roger Clarke has outlined a number of common privacy dimensions that have gained wide acceptance [11]. They are as follows:

- Privacy of the person
- Privacy of personal behavior
- Privacy of personal communications
- Privacy personal data

Personal data, also referred to as information privacy is the focus of this paper. Clarke provides a well referenced definition of information privacy after initially stating it is being a combination of personal communication privacy and personal data privacy. His formal definition of information privacy is "… the interest an individual has in controlling, or at least significantly influencing, the handling of data about themselves." [11]. The Common Criteria (CC) [12] provides a more formal requirements based definition for providing "… user protection against discovery and misuse of identity by other users.". As you can see from the CC definition it is information systems requirements focused, with emphasis on identity protection. Identity protection is a major component of information privacy but by no means represents the complete embodiment of its full meaning.

Solutions addressing information privacy issues are the most closely related to our own work and therefore of the most interest. Our solution of Shield Privacy is predominately an information privacy management and design tool. A similar line of work being is proposed by Borking in [10]. It is based around the concept of Identity Protectors. The authors approach is to question the amount of personal data that really needs to be collected in information systems. Once collected they divide a 'privacy-protected' system into two separate domains and use the identity protector to convert a user's actual identity into a pseudo-identity. The idea is to minimize the 'identity-domain', and maximize the 'pseudo-domain' for increased privacy protection. Their view of privacy-protection systems of the future would have the identity protector take the form of a user controlled smart card. The card seems to be initially 'separate' from the Information System and is primarily used to generate pseudo-identities. For anonymous transactions it is mentioned that the identity protector can be integrated into the information system.

Other related work and inspiration for our research on Hippocratic Privacy Policies for Information Systems [7, 8] and Shield Privacy was Hippocratic Databases [9]. This work was itself inspired by the Hippocratic Oath guiding physicians. It involves the designing and developing of databases to include responsibility for the privacy of data as a fundamental tenet. The privacy principles the database is responsible for are built upon the 'foundation' principles found in most current privacy legislations and guidelines. They normally have been derived from the Fair Information Practices (FIPs) [13] and the OECD Guidelines for Governing the Protection of Privacy and Transborder Data Flows of Personal Data [14]. Our own research recognized that not all personal information is restricted to database storage. Personal information flows through numerous processors and components of the information system. Our Hippocratic Privacy Policies address this issue and provide privacy guidelines for the whole Information System. The specific policies are detailed in Section 3.3.

3. Framework Components

Shield Privacy is made up of a number of unique privacy protecting system design and implementation 'rules'. These rules represent our own interpretations and approaches for better privacy protection. Essentially Shield Privacy aims at providing a holistic approach and solution to privacy protection in information systems. We feel that privacy is a design principle in its own right, made up of various rules and principles. System design guided by these rules and principles should produce an information system that goes a long way in protecting personal data and PII. Through the transparent privacy design philosophy administrators, designers, operators, and most importantly users can view and understand the privacy protection methodologies. It is felt that by providing a clear understanding of the privacy protection mechanisms it helps instill trust and confidence in the system. This is turn will hopefully result in greater

acceptance and usage of shield privacy, leading to integration into all information systems in one form or another.

From an architectural perspective Shield Privacy is founded upon four key elements, or design principles. Each component in its own right is not very complex or hard to implement. This makes Shield Privacy easy to integrate into a system. Most of the elements are based on well established IT Security and Privacy concepts, some even being referred to as foundational principles. It is our approach and application of each component that provides the unique contribution to information privacy. In addition, when applied together in a well designed, planned and complete manner it provides a distinctive solution to a broad spectrum of information privacy issues found in current information systems. The main elements that form the layers of shield privacy are the following:

• PDM-ADM Design and Implementation Rule: Our approach to personal data minimization and anonymous data maximization. Used for determining what personal information is really required. Once it is collected making it anonymous where ever possible for its use throughout the information system.

• SDD Design and Implementation Rule: Our approach to the separation of duty and data within the information system. Segregating system roles and data based on sensitivity, context of use, and owner assigned privacy use permissions for each data element stored.

• HPP Design and Implementation Rule: derived from our previous work on Hippocratic Privacy Policies for Information Systems [7, 8]. The research was inspired by the work on Hippocratic Databases [9]. It basically means that the database, or in our case the Information System, must take responsibility for the privacy of the data it manages.

The four elements are discussed in more detail in the following Sub-Sections. The final Sub-Section provides a holistic summary for how all four are combined in an information system during planning, designing, developing and implementing phases of a new system.

3.1 Personal Data Minimization – Anonymous Data Maximization

On the surface the idea of personal data minimization seems relatively easy and straightforward. Organizations and system owners should simply stop collecting so much personal information. However, the concept is much more complicated than this, and involves trying to change the mindset of organizations and system designers developed over the last few decades. It is widely perceived that the collection of as much personal information as possible is a profitable activity. Organizations feel that there are financial gains to be made by selling information to interested third parties. Additionally, from a marketing perspective the last decade has seen a dramatic increase in personalized services. However, Alan Westin has pointed out that customers are '… increasingly fed up with targeted marketing campaigns …' [2]. More importantly however is that many news laws and regulations are making these personal data practices illegal. The idea of collecting vast amounts of personal information from users with no details on how and by whom it is going to be used is opening the door for costly and damaging lawsuits from disgruntled users.

Organizations need to review their data collection practices and examine how their systems, and their partners, handle personal information. There are greater costs involved in trying to make systems privacy compliant after they are built. Additionally, potentially more costly are the legal complications arising from breaking privacy laws either intentionally or unintentionally. Privacy issues need evaluation from system inception. We feel, along with others working in this area [1, 6, 9, 10], that the current trend for collecting personal information needs to reversed. We have termed our approach Personal Data Minimization (PDM). PDM is one half of a personal information design and implementation rule. PDM can be used by system designers and developers to help provide better privacy protecting system configurations. It basic steps are outlined as follows:

• Analyze new system and system processors personal data requirements: this means that designers, developers and system owners need to determine what personal information is really needed. This should be done in such a way that the absolute minimum amount of information is collected.

• Follow personal information flows through the system: Once the decision has been made on the data to collect, the system needs to be analyzed to discover where the information goes. This helps associate usage and retention properties to the personal data elements collected.

• Determine contextual and subjective privacy sensitivity levels to each personal data element. By making serious attempts at this stage of the system design to determine the sensitivity of the data, it will make the task of data separation easier at later stages.

• Most importantly where the is still a perceived need to collect more personal information all form fields should be made 'opt-in' rather than 'opt-out' for users and data uses.

The second half of the PDM-ADM rule is termed Anonymous Data Maximization. Once we have determined the minimum amount of personal information we can collect we can now isolate the cases where the information can be used in an anonymous way. Borking [10] has identified that the only processes that really need to know a user's identity is Authorization and Accounting. Therefore, for most other system processes, the users personal data, and hence identity, can be anonymized. This means that for processes such as Identification and Authentication, Access Control, and Auditing no personal information is really required. For greater clarification, some parts of the personal information may be used, but in this context the identifying component has been made anonymous. As a result the privacy sensitivity of the personal information has been greatly reduced. So like PDM, each system process should be examined to determine if it can maintain its functionality using anonymous data. If so, then all personal data should be anonymized for those processors (Anonymous Data Maximization). See the diagram above (Fig. 1) for a graphical representation and summary of PDM-ADM Privacy Design and Implementation Rule.

3.2 Separation of Data and Duty

Separation of Duty and Data (SDD) is another two facet privacy design and implementation rule. The first half is concerned with the Separation of Duty. In the past the separation of duties has been considered valuable in deterring fraud and a number of other security benefits. What has become apparent more recently is its usefulness to the field of privacy. Separation of duty requires that for particular sets of transactions, no single individual be allowed to execute all transactions within the set. Its extension to privacy is that for particular sets of personal data and transactions dealing with this data, no single individual will be allowed to access all of the personal data or transactions within the set. Separation of duty can be either static or dynamic. Both static and dynamic approaches of separation of duty are based on some form of Role Based Access Control (RBAC). Static is the assignment of individuals to roles and allocation of transactions to roles. Dynamic provides a richer set of possible policies by controlling the activation and use of roles [15, 16].

For our purposes we use a type of dynamic approach, called the History-based Separation of Duty defined by Simon and Zurko in [16]. It basically means that ' … no role member is allowed to perform all the activities in a business task on the same target or collection of targets.' [16]. It is again a role based approach to access control which is beneficial due to the fact that many user functions can be conveniently separated by role. Allocating access rights according to role is also helpful in defining separation of duty in a way that can be enforced by an information system. The role division and allocation of rights in our approach are influenced by the personal information each would come in contact with during system operation. The greater the privacy sensitivity of the data the more divisional separation of roles, more restrictive access controls, and more limited allocation of rights.

Privacy and separation of duty are both subjective and context related. Therefore separation of duty is determined by conditions external to the system. For an effective application of separation of duty for privacy protection it also needs to be considered at the inception of a new information system. Having determined the processors that take place in the system then designers need to determine the separate duties and roles that have access to each of them. This relates back to the personal information each process might handle, and which roles should therefore have access to those personal data elements.

The separation of data works in conjunction with separation of duty. To provide a number of extra layers of privacy protection personal data, along with normal operating data should be classified into different classes. Each class would be categorized with a privacy sensitivity rating that is used to determine certain access privileges for the separation of duty role based access controls. From a physical implementation perspective it could mean that personal information is stored in a totally separate databases and locations. Which means it could be subject to not only tighter access controls, but also subject to better security protection mechanisms. Such as enforcing encryption of all personal data at rest with very strong protocols, auditing and recording of all access to the personal information database, and ensuring all information passing to and from the database is in encrypted form.

3.3 Hippocratic Information System Policies

Like the Hippocratic Database principles that govern the design and implementation of the databases that should be used in information systems protected by shield privacy, the rest of the system should be designed through the guidance of the Hippocratic policies [7]. There are seven in total and are the following:

1. Anonymity: Whenever and wherever possible the personal information collected and stored in the information system should be done in a way that supports anonymity for the individual user.

2. Limited Collection and Use: The personal information collected must be the minimum necessary for the primary purpose specified to the individual providing that information. Once collected the information system will only use that information for the primary purpose specified to the individual that providing the information. In both cases the individual must give explicit consent for the primary collection purpose and use.

3. Limited Disclosure and Retention: The information system may not disclose the personal information other than for the primary purpose of collection or keep it for a period longer than the primary purpose requires without the individuals explicit consent.

4. Security and Sensitive Information: The information system must take all reasonable steps to protect the personal information it holds from misuse and loss and from unauthorized access, modification or disclosure. Sensitive information that has been collected with the explicit consent of the individual must always be protected by 'stronger' security safeguards. It is recommended that all sensitive information at rest and in transit within, being transferred to and from the Information System be encrypted with suitable strong protocols.

5. Openness, Access, and Integrity: The information system must have documented and make easily available its policies and procedures for the management of personal information. The information system must also make all personal information about an individual available to that individual and allow them where possible to make corrections to ensure the information is always correct and up-to-date. If the individual is unable to make the corrections then the information system must take all reasonable steps to ensure the integrity of the personal information it stores.

6. Third Party and Transborder Uses: The information system may not transfer information to a third party or foreign country without the consent of the individual. Where the user is not aware of the privacy policies in the target information system and/or foreign country then the information system is responsible for ensuring the destination information system and/or country follows a set of privacy principles ATLEAST as restrictive as its own.

7. Identifiers: The information system must not adopt as its own identifier of an individual an identifier of the individual that has been assigned by external agency or government body. Further, an identifier assigned by an external agency or government body that is stored by the information system but not as an identifier should never be disclosed.

3.4 Information Security for Privacy Protection

The fourth major pillar of shield privacy is concerned with again meeting one of the privacy principles requirements set out in most privacy policies. It is in regards to the organizations responsibility for the safe and secure handling and storage of user's personal data. This means that in all privacy protecting information systems strong information security methodologies and technologies should be applied for the protection of personal information, and the information system in general. Again this requirement is subject to many influences including economic and resource constraints, the organizational security environment, and the rapidly evolving technologies of information security. Therefore this design consideration will vary determining on the organizations that are integrating shield privacy into their information systems. Therefore in this paper we do not go in to detail of all the latest IT Security products and technologies that could and should be used. Rather we provide a baseline list of what needs to be protected, and includes the following:

• The use of Strong Encryption for Personal Data at Rest.

• The use of secure and encrypted communication lines for all transmission of personal data both internal and external to the system.

• Issuing of public-private key pairs unique to all users. This comes from our perceived future need and improvement of protecting personal information from system administrators. It is envisaged that users would be able to encrypt where ever feasible their personal information with their public key to ensure that without their express permission and interaction, no one would have access to certain elements of their personal information. Alternatively the key could also be given to a trusted third party for appropriate use when required.

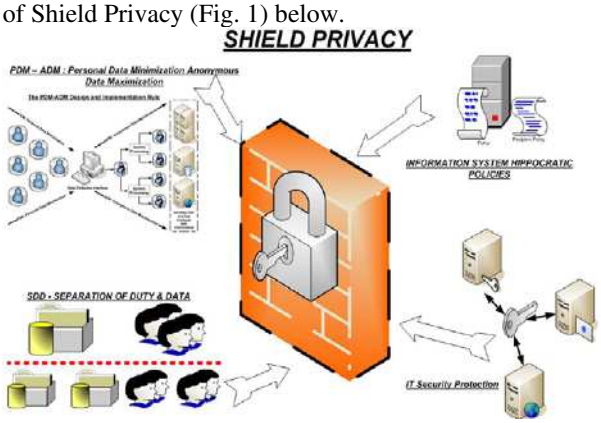
• The use of anonymous re-mailers, firewalls, IP address hiding and other technologies to increase the level of identity protection available to system users. form.

3.5 Shield Privacy

Each of the privacy components listed in the previous sub-sections provide another layer of privacy protection in their own right. Combined, there are then four key privacy designs and implementation 'rules' that make up Shield Privacy. They are the following:
• PDM-ADM: Personal Data Minimization – Anonymous Data Maximization.
• SDD: Separation of Duty and Data.

• HPP: Hippocratic Privacy Policies for Information Systems.
• IT Security for Privacy Protection.

However, it is through their concerted application that they significantly increase information privacy protection to a holistic level. It provides the necessary privacy tools organizations need to build privacy protecting information systems. Shield Privacy is aimed at making information system designers and developers to consider information privacy in its own right as a design objective. It forces information system owners to be aware of privacy requirements in their systems, from both a legal standpoint and as a way of building customer trust and satisfaction. Privacy is now seen as a product differentiator and can provide organizations with a competitive advantage. Shield Privacy provides a simple to use and integrate tool to achieve this objective. A visual representation of Shield Privacy is provided in the diagram of Shield Privacy (Fig. 1) below.



4. Framework Implementation

The application that we have developed for proof of concept is built on the notion of an online anonymous enabled health consultation. The idea is that individuals are able to join the system through a secure registration process. Once registered the users are able to log in to and interact with the collaborative health environment. Their identity interaction may be performed in a number of ways that include:

• As an anonymous user.

• As a pseudonymous user (a number of pseudonymous variants would be available that include transaction, role, and relationship pseudonyms).

• As their real identity or verifiable alias user.

These user accounts represent the first of three different system roles. The ideal conditions would provide for more than the three roles. For testing conditions and evaluation only the three roles are required without any loss of functionality or validity. These roles are the following:

• Information Consumers: these roles represent the users and their respective interactive sub-roles outlined above.

• Information Providers: these roles represent the health professionals that provide the advice to the information consumers.

• System Administrators and Operators: these roles represent the individuals or groups assigned the roles of managing and maintaining the system.

Actually in an ideal setting the roles would scale easily to the size and complexity of the application or organization in question. In our prototype we could further divide the roles of Information Providers depending on their health professional qualifications. For example Medical Practitioners or 'Doctors' could be assigned a role, Psychologists another role, 'alternate' medicine providers another role, and so forth.

In a commercial setting some form of financial infra-structure would be overlaid and integrated into the application. For example, users may be required to pay a small registration or usage fee when joining the Health Consultation Collaboration. Additionally, the Information Providers may also be charged a small fee to become a certified member of the collaboration to offer their services. Alternatively depending on the level of success of the application and how far information consumers are willing to reveal their health issues, a small referral fee may be charged. This is for the cases where an information consumer consults an information provider in the physical world face-to-face for health treatment. In our test setting we have not placed payment processing features in the application. However, we are still collecting the minimum amount of personal information required to process such payments and collections if necessary. This ensures that the same personal information is collected by the system and therefore protected by Shield Privacy. This is required to ensure the validity of the prototype.

Due to the size of the ethical and legal setup work and costs involved with offering health advice in this manner we were unable to use certified health professional at the time of initial testing and evaluation. As a result both the roles of the information consumers and providers had to be simulated by our own people. It is planned to incorporate 'real' consumers and providers as soon as possible. We are currently investigating the similarities to our application and the numerous health advice 'columns' and 'self-help' advice found online and in other media forms, such as magazines. They are also of interest for a privacy review as they always offer a number of

disclaimers and privacy statements with little validation to confirm they are actually enforced. Examples of such sites include: 'Ask The Doctor' [17], NHS Direct Online [18] and 'Health Advice for Students' [19].

5. Conclusion

The X Shield privacy provides a holistic solution to many of the privacy protection issues faced in information systems. It incorporates the four major models of privacy protection including comprehensive and sectoral laws, self-regulation, and technologies of privacy. Shield Privacy is built upon four unique design and implementation strategies that are distinct in terms of our approach to their application and integration into Information Systems. The four key components that deliver shield privacy are the following:

• PDM-ADM: Personal Data Minimization – Anonymous Data Maximization.

• SDD: Separation of Duty and Data.

• HPP: Hippocratic Privacy Policies for Information Systems.

• IT Security for Privacy Protection.

Shield Privacy will be of benefit to information systems owners, designers, developers and operators. Most importantly however is the benefit to users of the information system. Due to the transparent design and implementation nature of Shield Privacy users will be able to see and understand what is happening to their personal information. This addresses the major issues consumers have with current personal information collection practices. Shield Privacy allows users to see why, how, who and for what purposes their personal information is being collected. Additionally they have the ongoing ability to view and update not only their personal information but the privacy protection policies they have agreed to and stored alongside their personal information.

References

[1] Schwartz, P.M. (1999), Privacy and Democracy in Cyberspace. 52VAND. L. REV. 1609, 1610-11 (1999).

[2] Westin, A. and Scalet, S. (2003), Balancing Act. http://www.cio.com/archive/061503/

[3] Powers, C.S., Ashley, P., and Schunter, M. (2002), Privacy Promises, Access Control, and Privacy Management. Third International Symposium on Electronic Commerce (ISEC'02), October 18 - 19, 2002 , Research Triangle Park, North Carolina.

[4] Patrick, A.S. and Kenny, S. (2003), From Privacy Legislation to Interface Design: Implementing Information Privacy in Human-Computer Interactions. Privacy Enhancing Technologies Workshop (PET2003), Dresden, Germany, March, 2003.

[5] Palen, L. and Dourish, P. (2003), Unpacking "Privacy" for a Networked World. CHI 2003, April 5-10, 2003, Ft. Lauderdale, Florida, USA.

[6] Privacy and Human Rights 2003 – An International Survey of Privacy Laws and Developments. Electronic Privacy Information Centre and Privacy International.

[7] Skinner, G. and Chang. E (2004), Hippocratic Policies in Computer Based Collaborations. PHCRC 2004, Newcastle Australia, 2004.

[8] Skinner, G. and Chang. E (2004), Shield Privacy Hippocratic Security Method for Virtual Community. IECON2004, The 30th Annual Conference of the IEEE Industrial Electronics Society, Nov 2-6. 2004 Korea.

[9] R. Agrawal, J. Kiernan, R. Srikant, and Y. Xu; Hippocratic Databases. Proceedings of the 28th VLDB Conference, Hong Kong, China, 2002.

[10] Hes, R. and Borking, J. (2000), Privacy-Enhancing Technologies: The path to anonymity. Registratiekamer, The Hague, August 2000.

[11] Clarke, R. (1999), Introduction to Dataveillance and Information Privacy, and Definitions and Terms. http://www.anu.edu.au/people/Roger.Clarke/DV/Intro.html

[12] Common Criteria (2004), Common Criteria for Information Technology Evaluation. January 2004, http:// www.commoncriteria.org.

[13] Federal Trade Commission (FTC) (2003), Fair Information Practise Principles. Federal Trade Commission Online Privacy, http://www.ftc.gov/reports/privacy3/fairinfo.htm

[14] Organization for Economic Co-operation and Development. OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. http://www.oecd.org.

[15] Nash, M.J. and Poland, K.R. (1990), Some Conundrums Concerning Separation of Duty. IEEE Symposium on Research in Security, 7-9 May, Oakland, California, USA, 1990.

[16] Simon, T.S, and Zurko, M.E. (1997), Separation of Duty in Role-Based Environments. IEEE Computer Security Foundations Workshop, pages 183-194, 1997.

[17] Ask The Doctor, http://www.onlineambulance.com/AsktheDoctor.htm

[18] NHS Direct Online, http://www.nhsdirect.nhs.uk/

[19] Health Advice for Students, http://www.studenthealth.co.uk/