



This item was submitted to Loughborough's Institutional Repository (<https://dspace.lboro.ac.uk/>) by the author and is made available under the following Creative Commons Licence conditions.

 **creative commons**
C O M M O N S D E E D

Attribution-NonCommercial-NoDerivs 2.5

You are free:

- to copy, distribute, display, and perform the work

Under the following conditions:

 **Attribution.** You must attribute the work in the manner specified by the author or licensor.

 **Noncommercial.** You may not use this work for commercial purposes.

 **No Derivative Works.** You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work.
- Any of these conditions can be waived if you get permission from the copyright holder.

Your fair use and other rights are in no way affected by the above.

This is a human-readable summary of the [Legal Code \(the full license\)](#).

[Disclaimer](#) 

For the full text of this licence, please go to:
<http://creativecommons.org/licenses/by-nc-nd/2.5/>

Grid-enabled Mammographic Auditing and Training System

M.H. Yap & A.G. Gale

Applied Vision Research Centre, Loughborough University, Garendon Wing, Holywell Way,
Loughborough, LE11 3TU, UK

ABSTRACT

Effective use of new technologies to support healthcare initiatives is important and current research is moving towards implementing secure grid-enabled healthcare provision. In the UK, a large-scale collaborative research project (GIMI: Generic Infrastructures for Medical Informatics), which is concerned with the development of a secure IT infrastructure to support very widespread medical research across the country, is underway. In the UK, there are some 109 breast screening centers and a growing number of individuals (circa 650) nationally performing approximately 1.5 million screening examinations per year. At the same, there is a serious, and ongoing, national workforce issue in screening which has seen a loss of consultant mammographers and a growth in specially trained technologists and other non-radiologists. Thus there is a need to offer effective and efficient mammographic training so as to maintain high levels of screening skills. Consequently, a grid based system has been proposed which has the benefit of offering very large volumes of training cases that the mammographers can access anytime and anywhere. A database, spread geographically across three university systems, of screening cases is used as a test set of known cases. The GIMI mammography training system first audits these cases to ensure that they are appropriately described and annotated. Subsequently, the cases are utilized for training in a grid-based system which has been developed. This paper briefly reviews the background to the project and then details the ongoing research. In conclusion, we discuss the contributions, limitations, and future plans of such a grid based approach.

Keywords: grid-enabled, mammography, breast cancer, auditing, training.

1. INTRODUCTION

Screening for breast cancer is carried out nationally in the UK for every woman aged between 50 and 70 years where such women are invited for screening every three years [1]. The UK has over 100 screening centres distributed nationally, together with additional mobile screening vans so that each woman can easily attend for her regular screening, as and when she needs to. Screening involves having both medio-lateral and cranio-caudal views of each breast taken. These are then examined and are typically dual read by breast screening radiologists and/or specially trained radiographers (known in the UK as 'advanced practitioners') - both are termed 'mammographers' here.

Where cancer is suspected then other imaging procedures (e.g. ultrasound) or other clinical investigations are subsequently undertaken. The role of the mammographers is pivotal then in first identifying suspicious mammographic appearance and so the skill of these personnel is crucial to the overall success of the national scheme. The most recent data from the NHS Breast Screening programme (NHSBSP) which is for the period 2005/2006 indicate that 2,382,122 women were invited for screening and some 1,891,408 were actually screened (including here additional numbers of women referred by their GPs for screening). Of this, some 87,469 were recalled for assessment of which 14,841 cancers were identified. The aim of the NHSBSP is to reduce the mortality caused by breast cancer, and the latest data show that it is now saving some 1,400 lives every year [1].

In recent years there has been a workforce issue within UK breast screening in that there has been a shortage of consultant radiologists in the specialist domain; however this has been offset by a growth in specially trained radiographers as part of the changing landscape for the radiographer profession within the UK. Currently there are some 109 screening centers regularly performing over 1.5 million screening examinations per year and a growing number of individuals (circa 650) nationally who report on these screening cases [1]. Currently in the UK, screening is performed using mammographic film and several trials of digital mammography have been undertaken. It is highly likely that nationwide full field digital mammography screening will occur within a few years. Several research projects are underway in the UK examining the utility of digital screening and the associated use of CAD, building on research

findings from other countries. Elsewhere within the UK health system radiological imaging is predominantly digital with PACS systems very prevalent. The large high resolution digital images required for breast imaging examination require associated high resolution workstations for their examination, coupled with robust hardware and software to enable these images to be viewed, when required, at other centres besides the originating screening centre.

The grid provides perfect platform for the wide-scale and secure deployment of different types of medical images. The purpose of the present paper is to describe the development of a mammographic training system based on using a grid system and digital mammography images. In recent years there has been an increase in interest in Intelligent Tutoring systems (ITS) in the area of medical training, both in mammography as in other areas [2, 3].

Grid Computing provides the technology opportunities that led to the possibility of using distributed computers as a single, unified computing resource [4]. The EU-funded MammoGrid project [5] and the University of Pennsylvania - based NDMA project [6] are the key relevant projects here that employed grid computing. The National Health Service in the UK constitutes a virtual organization, where each hospital trust is legally responsible for the data held at its sites, and each trust determines who can access its data [7]. The Grid enables a seamless, integrated computational and collaborative environment. However, the security and bandwidth are two major issues that have hindered the practicality of the grids. Hence, GIMI is a research project that aims to provide a secure network with better data transfer rate, and demonstrate this concept via several different applications. The current paper describes the use of the grid in an application utilising geographically distributed medical images.

2. BACKGROUND

The UK is in a good position to support the development of novel technical solutions to support both research into, and clinical implementation of, individualised patient-specific healthcare as it has a National Health Service which provides free medical care for patients. As elsewhere in the world, IT is playing an increasingly important role in healthcare delivery in the UK and the facility for different hospital trusts to share patient images and data in a secure manner is vitally important as it helps to enable cost-effective provision of healthcare for patients. Thus it is important that the IT infrastructure is capable of interfacing with deployed systems (both now and in the future) within the NHS in both a secure and an ethical fashion. Widespread deployment of PACS systems in the UK enables images and other patient data to be shared as required.

GIMI is a large-scale collaborative research project, funded by the Department for Trade and Industry, which is concerned with the development of a secure IT infrastructure to support medical research and clinical practice. The GIMI consortium involves Oxford and Loughborough Universities, as well as University College London (UCL). IBM, e-San, and Siemens are industrial partners and the National Cancer Research Institute is also involved. The long-term goal of the research project is to allow data aggregation from a large number of different resources so as to provide an archive which can be used for commercial and academic research, as well as to support clinical trials. Additionally, the work will facilitate healthcare monitoring and patient self-management, facilitate the development of training algorithms and also serve as a training resource.

The GIMI project is developing software to allow the secure and ethical aggregation of data from distributed remote sources to support a number of different applications. Some applications involve large complex images such as digital mammography and others involve large volumes of small data quantities - such as in diabetes monitoring. The underlying infrastructure of GIMI builds on previous related research projects such as e-DiaMoND and NeuroGrid. This paper presents only the mammographic training aspect of the GIMI project.

3. MOTIVATION

3.1 GIMI Middleware: Service-oriented interoperability framework (sif)

The fundamental aspect of the GIMI project is suitable middleware [8] which provides a service oriented interoperability framework (sif) enabling existing and legacy data to support healthcare delivery, research, and training. Sif can be considered as a conduit which permits a number of applications to communicate with a disparate collection of geographically distributed data sources [9]. Such sif middleware is essentially a collection of web services and sif

provides an API (application programming interface) for application developers. The current mammographic work uses this sif middleware to build a suitable auditing and training application.

3.2 PERFORMS

The PERFORMS (PERsonal PerFORmance in Mammographic Screening) self-assessment scheme has run annually in the UK for some 19 years [10]. It allows an individual mammographer to self-assess their skills in interpreting recent difficult screening cases and judge their performance confidentially against that of all the other mammographers in the UK. The scheme is educational in nature and functions as part of the quality assurance of the NHSBSP. Fundamentally it allows some measure of an individual's skills in identifying key mammographic features. An individual can then opt to participate in further training if they wish. Additionally, individuals gain feedback on how many cases containing early signs of difficult cancers they correctly identify, as measured against known pathology, and also in identifying cases which radiologically appear as if they may contain cancer - as judged by a panel of experienced breast screening radiologists [11].

4. METHODOLOGY AND IMPLEMENTATION

A software development platform has been constructed utilising Eclipse which includes an integration of Flash and Java (figure 1). For a grid-based training system to be useful a key factor is that the images need to be recalled and displayed very rapidly. For this purpose a fast rendering DICOM image viewer was selected [12], and modifications are made to fit into the needs of the project.

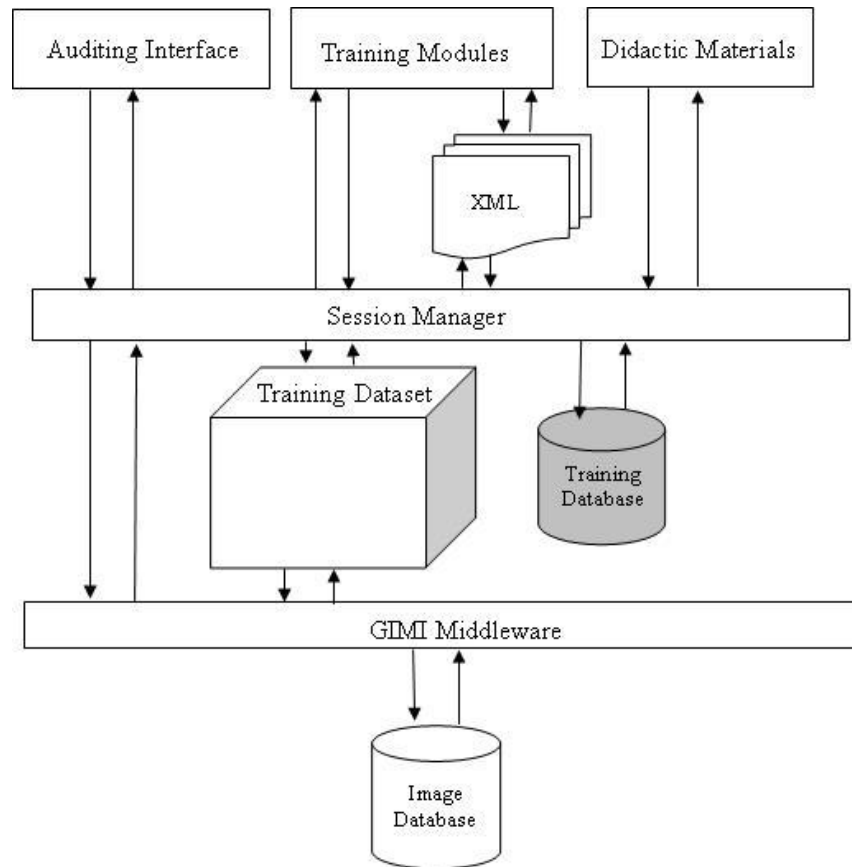


Figure 1. Components and interoperability in the Mammographic Training System.

4.1 Source Image Database

In a previous UK grid project (e-DiaMoND) a database of some 800 anonymised screening mammography cases was constructed [7]. Each case comprises at least four mammographic images; medio-lateral views (MLO) and cranio-caudal views (CC) of each breast. These images originated as mammographic film images from three hospital sites, each contributing a different number of images and the resultant images having differences in image annotations by the radiologists, and were subsequently digitized and added to the database. Images are in DICOM format [12] and each image size is some 33Mb. All cases had been reviewed by an experienced screening radiologist who both annotated and described the key mammographic features present. In the current GIMI research this initial source image database has been split up (as a read-only database) and spatially distributed across secure servers which are based in three universities: Oxford, Loughborough, and UCL. Two important fields in the database, which specify the difficulty of the MLO and CC views are used in the current training system design as guidelines for case selection for different training modules for different skill levels of users.

4.2 Servers Selection

Whilst in the final training system the user will not be aware of where particular cases are based, in the initial auditing of cases for selection in training it has been found useful to be able to identify a particular server. The user can establish a single connection to a server, or to more servers so as to enable federated queries.



Figure 2. Server Connection user interface

4.3 Case Auditing System

Prior to potentially using particular cases for training purposes it is important to be able to access and review individual cases to determine both their usefulness as well as what other information is included. Consequently, the software firstly permits any individual case to be recalled from the source database for examination based upon a patient identifier or upon specific case information. This means that the radiologist or researcher can audit images in the source database and check them for suitability for inclusion as training images in GIMI. If suitable then an identifier is saved into a training database (figure 3). New mammographic images and descriptors can further be added into this training database as needed. Initially the training database has been located at Loughborough University on a local database. Subsequently this will be migrated to the GIMI servers. Besides single server queries, case recall can comprise federated queries across the servers, via the GIMI API. An example of federated query:

```
<QueryRequest>
  <Union>
    <Query>
      <sql>SQL Statement</sql>
      <dbldent>dbldent 1</dbldent>
      <server>server 1</server>
    </Query>
    <Query>
      <sql>SQL Statement</sql>
      <dbldent>dbldent 2</dbldent>
      <server>server 2</server>
    </Query>
  </Union>
</QueryRequest>
```

For the patient-based auditing (figure 3) the case is recalled from the database and both the visual characteristics as well as the known details of the woman are displayed. This includes information about age, relevant medical history and potential hormone replacement therapy (HRT) usage – whilst all cases remain anonymous. The facility to add new cases in the future or to edit details of existing cases exists with such information being stored in the training database. Figure 4 shows the interface used for auditing cases based on displaying key image descriptors, rather than patient factors. The anonymised mammograms themselves are displayed (figure 5) using DICOM viewing software [12].

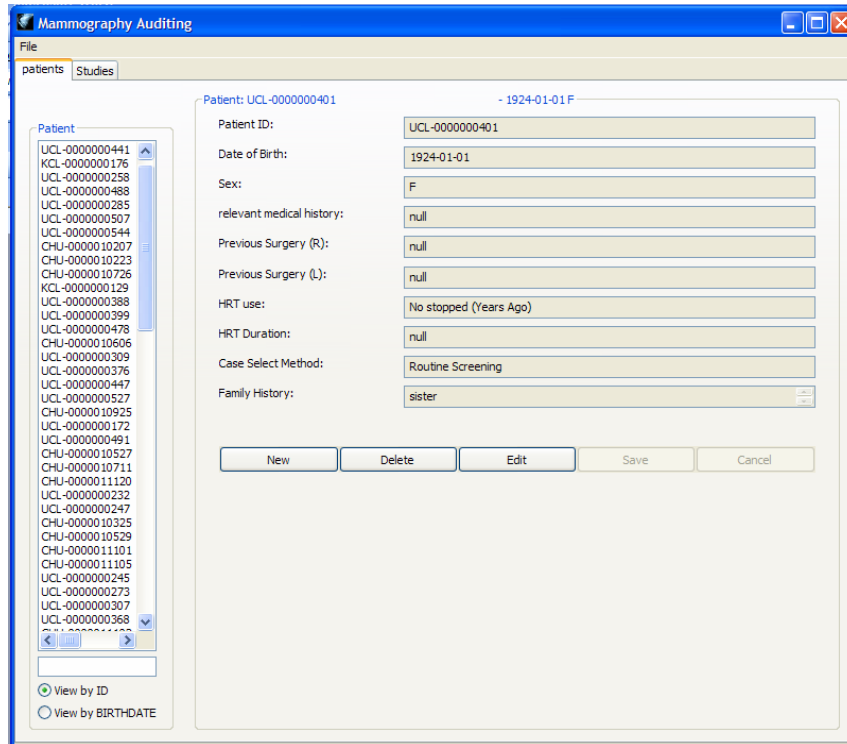


Figure 3. Layout of User interface for examining patient information

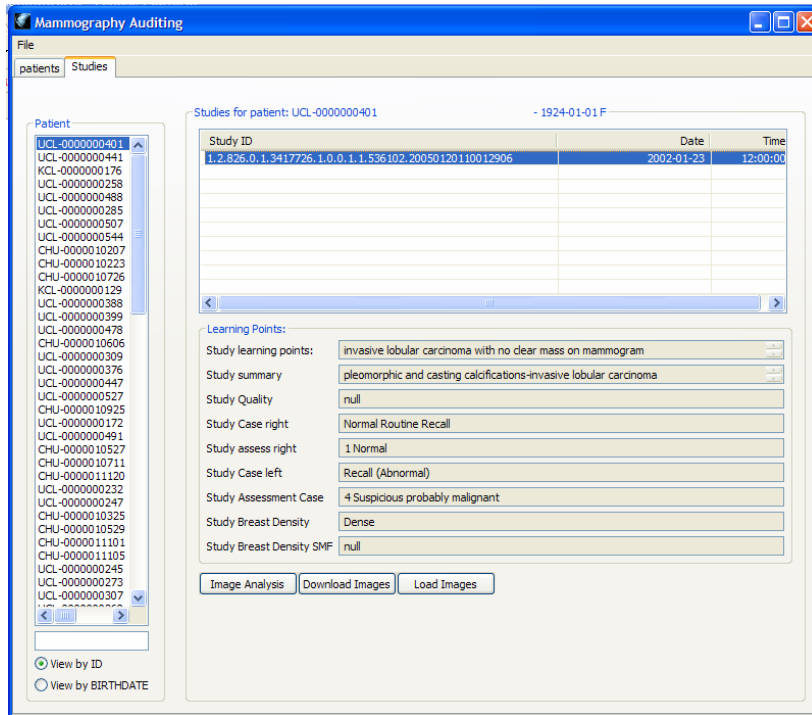


Figure 4. Layout of user interface for examining the study information.

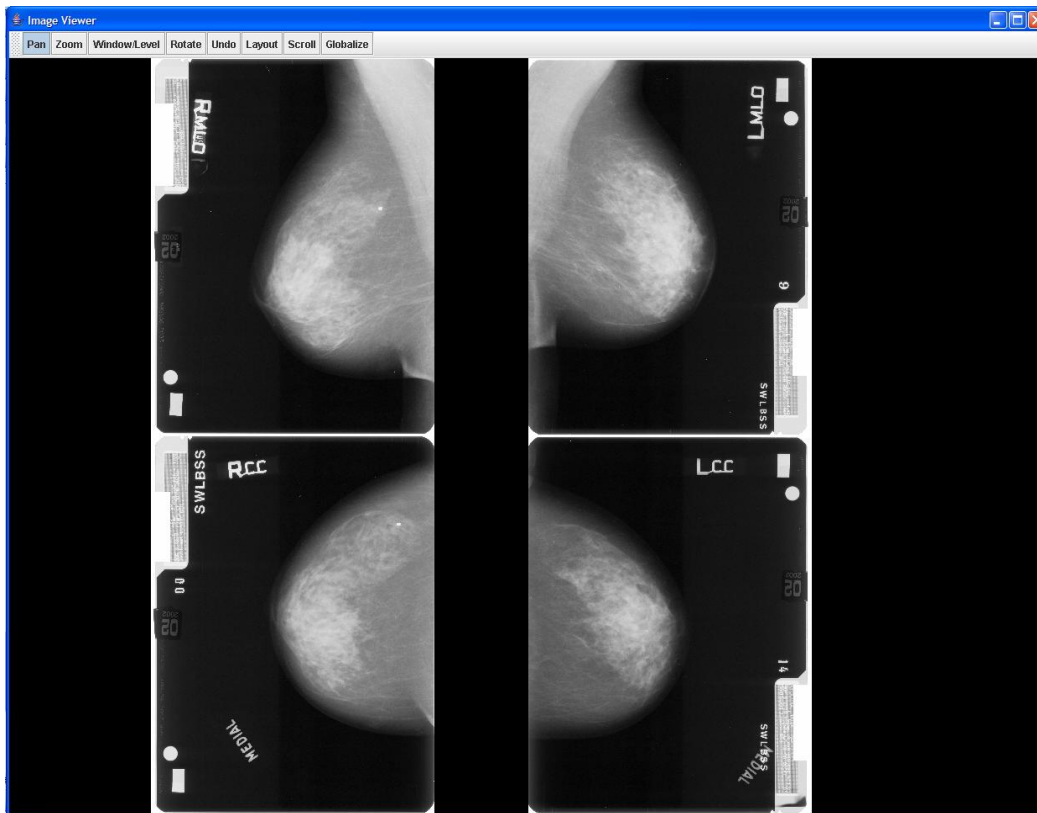


Figure 5. Image Viewer (Courtesy of UCLA freeware - jViewbox)

4.4 Mammography Training System

The training system which is currently under development has integrated the relevant software strengths of Java for networking and communication, and Flash for presentation of information to the user. Additionally, XML plays an important role in the system as it works as the bridge to pass information between the Java and Flash applications.

The initial simplified schematic of the training system is shown in figure 6. The user is identified as a new trainee or if not then their profile is checked from the training database. A profile can be generated by the user traversing through the training system and/or by other assessment criteria (such as their confidential data from the PERFORMS scheme if they wish to use that).

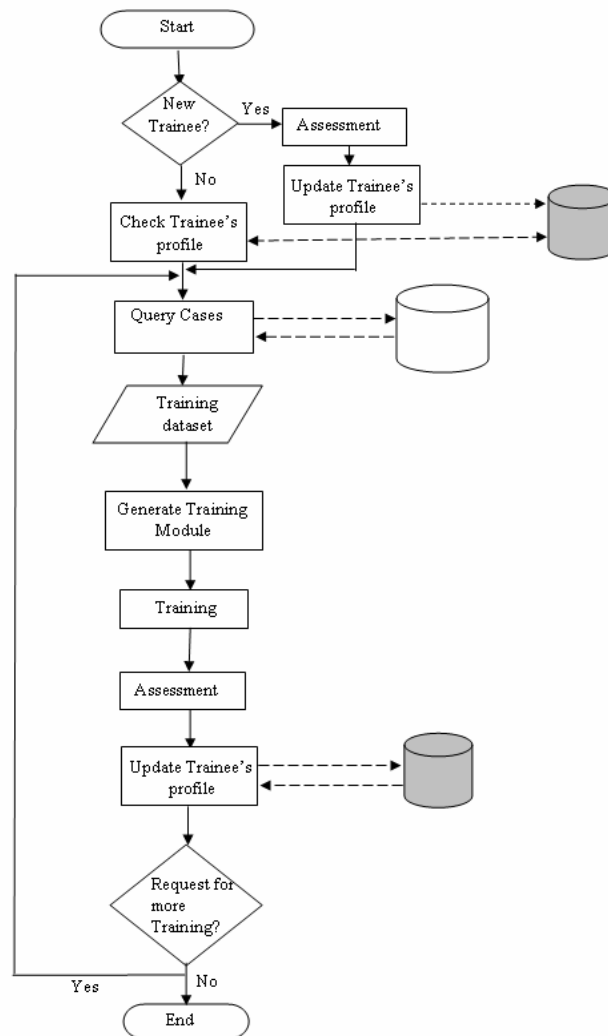


Figure 6. Illustration of the initial training system concept

The four important components in the training system are:

- *Assessment*: An initial test is provided to the non-registered trainee so that he/she can assess his/her own performance on known screening cases and gain immediate feedback. After the test is complete, performance statistics are presented, and the appropriate training module will then be assigned to the trainee based on their performance score.

- *Trainee Profile*: For registered trainees, performance statistics will feed into the trainee's biography, in this case, their profile is stored in the training database. This database is updated to keep track of the trainee's performance.
- *Case Selection for Training Module*: Current research work is investigating automatic case selection for generating specific training modules. For instance, one idea is to select the cases and generate the appropriate test set of cases automatically based upon trainee skills and the known case level of difficulty.
- *Didactic Materials*: Didactic teaching materials can be presented to the trainee via several modes of multimedia presentation. These materials can be in the form of PowerPoint, Flash presentation, Video files, Audio files, Web links, etc.

5. CONCLUSION

This aspect of the GIMI project shows the practicality of a grid-enabled mammographic auditing and training system. The future trend in medicine is towards data integration, data sharing, and collaboration within hospitals/clinics through grid computing. However, the bandwidth of the network and security are crucial issues for such grid computing. In the first phase of this project as described here, we still have some limitations in the actual application - for instance our training system is currently a computer-based model. Ongoing research is concentrating upon developing this into an intelligent training system. To date the practicality of GIMI middleware in driving this mammography application has been demonstrated which we hope will serve as a roadmap for future work in this area.

6. ACKNOWLEDGEMENT

This research is funded by the DTI. We acknowledge the continued support of our colleagues from Oxford University.

REFERENCES

1. NHS breast screening programme. <http://www.cancerscreening.nhs.uk/breastscreen/index.html> [15 January 2008]
2. Crowley, R.S. and Medvedeva, O. "A general architecture for intelligent tutoring of diagnostic classification problem solving," AMIA. Annu. Symp. Proc., pp. 185-189, (2003).
3. Azevedo, R. and Lajoie, S.P., "The cognitive basis for the design of a mammography interpretation tutor", International Journal of Artificial Intelligence in Education, 9: 32-44, (1998).
4. Baker, M., Buyya, R. and Laforenza, D., "Grids and Grid technologies for wide-area distributed computing", Softw. Pract. Exper, (2002). DOI: 10/1002/spe.488.
5. Amedolia, S.R., Brady, J.M., McClatchey, R., Mulet-Parada, M., Odeh, M., Solomonides, T., "Mammogrid: Large-scale distributed mammogram analysis", Proceedings of the 18th Medical Informatics Conference (MIE2003) (Studies in Health Technology and Informatics, vol.95), St Malo, France, 194-199, (2003).
6. National Digital Mammography Archive. <http://www.i3archive.com>. [5 May 2005]
7. Power, D.J., Politou, E.A., Slaymaker, M.A. and Simpson, A.C., "Towards secure Grid-enabled healthcare", Softw. Pract. Exper, 35:857-871, (2005). DOI: 10.1002/spe.692.
8. Simpson, A.C., Power, D.J., Slaymaker, M.A. and Politou, E.A., "GIMI: Generic infrastructure for medical informatics", In Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems, pp. 564-566, (2005).
9. Simpson, A.C., Power, D.J., Slaymaker, M.A., Russell, D. and Katzarova, M., "On the development of secure service-oriented architectures to support medical research. International Journal of Healthcare Information Systems and Informatics, 2(2):75-89, (2007).
10. Gale, A.G., "PERFORMS - a self assessment scheme for radiologists in breast screening", Seminars in Breast Disease: Improving and Monitoring Mammographic Interpretative Skills, 6:148-152, (2003).
11. Scott, S.J. and Gale, A.G., "Breast Screening: PERFORMS identifies key mammographic training needs", The British Journal of Radiology, 000: 1-7, (2006). DOI: 10.1259/bjr/25049149
12. Digital imaging and communications in medicine. <http://medical.nema.org/> [01 March 2007]
13. Laboratory of Neuro Imaging, UCLA, <http://www.loni.ucla.edu/Software/> [02 June 2007]