

e-DiaMoND: The Blueprint Document

Overview

This e-DiaMoND blueprint document serves two key purposes: the first is that, by complementing the e-DiaMoND prototype and the research data collected during the project, it constitutes the third major deliverable from the two-year e-DiaMoND project; the second is that this document should provide the basis of a framework for projects—both technological and clinical—that might follow on from e-DiaMoND.

The original intention of this document was that it would determine how a full e-DiaMoND system might be deployed within the UK to support the National Health Service Breast Screening Programme. For example, such a system would have to be capable of interfacing with systems already operating within the National Health Service and would also have to meet the legal and ethical requirements associated with the storage and transfer of medical data. Recent developments have made the full deployment of such a system—at least within the foreseeable future—unlikely. As such, a significant proportion of this document presents a rather more generic blueprint that prescribes the requirements for an idealised health grid. This change of focus should ensure that the document is of use not only to the e-DiaMoND consortium, but also to the wider e-Health community.

We consider the technical, ethical and legal requirements for deploying a health grid within the United Kingdom. The document considers issues specific to an e-DiaMoND-like system that might support the NHS's Breast Screening Programme, as well as more generic requirements that any UK-based health grid would have to meet. It also considers issues pertaining both to health grids where the main focus is to support research and health grids where the main focus is to health care delivery.

We can, then, think of this blueprint document as being in three 'parts'. The first part, consisting of Sections 1 and 2, describes where we are and how we got here: that is, we discuss the motivation for e-DiaMoND and describe the e-DiaMoND pilot system. The second part, consisting of Sections 3- 5, describes the constraints that a UK-based health grid must work within. The third part, consisting of Sections 6-7, describes the possible ways forwards, both technologically and clinically.

The information presented in this document is, of course, subject to change and represents what was understood by the authors to be correct at the time of writing. To reflect the evolving nature of the topics under investigation, it is intended that a version of this blueprint will be taken forward in future projects, where it will continue to evolve as a 'living' document.

Mike Brady, David Gavaghan, Steve Harris, Marina Jirotko, Alan Knox, Sharon Lloyd,
Eugenia Politou, David Power, Andrew Simpson, Mark Slaymaker, and John Williams
November 4, 2005

Contents

1 Introduction	2
1.1 The breast screening process	2
1.2 The e-DiaMoND project	4
1.3 Project structure	6
1.4 Critical success factors	6
1.5 Related projects and initiatives	7
1.6 The structure of this document	9
2 The e-DiaMoND pilot system	10
2.1 Relevant concepts	10
2.2 A functional description	11
2.3 Architecture overview	11
2.4 Usage scenarios	12
2.5 System design	13
2.6 Component interactions	15
2.7 Security assumptions made for the pilot system	15
2.8 Database structure	16
2.9 Benefits	19
3 Legal and ethical constraints	21
3.1 The legal structure of the NHS	21
3.2 Legal constraints	22
3.3 Ownership of data	24
3.4 The complexities of obtaining ethical clearance for research	25
3.5 The social issues of enabling digital patient data	27
3.6 Summary of requirements and process for real system deployment	27
4 External dependencies	28
4.1 Standards, initiatives and policies	28
4.2 Systems	35
4.3 Services	38
4.4 Summary	40
5 Security issues	41
5.1 Security	41
5.2 Security use cases	43
5.3 Technological issues	47
5.4 A gap analysis	48
6 Technology options	51
6.1 Web services	51
6.2 SOA	51
6.3 SOAP	52
6.4 ebXML	53
6.5 Other web services standards and specifications	55
6.6 Grid services	56
6.7 WS-RF	56
6.8 Other relevant middleware	57
6.9 Web services and the NHS	58
6.10 Summary	59
7 e-DiaMoND-specific issues	60
7.1 e-DiaMoND applications	60
7.2 The breast screening domain	60
7.3 Screening	60
7.4 Training	62
7.5 Epidemiology	63
7.6 Data mining	64
7.7 Technical constraints	64

Part I: The Past

1 Introduction

e-DiaMoND [1] is a project that has been deemed to be of significant potential benefit to the NHS and to the female population of the United Kingdom. The stakeholders of the project and the project team have together developed a vision of how the work undertaken within the project might play a role in the future of digital mammography within the United Kingdom, and might also influence the direction of e-health and, indeed, e-research in general.

This e-DiaMoND blueprint document serves two key purposes. The first of these key purposes is that, by complementing the e-DiaMoND prototype and the research data collected during the project, it constitutes the third major deliverable from the two-year e-DiaMoND project. The second of these key purposes is that this document should provide the basis of a framework for projects—of both technological and clinical varieties—that might follow on from e-DiaMoND.

In particular, we consider the technical, ethical and legal requirements for deploying a health grid to support healthcare delivery and research within the United Kingdom. (The aims in this respect, are rather more focused than those of the EU health grid initiative (see, for example, [2, 3]).) The document considers issues specific to an e-DiaMoND-like system that might support the NHS's Breast Screening Programme [4, 5], as well as more generic requirements that any UK-based health grid would have to meet. It also differentiates between issues pertaining to health grids where the primary aim is to support research and health grids where the primary aim is to support health care delivery. (Although it would appear to that there is a convergence path is emerging.)

To provide an appropriate backdrop for the issues documented, we first reprise the motivation for the e-DiaMoND project.

1.1 The breast screening process

The original motivation for the e-DiaMoND project was presented in [1]. The media coverage accorded to the Labour Party's ambitious pre-election pledge that "every suspected breast cancer case, not just those deemed urgent, will be screened within two weeks by 2008," [6] ensure that the driving forces behind the original project are no less relevant at the end of the project as they were at the start. As such we recall that original motivation here.

Breast cancer is a major public health issue in the western world, where it is the most common cancer among women. In the European Union, for example, breast cancer represents 19% of cancer deaths and fully 24% of all cancer cases. Breast cancer is diagnosed in a total of 348 000 cases annually in the USA and EU [7], and kills almost 115 000 annually. Approximately one in eight of women will develop breast cancer during the course of their lives, and one in 28 will die of the disease. The threat is negligible for women under 30, with the threat rising sharply until the age of 50, and continuing to rise (but less sharply) thereafter [8]. There were 900 000 new cases of breast cancer worldwide in 1997. Such alarming statistics are now being replicated in eastern countries as diets and environment become more like their western counterparts.

The earlier a tumour is detected the better the prognosis. A tumour that is detected when its size is just 0.5cm has a favourable prognosis in about 99% of cases, since it is highly unlikely to have metastasized.¹ Few women can detect a tumour by palpation (breast self-examination) when it is smaller than 1cm, by which time (on average) the tumour will have been in the breast for up to 6–8 years. The five-year survival rate for localized breast cancer is 97%; this drops to 77% if the cancer has spread by the time of diagnosis and to 22% if distant metastases are found [9].

The World Health Organisation's International Agency for Research on Cancer (IARC) concluded in 2002 that mass screening via mammography reduces mortality [10]. The findings are based on the work of an IARC working group, comprising 24 experts from 11 countries. This working group evaluated all currently available international evidence on breast screening. The working group found a 35% reduction in mortality from breast cancer among women in the 50–69 age group who were screened as opposed to those who were not screened: this equates to one life being saved for every 500 women screened.

Although the precise figures vary from country to country, the above discussion presents a clear rationale for mass screening. The United Kingdom was the first country to develop a national screening programme, though several other countries have now also established such programmes, including Sweden, Finland, The Netherlands, Australia, Ireland, France, Germany and Japan. (See [11] for an overview of national screening programmes.) In the United States, on the other hand, most mammograms are concerned with symptomatic patients [11].

The UK programme resulted from the Government's acceptance of the report of the committee chaired by Sir Patrick Forrest [12]. The report was rather bullish about the effects of a screening programme:

¹Metastasis is the process by which malignant cancer cells can spread via the blood and lymphatic systems to distant organs.

“by the year 2000 the screening programme is expected to prevent about 25% of deaths from breast cancer in the population of women invited for screening ...On average each of the women in whom breast cancer is prevented will live about 20 years more. Thus by the year 2000 the screening programme is expected to result in about 25 000 extra years of life gained annually in the UK.”

When the Breast Screening Programme (BSP) was first set up in 1988, research suggested that women aged 65 and over were less likely to accept their invitations for screening—it was understood that uptake decreased with increasing age. However, a woman’s risk of breast cancer increases as she gets older, and this, coupled with longer life expectancies means that older women now gain greater benefits from the early detection of breast cancer [5].

When the e-DiaMoND project started, in early 2003, the BSP was inviting women between the ages of 50 and 64 for breast screening once every three years, with this age range being derived from the fact that the breasts of pre-menopausal women, particularly younger women, are composed primarily of milk-bearing tissue which is calcium-rich; this milk-bearing tissue involutes to fat during the menopause—and fat is transparent to x-rays.

If a mammogram displays any suspicious signs, the woman is invited back to an assessment clinic where other views and other imaging modalities may be utilised. There are 92 screening centres in the UK, employing a total of 230 radiologists. Each radiologist reads, on average, 5 000 cases per year, with some reading up to 20 000.

In December 2004 the Breast Screening Programme pledged a major extension to women up to the age of 70 and available on request to women over 70. As a result of this, in 2002/3 an additional 120 000 women were invited for screening and the expansion continues [5].

To date, the Breast Screening Programme has screened more than 14 000 000 women, and has detected over 80 000 cancers. (This figure is, of course, independent of those cancers detected as a result of symptomatic screening.) Research published in the British Medical Journal in September 2000 [13] demonstrated that the NHS Breast Screening Programme is saving at least 300 lives per year. It is predicted that this number will rise to 1 250 per year by 2010. More precisely, it was demonstrated that the National Health Service Breast Screening Programme resulted in substantial reductions in mortality from breast cancer between 1987 and 1998. By 1998, mortality was reduced by an average of 14.9% in those aged 50–54 and 75–79, which would be attributed to treatment improvements. In the age groups also affected by screening (55–69), the reduction in mortality was 21.3%. Hence, the estimated direct contribution from screening was calculated in [13] as 6.4%.

It is estimated that 25% of cancers are missed at screening; the actual rate of such cancers, which are termed *interval cancers*, has been greater than the anticipated rate presented in the Forrest Report [14].

Increasingly, there are calls for mammograms to be taken every two years and for both a cranio-caudal (head-to-toe direction) and mediolateral oblique (armpit-to-opposite hip) image to be taken of each breast. At present only mediolateral oblique images are taken.

The opportunity for Information Technology to assist healthcare professionals in this particular domain can be illustrated by consideration of available statistics for screening in the USA. Currently, some 26 000 000 women are screened in the USA annually (that is, nearly half of the 55 000 000 women who are screened each year worldwide). In the USA there are 10 000 mammography-accredited screening centres. Of these, 39% are community and/or public hospitals, 26% are private radiology practices, and 13% are private hospitals. Although there are 10 000 mammography centres, there are only 2 500 mammography specific radiologists—there is a worldwide shortage of radiologists and radiologic technologists (referred to in the UK as radiographers).² Whereas expert radiologists have cancer detection rates of 76–84%, generalists have rates that vary from between 8–98% (with varying numbers of false-positives). The number of cancers that are deemed to be visible in retrospect, that is, when the outcome is known, approaches 70%.

There is some evidence that staff shortages in mammography seem to stem from the perception that it is ‘boring but risky’: a significant proportion of all malpractice lawsuits in the United States are against radiologists, with the failure to diagnose breast cancer becoming one of the leading reasons for malpractice litigation. The shortage of radiologists is driving the development of specialist centres and technologies (both hardware and software) that aspire to replicate their skills. Screening environments are ideally suited to information technology solutions, as they are repetitive and require objective measurements.

As we have noted, screening has already produced encouraging results. Process changes can help detection rates and reduce recall rates. Indeed, recall rates drop by 15% when using two views of each breast [15]. It has been demonstrated empirically that double reading (where two radiologists look at each mammogram) greatly improves screening results. The number of cancers missed when mammograms are double read is half that associated with single reading. However, double reading is expensive and in any case there are too few screening radiologists. In addition, a study at Yale of board-certified radiologists demonstrated that the

²Note that (in UK terminology) radiographers are the staff who take the mammograms and radiologists are the clinicians who assess, or ‘read’, mammograms and provide the initial diagnosis and assessment of patients.

radiologists disagreed 25% of the times about whether a biopsy was warranted and 19% of the time in assigning patients to one of five diagnostic categories. Recently, it has been demonstrated that single screening plus the use of computer-aided diagnosis (CAD) tools—image analysis algorithms that aim to detect microcalcifications and small tumours—also greatly improves screening effectiveness, perhaps by as much as 20%.

Post-screening, the patient may be assessed by other modalities such as palpation, ultrasound and, increasingly, by MRI: 5-10% of those screened do have such an extended assessment. Following assessment, around 5% of patients have a biopsy taken. In light of the number of tumours that are missed at screening (which reflects the complexity of diagnosing the disease from a mammogram), it is not surprising that clinicians err on the side of caution and order a large number of biopsies. In the US, for example, there are over 1 000 000 biopsies performed each year: a staggering 80% of these reveal a benign (non-cancerous) outcome.

It has been reported in the USA [16] that between screenings 22% of previously taken mammograms are unavailable or are difficult to find, mostly due to the fact that they have been misfiled in large film archives: lost films are a daily headache for radiologists around the world. In addition, in the same study it was reported that 50% of previously taken mammograms were found only after significant effort [16].

To illustrate the problem and how it is possible to lose a patient's mammograms, consider the following scenario. One of the e-DiaMoND partner hospitals screens around 30 000 women per year. These mammograms are stored in racks similar to large bookcases; however, the 'books' are only a few millimetres thick and don't have spines to distinguish them from their neighbours. If a patient's records are misfiled then they are virtually impossible to trace. The storage of such images and associated information electronically has the potential to eliminate the potential for the loss of images. In addition, the judicious use of unique identifiers and referential integrity can ensure that different 'bits of information' pertaining to the same patient remain 'tied together'.

1.2 The e-DiaMoND project

The primary focus of the e-DiaMoND project was to build a pilot of a distributed computer system to determine potential benefits for the NHS Breast Screening Programme if it were to adopt such an approach. In this respect, e-DiaMoND was one of the largest projects funded by the UK's national e-Science Programme [17], the main aims of which were:

- to build a computational infrastructure, or 'grid', to support large-scale research, and
- to identify potential applications—from varying domains—for such an infrastructure.

The e-Science Programme adopted a phased approach—termed the *Core e-Science Programme* [18], with the first phase consisting of the following [18]:

- a National e-Science Centre linked to a network of regional grid centres;
- generic grid middleware and demonstrator projects;
- IRC (Interdisciplinary Research Centre) research projects;
- support for e-Science pilot projects;
- participation in international grid projects and activities; and
- the establishment of a grid network team.

The second phase consisted of [18]:

- a National e-Science Centre linked to a network of regional grid centres;
- support activities for the UK e-Science community;
- an Open Middleware Infrastructure Institute (OMII);
- a Digital Curation Centre (DCC);
- new exemplars for e-Science; and
- participation in international grid projects and activities.

The e-DiaMoND project was funded as a first phase demonstrator project.

The long-term vision for the e-DiaMoND consortium—which the e-DiaMoND project was intended to serve as an initial step along the path for—is to recognise the current constraints which exist with the non-digital

and low technology environment in which clinicians work, and seek to develop solutions which increase the efficiency of the service and improve the tools available to those staff.

As an example, one benefit of a digital environment might be associated with data transport. At present, if a Breast Care Unit is short staffed, the movement of data for reading to be performed at another centre is problematic. (Although, ironically, the longer the staff shortage, the easier it is to deal with in terms of establishing procedures and scheduling.) An infrastructure such as that developed within e-DiaMoND would enable easy, dynamic and real-time movement of information with high levels of security to facilitate the sharing of workloads. (The issues pertaining to the movement of such data within a high-speed network environment are being investigated within the ESLEA project [19].) It would be practical with such a solution to request second readings from another radiologist from another Hospital Trust, or even to seek arbitration services. Further, image processing and data mining are aspects of digital mammography which could offer benefits to clinicians to aid in both the detection of abnormal features as well as in the diagnosis of cancers: such techniques could be realised 'on-demand'.

In the above example, we see the two complementary aspects of e-DiaMoND that are reflected in Part III of this document: the technological and the clinical. The former aspect in this case—the secure transfer of clinical data—is concerned with generic problems; the latter aspect—the use of data mining and image processing techniques specific to breast cancer detection—is concerned with domain-specific problems. This marriage of complementary concerns has served the e-DiaMoND project well.

The present authors feel that this marriage of complementary concerns has served the project well.

As some further examples of the potential benefits of a digital approach, consider the following.

- Standardisation of images from different centres in the UK would enable a database to be built using scans taken on different machines or at different sites, with the effect being that they would all appear as if they were produced on the same machine under the same conditions. This would enable temporal comparisons to be performed, which—if coupled with methods for removing known changes with age, HRT and other contributing factors—would offer the potential to detect any abnormal changes over time.
- Data mining techniques could also be used to prompt radiologists for likely areas of interest on images and could possibly be considered for second or arbitration reading.
- Web-based training could enable the sharing of interesting cases across the UK as well as removing the need for intensive use of senior radiologists' time to train more junior staff. Training would also be auditable thus enabling the Breast Screening Programme to monitor performance of staff more closely.

Each of the above avenues have been explored within e-DiaMoND, and each remains an avenue of fruitful research for future projects.

With the above vision in mind, the e-DiaMoND project has developed a pilot system that demonstrates some of the key features of the future digital world, as well as providing a resource of anonymised breast imaging data to aid in the development of innovative techniques to assist with detecting and diagnosing breast cancer.

The long-term benefits offered by the 'real' deployment of such a system would be numerous, and would include the following [1].

- Patients would benefit from secure storage of films, better and faster patient record access, better opinions, and the lowering of the likelihood of requiring a biopsy.
- Radiologists would benefit from computer assistance, massively reduced storage space requirements, instant access to mammograms without loss of mammograms, improved early diagnosis (because of improved image quality), and greater all-round efficiency leading to a reduction in waiting time for assessment. Radiologists will also benefit from the application of relevant data mining technology to the database to seek out images that are 'similar' to the one under consideration, and for which the diagnosis is already known.
- Administrative staff would benefit significantly from faster image archiving and retrieval, and automated support for in the administration process.
- Hospital Trust managers would benefit from the overall reduced cost of providing a better quality service.
- Hospital IT Managers would benefit from the greatly reduced burden on their already over-stretched resources. The regional consolidation in storage implied by a system such as e-DiaMoND would have the potential to reduce drastically their costs through reduced equipment and staffing requirements and support contracts.
- Researchers would benefit, as such a national resource—together with associated computing tools—would provide an unparalleled resource for epidemiological studies based on images.

1.3 Project structure

A project as ambitious as e-DiaMoND, featuring complementary technological and clinical strands, could never be undertaken by any single organisation. To ensure that a pilot system might engender interest in the end-user community to guarantee the possibility of follow-on projects, the needs and requirements of those potential end-users—both from the clinical and research domains—must be carefully considered.

This project could not be undertaken by computer scientists without detailed understanding of the procedures and needs of the radiologists and researchers for whom the system is being built to support. In common with other complex multidisciplinary UK e-Science projects, it was, therefore, essential to the success of the e-DiaMoND project that it was undertaken by a project team with diverse and complementary skills. As such, the e-DiaMoND project brought together a blend of expertise in clinical medicine, software engineering and medical imaging from both the research and the development perspective.

For example, the requirements pertaining to patient confidentiality and ethical use of data within projects such as e-DiaMoND are very complex (as we shall see in detail in Section 3). Only via a combination of those understanding the domain and the requirements associated with it and those that might prescribe potential solutions to these problems could any kind of success be guaranteed.

With respect to the underlying technology, the e-DiaMoND partnership involves a three-way collaboration to develop the system, which is based on a combination of open source, proprietary off-the-shelf products, and bespoke code. This three-way collaboration involved IBM (UK), the University of Oxford (in the form of the Computing Laboratory, the e-Science Centre, and the Engineering Science department), and Mirada Solutions Limited (now Siemens Molecular Imaging), a spin-out company from the University of Oxford.

IBM provided the hardware for the project under a Shared University Research (SUR) grant. The terms of this grant meant that IBM donated equipment to a university (in this case, the University of Oxford) for use in a research project. IBM also provided licenses for software such as DB2 and the Tivoli storage management package. In addition, a number of IBM staff were assigned full-time to the project.

The University of Oxford, which had staff in the Computing Laboratory and the Engineering Science Department, was the main academic partner. The development of the underlying database and its grid services was the responsibility of the Computing Laboratory; the development of data mining tools and techniques was the responsibility of the Engineering Science Department.

Two very important aspects of the project were undertaken by (what was) Mirada: the development of user applications for screening and radiographer training, and the development of a range of leading edge technologies for processing mammograms.

A key role is also played by the project's clinical partners, who, as well as being charged with progressing potential applications, were responsible for collecting the data to be used within the project. The clinical partners were as follows.

- King's College London and Guy's and St. Thomas' NHS Trust Hospitals, London, where strong expertise exists in medical imaging, including image-guided intervention, tissue modelling, and the measurement of change using medical images.
- The John Radcliffe Hospital Trust, Oxford, contributed a strong history of research and practice in medical imaging and medical physics, and has existing extensive links with the groups in the Engineering Department.
- St. George's Hospital, London (with University College London) has a large breast screening centre located at the hospital and provided expert opinion on user requirements for screening and teaching.
- The Edinburgh Breast Care centre (with Edinburgh University) was also focused on training aspects, and has an extensive track record in evaluating the use of IT systems to support the work of breast-screening radiologists.

1.4 Critical success factors

The e-DiaMoND 'Critical Success Factors' document [20] provides a detailed definition of success for the e-DiaMoND project, which, in turn, was driven by the definition of the scope of the project as defined in the 'Requirements Overview' document and the 'Project Proposal' document. It is worth repeating these critical success factors here.

"The primary goal of the e-DiaMoND project is to develop a prototype working system by the project end date where a working system is defined as having the following characteristics.

- It has a significantly large distributed database of mammograms (400 cases per site with a majority annotated).

- It aligns with and complies with new IT policies for the NHS in that it is secure and wins the confidence of the relevant legal, ethical and NHS Trust IT officers. In addition, the system follows all known guidelines for the deployment of NHS patient and health records.
- It is scalable and is designed in such a way that it could scale to cope conceptually with millions of images spread around the 90+ Breast Care Units in the UK.
- It is effective in that it is fast, it is useful to the clinicians in the areas of screening, training, epidemiology and computer aided detection, and it is intuitive for the users.
- It must be built such that upgrades of platform or image analysis software are graceful.
- It is reusable, in that the platform could be used as a foundation for other e-health projects.
- It is based on grid architecture.”

At the end of the two-year e-DiaMoND project, the team has:

- delivered a pilot solution to demonstrate the feasibility of utilising a grid-based infrastructure to support the storage and transport of breast imaging information across four clinical sites;
- developed prototype platforms to support a number of important application areas; and
- collected a significant number of high-quality anonymised annotated mammograms from four sites.

We are now in a position to consider each critical success factor in turn:

1. **It has a significantly large distributed database of mammograms (400 cases per site with a majority annotated).** While the initial ambitious target of 400 cases per site has not been achieved, we have still collected over 1 000 cases of high quality annotated mammograms.
2. **It aligns with and complies with new IT policies for the NHS in that it is secure and wins the confidence of the relevant legal, ethical and NHS Trust IT officers. In addition, the system will follow all known guidelines for the deployment of NHS patient and health records.** This is one of the motivations for this document.
3. **It is scalable and is designed in such a way that it could scale to cope conceptually with millions of images spread around the 90+ Breast Care Units in the UK.** The pilot system was designed to be scalable.
4. **It is effective in that it is fast, it is useful to the clinicians in the areas of screening, training, epidemiology and computer aided detection, and it is intuitive for the users.** This is a property of the individual applications, and beyond the scope of this document.
5. **It must be built such that upgrades of platform or image analysis software are graceful.** This has been accomplished as far as is possible.
6. **It is reusable, in that the platform could be used as a foundation for other e-health projects.** This is one of the motivations for this document.
7. **It is based on grid architecture.** This has been accomplished.

From the above, we see that there are two motivations for this document, derived from two separate critical success factors (2 and 5). Further, each is necessitated as a result of change: in terms of the introduction of a National IT Programme for the NHS [21]—which had an impact on the second critical factor—and the relative immaturity and subsequent evolution of the grid technologies upon which the e-DiaMoND pilot system was built—which had an impact on the fifth critical success factor. With respect to the second of the above critical success factors, we describe in Part II of this document what a system such as e-DiaMoND would have to satisfy to fit within the prescribed context—as it currently stands; with respect to the fifth of the above critical success factors, we provide an overview of the current e-DiaMoND pilot system in Section 2.

1.5 Related projects and initiatives

MammoGrid and NDMA

Key to the e-DiaMoND project is the facility to standardise digital mammogram images, a capability which will help radiologists to compare and evaluate mammography scans stored within e-DiaMoND accurately—no matter where or when they were created. By focusing on providing a solid grid infrastructure using open source and commercially available products where possible, and developing effective clinical applications for data acquisition, screening, training and epidemiology as well as developing data mining techniques to improve

cancer detection and treatment in the future, the e-DiaMoND project was distinct from other similar projects as detailed below.

MammoGrid [22] was an EU-funded project with the aim of developing a demonstrator for use in epidemiological studies, quality control and validation of computer aided detection algorithms for mammographic images. Whereas e-DiaMoND concentrated upon teaching, tele-diagnosis and algorithm development for data mining to influence the future of breast screening programmes by encouraging them to embrace and pioneer the SMF technology, MammoGrid focus on quality control for breast cancer screening, epidemiology of breast cancer from a European perspective, and the creation of CADe algorithms.

NDMA [23] was an IBM SUR grant backed project, managed out of the University of Pennsylvania. The goal of this project was to develop an Electronic Medical Record (EMR) data grid and repository. The focus has very much been on the development of architecture to support the storage and transfer of mammography information for the US health market, where patients typically do not receive treatment from a central trust or health organisation. The drivers for this development were on reducing the cost of treating breast cancer in an environment where there is no breast screening programme and where the needs of the insurance market drives these types of developments. In addition, the UK Breast Screening Programme is primarily film-oriented, whereas full field digital environments are the norm in the USA.

HealthGrid

The HealthGrid White Paper [3] describes the rationale behind the EU HealthGrid efforts. The White Paper has similar (albeit broader) aims to this document, and consists of nine sections:

1. From Grid to HealthGrid: prospects and requirements.
2. Creating a compelling business case for HealthGrid.
3. Medical imaging and medical image processing.
4. Computational models of the human body for therapy planning and computer assisted intervention.
5. Grid enabled pharmaceutical R&D: PharmaGrids.
6. Grids for epidemiological studies.
7. Genomic medicine grid.
8. From Grid to HealthGrid: confidentiality and ethical issues.
9. Legal approaches to the HealthGrid technology.

It is clear that there will be an overlap of concerns between this document and that White Paper—with confidentiality, ethical, and legal approaches being of particular relevance in this regard. However, while the HealthGrid initiative concerns itself with pan-European issues, our focus is the UK. In addition, while a key aspect of the HealthGrid White Paper involves explaining the rationale for adopting a ‘grid’ approach, we do not feel that making such arguments is within the scope of this document.

Other e-Science healthcare projects

There have been a number of other healthcare-related projects funded by the UK e-Science Programme and by the UK’s Medical Research Council, with the notable examples including the following.

- Axiope [24].
- CancerGrid [25].
- CLEF[26].
- eFamily [27].
- Equator [28].
- Integrative Biology [29].
- IXI [30].

Note that this is by no means meant to be an exhaustive list. Most, if not all, of these projects have met (or are meeting) similar social, legal, ethical and technical challenges faced by the e-DiaMoND project team.

1.6 The structure of this document

The structure of the remainder of this document is as follows.

First, in Section 2, we provide a necessary, but brief, overview of the e-DiaMoND pilot system that has been implemented: this should help to put the discussions of the later chapters into some sort of context. Together with this section, Section 2 forms Part I of this document.

Taken together, Sections 3-5 constitute Part II of this document.

In Section 3, we consider the legal and ethical constraints that any health grid—supporting either healthcare delivery or research—deployed within the United Kingdom would have to satisfy.

In Section 4, we address the external issues that are relevant to the development of a deployable solution. Of primary concern in this regard are the IHE initiative, NPfIT, and the requirements of the NHS Breast Screening Programme.

Then, in Section 5, we consider the security requirements pertaining to a UK-based health grid. Whereas other non-functional issues, such manageability, dependability, and quality of performance are largely domain- and application-dependent, security issues are predominantly generic. This is primarily due to the fact that they are derived from British and European legislation.

Having considered the generic, application-independent constraints, we then, in Sections 6 and 7 start to consider ways of moving forward. In the former, we provide an overview of the current situation with respect to technology choices for the development of a web service-based health grid, and in the latter we consider those issues pertaining to BSP-specific applications associated with e-DiaMoND. These two sections constitute Part III of this document.

While we have attempted to divide the concerns of this document into discrete sections, there are (some-what inevitably) interdependencies between different aspects: some security requirements emerge due to legal constraints, some of our description of the pilot system relies upon information pertaining to standards, and so on. Where this is the case, we have tried to limit duplication—whilst at the same time retaining readability—as much as possible.

2 The e-DiaMoND pilot system

In this section we provide a brief overview of the pilot system that has been developed within the e-DiaMoND project, the key technical objective of which was to develop the blueprint design of a grid for digital mammography. The pilot system was developed as a proof of this design. It is for this reason that we focus exclusively on the underlying grid infrastructure developed within this section, and do not consider the prototype applications that have been developed. Domain-specific applications associated with e-DiaMoND are discussed in Section 7.

The project team assumed technical risk in adopting a grid approach to the problem that had to be solved: part of the nature of the project was to research the issues pertaining to the implementation of a practical, non-trivial grid. It was intended that the experiences gained and shared in this ‘demonstrator’ fashion would be valuable to the wider e-Science community.

Given the above, it was determined by the project team that the most appropriate way forward was to define a framework of services that provide some abstract solution.

Adopting a technical solution based on commercial products would have had the benefit that such an approach would lead to reduced risk and faster delivery; however, a drawback of adopting such an approach is that it is more difficult to answer questions pertaining to dynamic virtual organisations, distributed security responsibility, etc. Adopting a ‘pure’ grid approach would address these issues, but would have been more time-consuming to build and would have involved significant technical risk. To balance the two, the project team attempted to steer a ‘third way’ in the development of the e-DiaMoND infrastructure: novel and immature grid technologies were used in conjunction with established commercial off-the-shelf products.

In describing the pilot infrastructure, we consider first the architecture, before providing an overview of the operation of the underlying database. The description of the architecture is based on [31]; the description of the database is based on [32]. We start, though, by introducing a number of concepts that are relevant to our discussion.

2.1 Relevant concepts

Before discussing the architecture, we first introduce four key concepts and technologies in order to aid the narrative. We will revisit some of these (in particular, DICOM) in greater detail when, for example, we consider the protocols and standards that any future system would have to adhere to.

Service-oriented architecture

The use of service-oriented architectures is an approach to distributed computing that treats software and data resources as services that are available on a network. This approach is typified by the Open Grid Services Architecture (OGSA) definition of *grid services* [33], where a grid service is essentially a stateful web service with a defined lifetime that conforms to a set of interfaces and behaviours with which a client may interact [34].

OGSA defines an architecture whereby service providers create a description of the service they offer and publish it in a registry. Service requesters ‘discover’ service descriptions in the registry and ‘bind’ to a service implementation offered by the provider.

The registry that a particular grid client node interacts with to discover e-DiaMoND services can be specified as a parameter. The contents of the registry effectively define the e-DiaMoND grid with respect to that particular client. A client may be easily re-configured to point to a different registry. All the clients in the grid may share the same registry, but are not forced to, and grid services may be added to and removed from the registry dynamically. These features have been exploited fully to ensure that the e-DiaMoND grid is both flexible and extensible.

DICOM

DICOM (Digital Imaging and Communications in Medicine) is a standard for communicating and managing digital medical data [35]. DICOM defines communications protocols for manipulating medical objects in a distributed computing environment. It also defines a file format for storing those objects.

The e-DiaMoND system has a DICOM interface at the data acquisition workstation to capture images from DICOM compliant scanners and digital x-ray machines. The *DICOM file format* is used within the e-DiaMoND system, but OGSA grid services—as opposed to *DICOM protocols*—are used internally for communication. Specifically, the x-ray images are stored as DICOM image objects whereas the medical and patient reports are stored in a DICOM Structured Report (SR) format.

Standard Mammography FormTM (SMF)

SMF technology was developed by Mirada Solutions and Oxford University, and models the complete image creation process for x-ray mammography to determine the height of non-fatty tissue in the breast for each

pixel in the image [36]. In essence, it provides a normalisation of x-ray mammograms that allows comparison of images and supports the development of data mining algorithms and computer-aided detection technologies. The SMF algorithm takes a DICOM image file as input and generates three outputs: a segmentation map, a breast tissue density map, and some simple metrics.

Within e-DiaMoND, a newly captured DICOM image file is sent to the grid from the acquisition workstation. Then, the DICOM file is processed to create its breast density map, that is, the SMF version of the image which is also stored.

OGSA-DAI

Open Grid Services Architecture–Data Access and Integration (OGSA-DAI) was a UK e-Science project, the aim of which was to provide an extension to OGSA to allow data resources, such as databases and other data sources, to be incorporated within an OGSA framework [37]. The e-DiaMoND architecture makes use of OGSA-DAI services to represent the non-image and image data resources in the grid.

2.2 A functional description

A ‘grid’ is a distributed computing system for flexible, secure, coordinated resource sharing between virtual organisations, where a virtual organisation is defined to be a dynamic collection of individuals, institutions and resources.

The core e-DiaMoND system consists of middleware and a virtualized medical image store to support the e-DiaMoND data grid concept. The virtualised image store is made of physical databases, each owned by a different organisation that is participating in the e-DiaMoND grid. (In a rolled-out system, these organisations would be Breast Care Units (BCUs).) The e-DiaMoND grid is formed by participating BCUs coming together as a virtual organisation that unites their individual databases as a single logical resource.

The key functions of e-DiaMoND are as follows.

- **Image acquisition.** This is the process of inputting x-ray mammograms into the system. A radiologist at the Image Capture Workstation takes scanned x-ray films and adds patient information. The result of this process is a DICOM image file, which is passed into the e-DiaMoND grid.
- **Query.** An administrator may query data in the system to set up a reading session as part of the screening process; a radiologist may make ad hoc queries in screening, or in constructing sets of images suitable as training cases; an epidemiologist may construct complex long-running queries that run across the entire archive.
- **Image retrieval.** This is the retrieval from the grid of specified DICOM image files and reports. DICOM files can be retrieved individually, or as a collection of all the files belonging to a particular series or study or patient.
- **Diagnosis reports.** The system captures and manages reports made by radiologists during screening.
- **Image processing.** This covers processes that categorise or manipulate the image data in support of data mining and Computer Aided Detection (CADe) services in the grid.

For a successful, deployable version of the e-DiaMoND system, these functions would have to be implemented in a grid that allows BCUs to collaborate with each other but maintain individual policies on how data for which they are responsible is distributed and shared. In addition:

- the system must allow BCUs to form a virtual organisation for breast screening without requiring any central authority or centralized IT resources;
- it must support a workflow for breast screening; and
- it must allow the same or different BCUs to form virtual organisations for other applications for the mammography resource—with training and epidemiology being the examples used to demonstrate this within e-DiaMoND.

2.3 Architecture overview

Figure 1 provides a functional view of the e-DiaMoND architecture.

Interfacing with this architecture is a GUI application developed by Mirada that allows the capture of images and patient data. In addition, Mirada has developed a GUI application for Breast Screening.

Client interaction with the grid is through a set of services managed by a registry. These services can be divided into data services—that allow each hospital to see all of the data owned by the participating nodes—and compute services—that can perform potentially complex and long-running calculations on the image data.

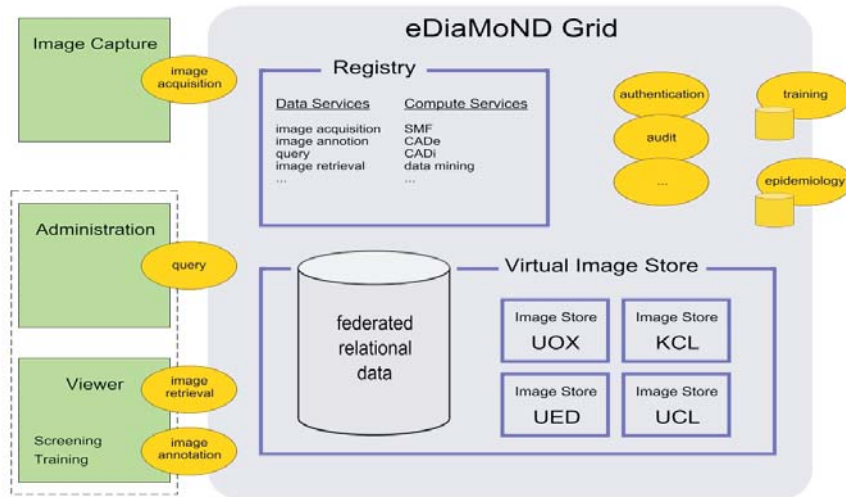


Figure 1: e-DiaMoND architecture: a functional view

The set of services in a registry effectively defines the function of the grid. Clients can interact with the registry to discover services. Flexibility is introduced by manipulating the services available in a registry. The registry may maintain several services with the same service interface offering different levels of quality of service.

Digitised mammograms are very large binary objects. It is possible to store these in a relational database, but this is not as efficient as storing the image in a filesystem and holding a reference in the database. As such, the digitised mammograms are stored as files in e-DiaMoND. Each node on the grid maintains its own image store consisting of a relational database of patient data and image metadata with records linked to images in the filesystem. The e-DiaMoND grid ensures that the individual image stores can be represented as a single large virtual image store that is accessible to clients at all nodes.

2.4 Usage scenarios

The following usage scenarios have influenced the development of the e-DiaMoND pilot.

Image acquisition

The scenario for image acquisition is as follows:

“A technician logs on to an acquisition workstation. She scans all available studies for the patient from a patient folder in which all the paper records for that patient are present, with 2-4 mammograms per study, some with magnification views and a limited number with two mammograms per view per breast representing retakes or partial view cases.

“The patient folders may typically contain screening forms, patient medical history information and past diagnosis reports. The technician will enter all available information about the patient from the paper forms and will associate the images with the patient details split by study. Identifying patient information will be removed and new names made up by the system (fictionalised). She will then anonymise the images and save them as individual DICOM files.

“These files will be transferred to the grid for persistent storage. Once the grid receives a file for storage, it will request CADE processing from a remote CADE service, and will store the resultant DICOM SR [Structured Report] files.”

Screening management

The scenario for screening management is as follows:

“The BCU administrator will manage the workload versus capacity within the system. On a regular basis she will log on to the work management console and review how many cases are outstanding for first or second reading and arbitration. She will balance this against her manual records of staff

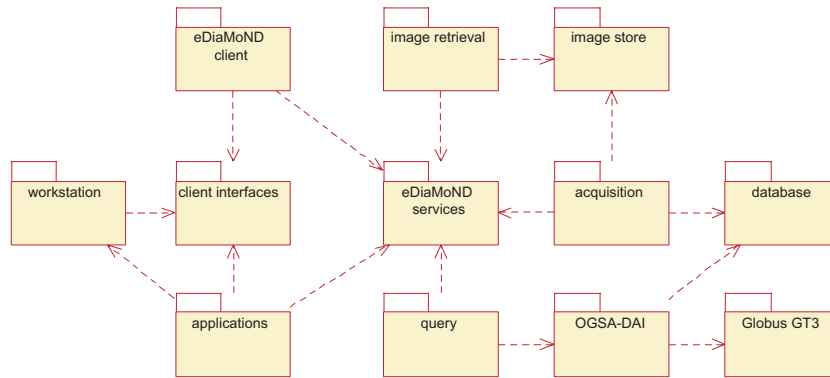


Figure 2: e-DiaMoND primary components

availability (i.e., number of radiologists performing reading and number of hours availability per radiologist) and average statistics for the BCU for time required to read a case.

“If the BCU administrator judges that the workload is greater than the reading capacity, she may decide to request that a selection of cases is read remotely at a site that has excess capacity. In this scenario we assume that the administrator, through telephone calls, email, etc. has determined that a particular BCU has capacity to help out. She then selects the cases that she wishes to be read remotely and selects the remote site to which she wishes them to be sent. Remote read requests could be for complete case reading, single reading or arbitration cases. The case references will then be sent to the remote site where then they will be added to the remote work-list.

“Once the required reading has been performed remotely, the SR file containing the results will be returned to the originator where it will be managed to completion. This is necessary since batched letter production is based on scheduling of radiologist availability to conduct the assessment clinics for recall patients.”

Reading

“A radiologist logs on to the workstation to perform a screening session. The oldest unread or ‘second reading’ case is retrieved from the grid at application start-up and presented to the radiologist. Images are presented in ideal reading conditions including correct alignment, auto-adjustment/balancing of brightness/contrast, pre-registered, labels removed and replaced with standard labels, etc. The radiologist typically will not need to make further adjustments to the images in order to form an opinion, however, in certain cases where there is uncertainty advanced image processing functions may be used. The radiologist will need to view each image pair of the current images side-by-side at real size and to look at each image in full size (not sub-sampled). She may also toggle CADe results on/off after preliminary review of the case. Likewise, she may toggle first reader opinion annotations and opinion on/off. The availability of certain options such as these could be configurable determined by site-policy.”

2.5 System design

Figure 2 depicts the key components of e-DiaMoND and the dependencies between them.

e-DiaMoND components

The workstation package encompasses the imaging application and the acquisition application. These applications are not directly dependent on the grid. Rather, they depend on client interfaces; a package of Java interfaces classes that are an abstract, grid-independent definition of the system functions.

The e-DiaMoND client package is a concrete implementation of these interfaces that uses the grid.

The grid function is defined by a set of OGSA grid services depicted as the e-DiaMoND services package. Concrete implementations of the services defined there are split into three broad categories and packaged as acquisition, query and image retrieval. The acquisition package is concerned with inputting DICOM image data and Structured Reports into the grid. Its key component is a DICOM parser that allows non-image data to be extracted from DICOM files and added to a relational database. The query package consists of services that allow this structured data to be searched. In general, a query service will be a facade for an OGSA-DAI service that abstracts relational data as a grid data resource. The image retrieval package is concerned with function that allows identified DICOM image files to be securely retrieved from the grid.

External system components

The OGSA-DAI, Globus GT3, database and image store packages represent external components used in the pilot e-DiaMoND system. The following table lists the external components used by e-DiaMoND.

Component	Version	Comment
Globus Toolkit 3 (GT3)	3.2	OGSA Grid Services container
OGSA-DAI	4.0	OGSA Data Access Integration framework
IBM DB2	8.1, FixPack 5	The database package; also used by Content Manager Library Server
IBM Content Manager	8.2, FixPack 3	The image store package
IBM Visual Age C++ Professional for AIX	6.0	Prerequisite for Content Manager
IBM WebSphere Application Server	5.0, FixPack 2	IBM Content Manager prerequisite;
Apache TomCat	4.1.30	Servlet container

Registry

A registry is a directory of services. Globus GT3 provides two implementations of registries:

- a `ServiceContainerRegistry` that is a list of the services running in some specific grid service container, and
- a `VORegistry` that is a list of services belonging to some virtual organisation.

Each OGSA Grid Service Container runs an instance of `ServiceContainerRegistry`. Any service that is created within a particular service container is automatically registered in its `ServiceContainerRegistry`. The `VORegistry` is also configured to run in the default OGSA Grid Service Container set-up. It is empty when the container starts up and services must register with it explicitly. It is capable of registering services in containers other than the one in which itself runs. A grid service container is capable of supporting more than one `VORegistry`. A registry is a grid service in its own right; a registry may therefore contain other registries. Service data is associated with each of the services in a registry and this service data can be used to locate particular service instances. Multiple instances of the same service, with different service data, may be registered with the same registry, and the same service instance may be registered with more than one registry.

Where a service is persistent and re-entrant a single instance may be registered and used by multiple clients whereas when a service is transient and needs lifetime management the registry refers to a factory service that creates instances. A client either knows the Grid Service Handle (GSH) of a service or can discover it by searching against service data in the registry. A GSH is the minimum information a client needs to begin interacting with a grid service instance.

An e-DiaMoND client will have the GSH of a `VORegistry` supplied as a parameter. This will be a `VORegistry` at the local grid node in the peer grid defined for the first implementation, but could easily be a regional or central registry in other grid topologies. The services available in this registry define the function of the grid with respect to that particular client. As such, the e-DiaMoND pilot has four equivalent registries (one at each clinical evaluation site), with each registry containing instances of query, image acquisition, etc. services that operate on the local node.

Image store

An image store consists of two parts: a managed repository of DICOM image files and a relational database model of the patient data and image metadata that describes those files. These are referred to as 'image data' and 'non-image data', and are defined as follows.

- **Image data.** Each image store is implemented using an instance of IBM Content Manager. This consists of a Resource Manager component, that manages the DICOM image files, and a Library Server component that maintains a data model of the objects stored by the Resource Manager.
- **Non-image data.** The relational data model for non-image data is described in more detail in Section 2.8, and has been implemented using IBM DB2.

The virtual image store is the federation of all data (image and non-image) from all sites.

This means that the data resource is essentially local to the image server and the registry contains services that operate locally. Each node in the grid has an equivalent registry, with the result being that the nodes are capable of operating either entirely independently or collaboratively using the replicated data. In this mode the data grid forms though sharing data at the underlying database level rather than at the grid service level. It is possible to easily move to a grid configuration with regional or central image resources simply by modifying

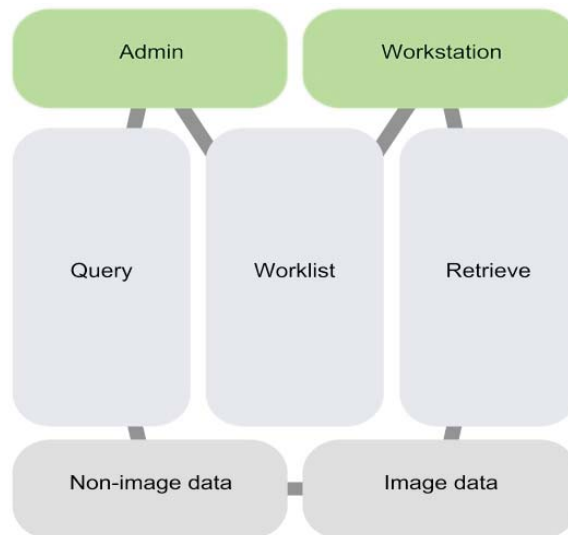


Figure 3: Major components with their interactions

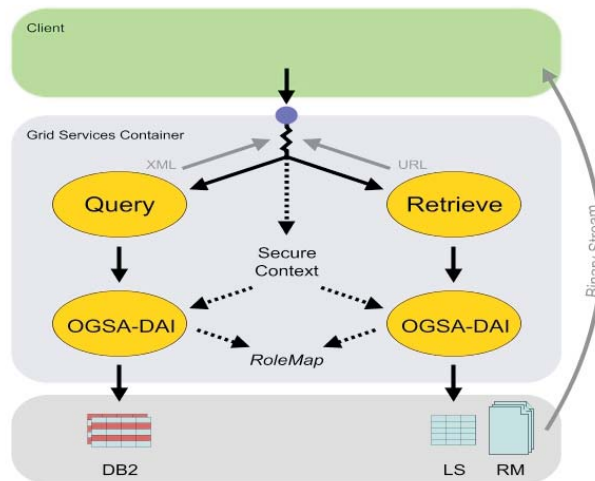


Figure 4: e-DiaMoND data flows

the contents of the registries at each node. In a centralised 'hub and spoke' model for example, the registry at the node would contain the same data service entries, discovered in exactly the same way by the clients, but referring to data resources held centrally rather than locally. In this mode, the data grid is formed through sharing data at the service level rather than the underlying database level.

2.6 Component interactions

Data flows for e-DiaMoND are defined in [38]. There are two main flows: the first is to query non-image data; the input is a query expressed in some form, and the output is an XML document that includes unique identifiers of DICOM objects; and the second is to retrieve DICOM files by ID. Both flows are linear and synchronous.

2.7 Security assumptions made for the pilot system

A number of simplifying security assumptions have been made for the pilot system. These are listed below.

1. References to DICOM image files are not sensitive and do not need to be secured. A reference to a DICOM image file must not allow direct access to the image file in the grid. The reference must be resolved to a URL by the image retrieval service to get access to image data. The image retrieval service must authorize every request it receives to resolve an image reference to a URL.

2. The grid hosting environment is responsible for authentication. Services within the hosting environment can check authentication in a secure manner. This assumption is in the spirit of the ongoing work on defining security architectures for grid services. It is true for the hosting environment chosen for e-DiaMoND grid services.
3. Each user of the system will have an X.509 certificate. Participating BCUs agree on a trusted Certificate Authority.
4. The distinguished name in an X.509 certificate is used to map individuals to a database user.
5. All nodes in the e-DiaMoND grid share the same authorisation model with respect to their individual policies, roles and authorisation rights.

Client enablement

There are three places in the system at which client interaction with the e-DiaMoND grid can be enabled.

- The workstation package is capable of extension. The e-DiaMoND workstation for training has been developed in this way.
- A client may be coded against the Java interface classes in the client interfaces package. The client will need the e-DiaMoND client package on the classpath.
- A client may interact directly with e-DiaMoND grid using OGSA grid services.

2.8 Database structure

e-DiaMoND stores both the raw, unprocessed images, and the images stored in SMF, together with associated non-image patient data.

Interactions with the e-DiaMoND database occur via grid services. These grid services are realised using the Globus Toolkit 3.2 [39], which was used to produce Open Grid Services Infrastructure(OGSI) [34] compliant grid services.

One such grid service is the query service, which allows users to query the database. Users are not permitted to query the database directly using SQL; instead, queries are sent in a pre-determined format as XML documents to the query service. Only allowing pre-determined queries affords numerous advantages, with the key ones being the hiding of low-level details from the end-user and the opportunity for static application-level query optimisation.

The query results are also returned as XML documents, with the data being represented in XML WebRowSet format. Being an XML document it is relatively straightforward to transform the results into any format that is required using XSL transformations.

As can be seen in Figure 5, the query service communicates with the database through an OGSA-DAI Grid Data Service [37]; this is also an OGSi compliant grid service built using the Globus Toolkit.

When the database is queried, the OGSA-DAI service ascertains which database user the user querying is mapped to. It achieves that by using the user's distinguished name that comes from an X.509 certificate and the mapping between that name and the database user is stored in the OGSA-DAI map file. Having determined the appropriate access permissions, the query is either rejected or translated into an appropriate SQL query. The SQL query is then placed inside an OGSA-DAI Request (Perform) Document and passed to the OGSA-DAI Data Service, which queries the database using a JDBC connection.

The result of the query is then returned by the OGSA-DAI Data Service as an XML document. The query service extracts the WebRowSet data from the document and passes the results back to the user.

The query service also ensures that the access attempts are logged correctly by inserting the user details into the logs; this allows several users to be mapped to a single database account while maintaining traceability.

Queries are made against a single logical database—although, in reality—this will be either a federation of different physical databases or a group of replicated databases.

While it is desirable to be able to handle all types of DICOM files, it is not practicable—with the resources available to us—to create a normalised database that could fully represent an arbitrary file. This problem can be easily understood if the system was required to deal with a new version of DICOM, which—perhaps—introduced new data fields or support for new image modalities. In effect, we are designing for forwards-compatibility. In addition, employing an unnormalised structure—with the potential performance benefits that may result—is inappropriate for an application that requires guaranteed consistency of data; data integrity is essential for an application such as e-DiaMoND. Despite the fact that it is undoubtedly the case that performance is important to us, delivering correct data in a longer time is more preferable than delivering incorrect data quickly.

For these reasons we divide the e-DiaMoND database logically into two parts, as illustrated in Fig.6.

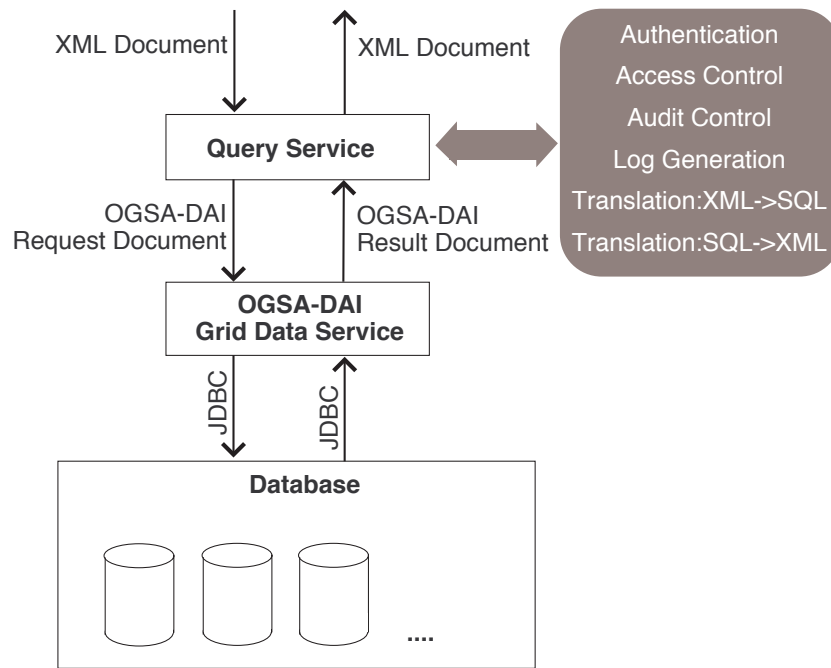


Figure 5: Grid Query Service Architecture

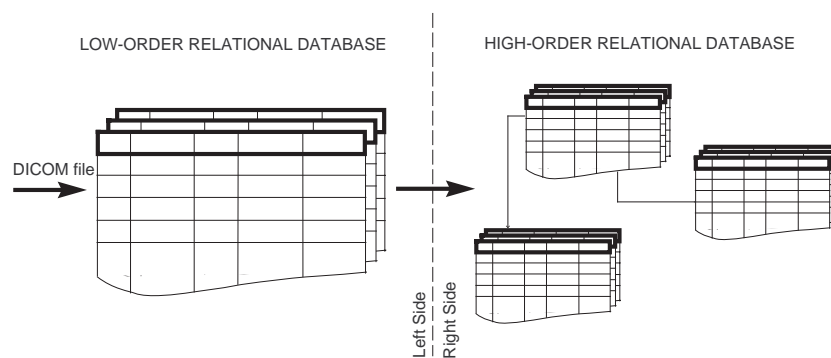


Figure 6: Database architecture

In the left-hand side—which we term *the repository*—the data is stored in a relatively unstructured fashion. In the right-hand side—which we term *the clinical information store*—data from specific IODs can be stored in a normalised fashion. When a DICOM file is inserted, it is parsed, with all non-image data being stored in the repository. Then, automatically, via automated constraint-enforcing procedures, the necessary data is inserted into the clinical information store. It is noted that in the repository, all the tags of the DICOM file are stored, including optional and private tags. In contrast, the clinical information store holds *only* the data that is currently useful for e-DiaMoND applications. If any tags were to be deemed relevant in future, it would be trivial to recreate the appropriate part of the clinical information store using the existing data stored in the repository.

In this architecture there are some rules that have to be obeyed.

1. The repository allows only INSERTs, but not UPDATES or DELETES; this ensures that there is no potential for data loss.
2. The clinical information store allows only INSERTs and UPDATES as a result of INSERTs to the repository; this ensures that the data in the clinical information store is consistent with that in the repository.

The schema for the underlying database is represented pictorially in Figure 7. (See [32] for further details.)

This schema has been designed in a way that allows an arbitrary set of DICOM files to be stored in the database, while ensuring that the relationships between Patients, Studies, Series and Equipment is maintained.

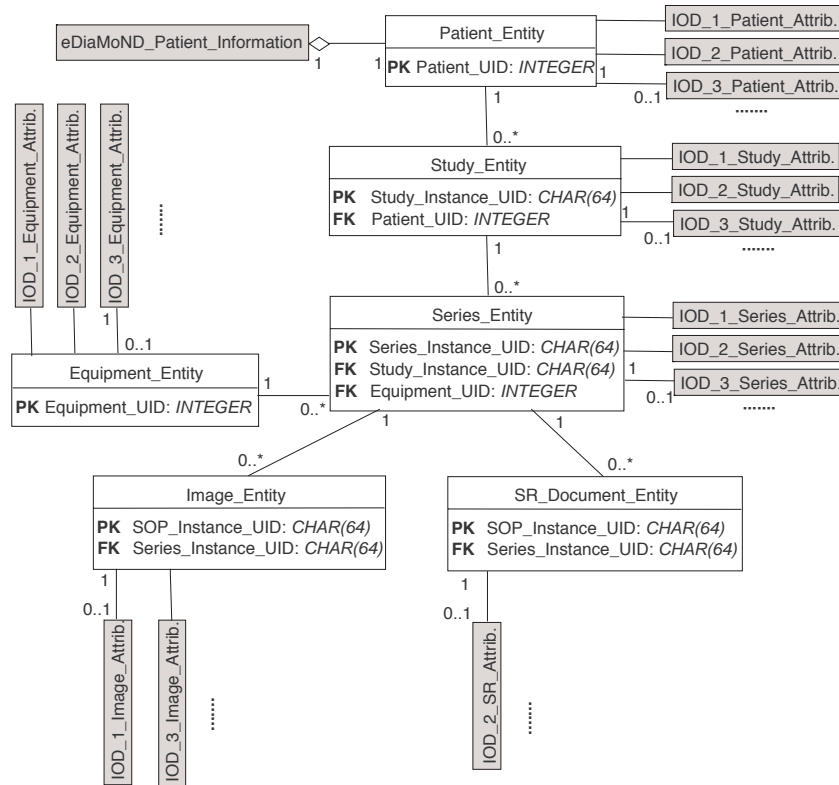


Figure 7: Schema to store DICOM data

The schema is logically divided into two parts: there is the common spine of DICOM entities and additional tables relating to specific IODs. In the diagram, three IODs are represented: *IOD_1* and *IOD_3* represent two different types of DICOM *image* IODs, while *IOD_2* represents a *Structured Report* IOD. While all three IODs have attributes related to Patient, Study, Series and Equipment, only the Image IODs have attributes related to the Image entity and only the Structured Report IOD has attributes related to the Document entity.

The modules of an IOD are all related to specific entities. Those modules contain attributes that may be present in more than one module, however the structure of DICOM is such that the same attribute will not be present in more than one entity. For this reason it makes more sense to group the attributes of an IOD by entity and not by module.

Although not shown in the diagram, the tables relating to each IOD are linked to the entity tables by a shared primary key, for example the *IOD_1_Image_Attrib* table has a primary key of *SOP_Instance_UID*, which has a foreign key constraint to the *SOP_Instance_UID* primary key of the *Image_Entity* table.

The *Patient_Entity* table is also linked to the *eDiaMoND_Patient_Information* table; this contains the additional patient data that was described previously, which is essential for epidemiology and integration with existing patient record systems.

The entity spine contains a bare minimum of data allowing efficient querying when using Unique Identifiers (UIDs). The *Patient_UID* and *Equipment_UID* are not part of the DICOM standard. The *Patient_UID* is represented as the *Patient_ID* attribute in the DICOM files, as an additional requirement on the users of the e-DiaMoND system it is essential that all *Patient_IDs* are unique. The *Equipment_UID* is stored in the *Device Serial Number* attribute, and is also unique.

One of the most critical issues when developing a medical database is the provision of appropriate mechanisms for allowing updates and tracking changes. This importance is derived from the legal and ethical requirements to record all updates of patient and screening data.

In the e-DiaMoND database, the deletion of previously captured DICOM files is forbidden. The principal reason for this is the necessity for keeping a history of previous data and cooperating—at the same time—with the existing health and legal regulations. In addition to this regulations, the e-DiaMoND database has the ability to support an audit trail mechanism for every update or insert that takes place.

When an update of information is needed, e.g., a change of name or a change of address, this will take place as a two-phase operation. The first phase involves the insertion of a new DICOM file that contains the updated—or corrected—data. The second phase—which is triggered by the successful completion of the first

BEFORE UPDATE					AFTER UPDATE				
Patient Table					Patient Table				
PK	Name		Parent Index	Audit Index	PK	Name		Parent Index	Audit Index
12	Jones Helen		NULL	879	12	Lloyd Helen		NULL	1043
					13	Jones Helen		12	879
Audit Table					Audit Table				
PK1	Res.Person	Date/Time	Reason	Place	PK1	Res.Person	Date/Time	Reason	Place
879	Dr. Miller	2002-2-7-5:23:03:786476	Insertion	John Radcliffe Hospital	879	Dr. Miller	2002-2-7-5:23:03:786476	Insertion	John Radcliffe Hospital
					1043	Dr. Baker	2003-12-20-8:20:32:700346	Marriage	Churchill Hospital

Figure 8: An example of the update mechanism

phase—involves a copy and an update.

This second phase of the update mechanism is described by the following example.

We consider a patient—Helen Jones—who has a unique ID: 12. Associated with this record is a related record—referenced by a foreign key called *Audit Index*—in another table, called *Audit Table*, which records the “who”, the “when”, and the “where” corresponding to the creation of a record in the *Patient Table*. If the patient were to get married, with a consequent change of surname, we would, of course, wish to update the patient’s name, with all other pertinent information remaining the same. To achieve this, the patient record in *Patient Table* has to be duplicated, resulting in a new record, with a primary key of 13. The information contained in the fields of the old record, having a primary key of 12, will remain the same, except the one that has to change, which in our case is the *Name* field, and also the field called *Audit Index*. The latter has to reference a new record in *Audit Table*, which will contain all the necessary information pertaining to the conditions under which the alteration has taken place. This new record is the one having primary key 1043 in the *Audit Table*.

In the above example we have used one *Audit Table* for a given table in the database. However, one *Audit Table* can be used to track the changes of many tables. The concept remains the same: the only modification is the addition of another field in the *Audit Table* containing a reference to the table that has been changed.

2.9 Benefits

e-DiaMoND allows individual nodes to store and manage mammograms as digital images. It further allows those nodes to share these mammography archives, allows radiologists to collaborate on diagnosis without being in the same physical location, and provides the potential to even out workload on radiologists by distributing reading across different nodes.

The use of SMF within e-DiaMoND makes it possible to compare images. Combining this technology with a significantly large repository of x-ray mammograms gives potential for developing useful new techniques for medical image processing and data mining. The manifestation of e-DiaMoND as a grid facilitates the addition of further applications represented as services.

Part II: The Present

3 Legal and ethical constraints

In [20], it is stated that “the e-DiaMoND project would be deemed a success as far as legal and ethical approval was concerned if the Information Commissions (Data Protection) and the relevant ethical approval bodies accepted the project as fully conforming to all processing and security needs to satisfy their levels of acceptance.” Such considerations have been key to the deployment of the e-DiaMoND pilot system and its associated virtual database of anonymised data.

A complex web of policies, acts and organisations governs the use of information originating from patient or case information. During the initial phase of the e-DiaMoND project, the team had to ensure that their use of anonymised information from the clinics involved in the project satisfied the requirements under the Data Protection Act; the project team also had to obtain clearance from the Thames Valley Ethics Committee. The situation was exacerbated by the need to use not only newly collected case data with explicit consent from the patients, but also archives of cases which had been selected by hospitals for training purposes and for which it would have been impossible to seek retrospective consent from those patients to use the data for the e-DiaMoND project.

The process of gaining ethical approval for e-DiaMoND took in excess of 18 months; our experiences suggest that the early consideration of such constraints is imperative.

3.1 The legal structure of the NHS

Before exploring the relevant legislation, it is important to understand the complexities of the legal structure of the UK Health Service. The National Health Service (NHS) came into being on the 5th July 1948 to provide healthcare for all citizens, based on need, not the ability to pay [40].

The NHS is funded by the taxpayer and managed by the Department of Health, which sets overall policy on health issues—it is the responsibility of the Department of Health to provide health services to the general public through the NHS.

The NHS was launched as a single organisation based around 14 regional hospital boards and was originally split into three parts:

- hospital services;
- family doctors, dentists, opticians and pharmacists; and
- local authority health services, including community nursing and health visiting.

Since 1948 there have been huge changes to both the organisational structure of the NHS and the way that patient services are provided, including an initiative in 2001 to give greater power to local trusts and front-line staff. This initiative, called ‘Shifting the Balance of Power’, was designed to offer a patient service that was faster, more convenient and offered more choice for patients. It launched the NHS Modernisation Agency as the lead organisation in reforming the way the NHS works. The aim was to design a service that puts both patients and staff at the heart of the NHS, and this has meant abolishing the previous health authorities and creating new ones that serve larger areas and have a more strategic role.

The NHS in the UK employs about a 1 000 000 people, which equates to approximately 5% of the working population. These employees are distributed across the UK and in clinics, hospitals and surgeries.

It should be noted that the NHS is not a single legal entity, but an amalgamation of over 400 legal entities. These legal entities comprise clinics, surgeries, governing bodies, trusts and regional centres which together provide the health service in the UK. The structure of these legal entities differs in each of the UK countries and within regions.

Below we summarise the current structure of the NHS within the different regions of the UK.

England

The Department of Health supports the government to improve the health and well-being of the population. The Modernisation Agency supports NHS clinicians and managers in their efforts to deliver improvements to their services. Special Health Authorities (SHAs) provide a health service to the whole of England, not just to a local community. An example of this is the National Blood Authority. Strategic Health Authorities manage the NHS locally and are a key link between the Department of Health and the NHS. Within each SHA, the NHS is split into various types of trusts, such as primary care trusts, ambulance trusts, mental health trusts, etc., that take responsibility for running the different NHS services in that local area [41].

As part of its programme to modernise and reform the NHS, the Government set up National Service Frameworks (NSFs) in order to improve patient care and reduce inequalities in a series of identified priority areas. The

NSFs work towards this by setting national standards and putting in place strategies to support the development and improvement of services within the areas such as cancer, paediatric intensive care, mental health, etc.

More details on the NHS in England can be found in [41].

Wales

There are 22 local health boards associated with the 22 local authorities, which assess health services their populations need and then pay hospital trusts, family doctors, dentists, and so on to provide those services. The CHC (Community Health Council) in each of the 22 local government areas in Wales take up a wide range of health issues on behalf of the public. There are 14 NHS Trusts in Wales, including one all-Wales ambulance trust. Between them, the Trusts manage 135 hospitals.

More details on the NHS in Wales can be found in [42].

Scotland

Scotland comprises 15 NHS boards and 28 NHS Trusts.

On the Scottish mainland, there are two main kinds of NHS Trust providing health care to communities—primary care trusts (PCTs) and acute hospital trusts (AHTs)—both of which report to their local NHS Boards.

More details on the NHS in Scotland can be found in [43].

Northern Ireland

There are currently four area Health Boards in Northern Ireland (east, west, north and south) and 40 hospitals, 10 acute and 30 local. Within these boards, NHS Northern Ireland utilises the services of local health and social care trusts but is also complicated by the definition of regional health councils and health agencies.

More details on the NHS in Northern Ireland can be found in [44].

Summary

The above section highlights the complex legal structure of the NHS. While we consider the NHS to be a service provided to the UK population for free, the segregation of this service into England, Scotland, Wales and Northern Ireland results in an near impossible domain in which to consider joined-up IT solutions. The National Programme for IT (NPFIT) [21] for the NHS, in fact only covers England, and other countries are instigating their own programmes of IT infrastructure improvement.

Considering the impact of this diverse approach, it is apparent that unless these programmes of activity consider interoperability, patient mobility will not be as seamless as first envisaged with the implementation of a digital environment.

If we look also at the issues of data ownership and management, this complex array of legal entities, all collecting and storing information about patients—as well as having to make this information available to other legal entities to provide healthcare—need to have data controllers in each of these legal entities whose role it is to ensure that data about patients is never misused or released into the wrong hands. Each legal entity may have its own records management policies detailing the local understanding of the legal and ethical constraints of data usage and the policies to be adopted by the clinical and administrative staff within that department or institution. Senior managers and Chief Executives are personally accountable for the records management of their organisation. The policies for records management stem from the legal constraints as dictated by UK Law and by best practice as guided by the ethics committees.

We now consider the legal and ethical constraints which drive the policies for patient records management.

3.2 Legal constraints

Legal concerns can be derived from a number of different areas. We address each in turn.

The Data Protection Act

The Data Protection Act of 1998 [45] provides a framework for the processing of *personal data* on living subjects on all forms of media. According to the Data Protection Act, the more stringent requirements of the act do not apply to anonymised unlinked data typically used in research.

The main provisions of the Data Protection Act state that personal data, in written or electronic form, must be

- fairly and lawfully processed;
- processed for limited purposes;
- adequate, relevant and not excessive for the purpose;

- accurate;
- kept no longer than necessary;
- processed in accordance with data subjects' rights;
- kept secure, and only transferred to countries with adequate data protection systems.

The following definitions are relevant within this context:

- *Personal information* is defined as all information about individuals, living or dead, e.g., medical records which are written or held on a computer system including images, recordings and medical opinions about the individual.
- *Personal data* is information about living people, which, in isolation or in combination with other data, may lead to the identification of the patient.
- *Confidential information* is information provided explicitly or implicitly on the understanding that it will not be disclosed to others outside the patient's care or without the patient's consent.
- *Sensitive information* is information about individuals which may have severe consequences if disclosed inappropriately. The Data Protection Act refers to it as 'sensitive personal data' and includes information about physical or mental health or a patient's sexual life.

Where personal data is used for research purposes, different techniques are used to ensure that the patient is protected from misuse of information. These techniques include:

- using coded data, where the patient identity is disguised but can be easily decoded by those in control of the data;
- anonymisation, where data has had all means of identity removed by clinical staff before being released for research purposes; and
- link anonymised where the data has identifying features removed but can be decoded again by the organisation supplying the data to the researchers.

Within each legal entity, in this case a hospital trust, a Caldicott Guardian [46] is appointed whose role is to ensure that data is processed correctly and that the principals of the data protection act are respected and acted upon. This process involves ensuring that any staff who require access to un-anonymised patient data for their research or for processing to create anonymised data are either NHS employees of that legal entity or are put onto honorary NHS contracts to ensure that they apply the same duty of care as NHS clinicians and staff.

Within e-DiamoND, this has applied to resources who are assisting with the data acquisition in the clinics as well as those who may come into contact with un-anonymised patient records through attendance at clinics for ethnographic studies. This process requires the project to have full ethical clearance before proceeding and will require staff on these contracts to provide a recent curriculum vitae for hospital staff.

The principles of the Caldicott Guardian, are stated below [46].

1. Justify the purpose(s): every proposed use or transfer of patient-identifiable information within or from an organisation should be clearly defined and scrutinised, with continuing uses regularly reviewed by an appropriate guardian.
2. Don't use patient-identifiable information unless it is absolutely necessary: patient-identifiable information items should not be used unless there is no alternative.
3. Use the minimum necessary patient-identifiable information: where use of patient-identifiable information is considered to be essential, each individual item of information should be justified with the aim of reducing identifiability.
4. Access to patient-identifiable information should be on a strict need-to-know basis: only those individuals who need access to patient-identifiable information should have access to it, and they should only have access to the information items that they need to see.
5. Everyone should be aware of their responsibilities: action should be taken to ensure that those handling patient-identifiable information—both clinical and non-clinical staff—are aware of their responsibilities and obligations to respect patient confidentiality.

6. Understand and comply with the law: every use of patient-identifiable information must be lawful; someone in each organisation should be responsible for ensuring that the organisation complies with legal requirements.

The deployment of a system into the NHS for real patient care stresses the needs of complex security to protect the data stored in the virtual archive. The data stored would be raw data with full identifying features. The movement of such information would be performed in the interests of direct care and should be in control of the clinicians, e.g., seeking assistance from a remote radiologist. The system should allow users to choose who should see the data, or elements of the data, and where data should move to e.g., if a patient moves clinics, the process of handing over records and taking receipt of them should be a controlled process. Each trust may have differing views of who should have control of information and its movement. A deployed system should be sympathetic to these local needs and understand that data ownership resides with a local trust and should never be moved or processes without the prior permission of that trust.

If data were to be moved outside of the United Kingdom, to potentially make use of resources from other countries, e.g., to buy the assistance of resources from the US where the UK hospitals have problems recruiting resources, there would need to be sufficient protection in that country from similar data protection rules to safeguard personal data, particularly as the data would likely be in raw format. Where possible, it would be desirable to ensure that only where it was essential to include patient identifiable information, should it be transmitted outside of the trust or outside the UK.

An area not yet explored is the need for data controllers if the NHS considers itself as a joined up entity through the use of grid technology to share information as required between clinics. At present each legal entity would have a data controller but would there need to be a legal entity and a data controller with sight of the needs of movement of data between clinics, or the storage of any information about a clinic held at a level outside of that clinic.

Common law

It has been established via Common Law that:

“information that has been confided must not be used or disclosed further *except as originally understood by the confider or with subsequent permission.*”

There are caveats: for example, this principle can be breached ‘in the public interest’ (such cases are examined on a case-by-case basis in civil courts).

The Human Rights Act

The Human Rights Act of 1998 [47] establishes the following.

- A right to life.
- A right not to be subjected to degrading or inhuman treatment.
- A right to liberty.
- A right to a fair trial.
- A right to ‘respect for private and family life’.
- A right to the enjoyment of rights and freedoms without discrimination.

The right to ‘respect for private and family life’ details that:

“there shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.”

3.3 Ownership of data

Ownership of data is an area where, in the UK, there are currently no clear guidelines. Whilst a patient’s medical record is a dispersed entity residing in many locations rather than in a single computerised or paper based archive, much of the information in the records is obsolete, redundant, duplicated, or indecipherable to the extent that it sometimes fails to benefit the patient at the point of care.

Many hospitals consider the records in their systems to be their property, whereas many patients argue that their medical information is their own. Consequently, a distinction is made between ownership of the physical record and the right to access (or duplicate) data that are stored in it. Policies on this issue differ substantially between delivery networks, regions, and countries. That said, it is typically agreed that patients have the right to be informed of the general content of their medical record and that patients' care providers must be allowed access to any information that is relevant to a patient's treatment. This approach endorses locking sensitive information (such as psychiatric evaluation or various serological findings) from some care providers but promoting access to what is 'needed to know' for the provision of appropriate care. It is thus reasonable to assume that, between the patient and his or her primary healthcare coordinator (such as the family doctor), most of the 'critical information' is within reach.

Database rights (which are a right akin to copyright, but arise under the EU Database Directive) may subsist in databases of patient, clinical or other data which are compiled and maintained in electronic form. Given this link with patient records and other clinical information, careful thought needs to be given as to where the intellectual property rights in such data and the resulting database vests, particularly as the ownership of copyright and database rights does not arise in the same way.

Considering projects like e-DiaMoND, there is no indication whether database rights may be claimed for valuable research archives as created here. The data digitised, anonymised and collated has been provided by clinics who are unsure themselves whether they have rights in the data but have assumed that they do and may 'license' it for use in research, supported occasionally by basic legal agreements. Whether these rights extend to being able to receive payment for such use of this data if it were ever to be used for something other than research e.g., key datasets for training applications, are questionable.

3.4 The complexities of obtaining ethical clearance for research

The ethical clearance required for the use of genuine patient data for research purposes requires individual projects to seek clearance either from a local ethics committee for research involving just a local site, or to a multi-site ethics committee for clearance to use data across many sites. Within the United Kingdom, gaining ethical clearance for the use of data in research projects is a complex task.

At present, the process for obtaining ethical clearance is a time-consuming and prohibitive process. Projects have taken a variety of approaches in dealing with patient originated data, from obtaining revocable consent for every patient case used, even for anonymised data, to a 'best endeavour' approach as made by the Biobank project [48]. For the patient, if case information has been truly anonymised for research, there is no way this data may be later removed from the research databases as there would be no way of identifying the patient records. This approach often results in data being used for specific research, and subsequent use of this data may require further clearance from ethics committees. Often the consent process requires the patients to sign up to a specific project, rather than any future possible use of it. Improvements to the ethical clearance process may allow patients to approve the use of data for any research under conditions. Where there is a need to be able to remove records if a patient requests this, a link has to be kept in the clinic to enable the clinician to provide ID information to the researchers to inform them of the request. This process protects the individuals but does also enable the revoking of rights of access to be respected.

Recent announcements by the Medical Research Council have indicated a move towards reviewing the process of clearance for using medical data for research and a work package has been defined to look at the process for enabling generic consent to be granted by patients for their anonymised data to be used for research purposes.

The e-DiamoND project was required to apply for multi-site ethical clearance as well as local ethical clearance. This enabled the project team to take data from the selected clinics in anonymised form and keep it for ten years for the purpose of the original ethics application. Any further use of this data would need further ethical clearance.

It should be noted that it is the duty of the principal or lead investigator to ensure the appropriate archiving of the data once the research has completed, and to ensure that data is kept securely at all times. For many patients, the project will have had to seek explicit consent to use the data. The process of anonymisation means that we will be unable to remove records from the virtual database data once all identifying features have been removed.

A future deployment of e-DiaMoND or a grid healthcare solution supporting clinical practice, would not need to apply for ethical clearance as the data is used in the interest of direct patient care. Only where data were to be released for research purposes would ethical clearance be required and potentially patient consent sought. A tiered approach to data management may be appropriate as shown in Figure 9.

These issues also raise the concern over the use and management of medical research data in general. It is often the case that valuable data is generated out of innovative research, but the issues of ownership and what researchers are allowed to do with this data are perplexing, as detailed above. With current constraints on

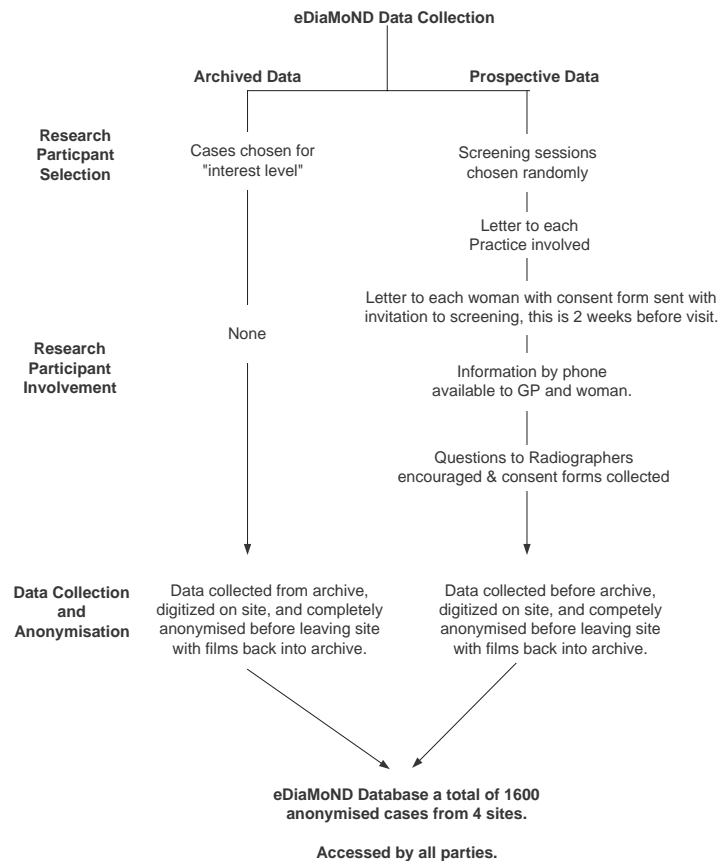


Figure 9: Data management in e-DiaMoND

applying for reuse of data and the need to request specific use to use previously consented data for follow-on research projects, it is envisaged that a significant amount of valuable medical data fails to be reused. The Medical Research Council are currently embarking on a programme of research to cover this area, but it is believed that a framework of ownership needs to be developed to support those utilising clinical research data. This may include rules on who the data controller is, what their responsibilities are, rules on commercial exploitability of that data, and the rules on how that data may be used beyond the original requirements for that data. In addition to a framework to advise the researchers on their rights, it is evident that the persistence of that data relies on the continuing employment of the project staff to ensure this happens. A clear process and management infrastructure to support the long term storage of such data would provide an essential resource to the academic and commercial communities across the world. Clearly a move towards 'generic consent to use my anonymised data for research and training purposes' would benefit the proliferation of such data and consequently research that relies on it.

Occasionally, it would be unreasonable to be able to seek explicit consent to use patient data for research. These cases are rare but require special treatment to ensure that such research is not prohibited. In 2001, a new statutory body, the Patient Information Advisory Group (PIAG), was created giving the Secretary of State the power to exempt specific research projects from the GMC's guidance. Section 60 of the Health and Social Care Act 2001 provides a power to ensure that patient identifiable information needed to support essential NHS activity can be used without the consent of patients. The power can only be used to support medical purposes that are in the interests of patients or the wider public, where consent is not a practicable alternative and where anonymised information will not suffice. It is intended largely as a transitional measure whilst consent or anonymisation procedures are developed, and this is reinforced by the need to review each use of the power annually. Additionally, PIAG will be disbanded once NPfIT is completed.

It is important to note not only the time-scales required to achieve PIAG clearance which are estimated to be approximately 18 months, but also that the permission given is temporary and may be removed at any time.

In this section we have looked at the complexities of the legal and ethical constraints of working with patient data both in the NHS and for medical research. The use of grid technology opens up opportunities for collaborative working both within the UK and across the world. This, of course, requires an understanding of

the legal and ethical considerations in other countries. This makes collaborative working more complex.

3.5 The social issues of enabling digital patient data

Whilst it is important to ensure adherence to all legal and ethical guidelines, it is also important to understand public perception and concern over the use of electronic patient information. In [49], the authors indicate the wide reaching concerns over the digital enablement of information and the fundamental requirements of keeping secure all aspects of information, and, in particular, ensuring only those required to see data in order to perform their role are allowed to see it. There is an element of mistrust of administrative staff, and also clear indication that the public wish to consent to the use of their information and to exercise their rights in determining what information is stored about them. Additionally, the rights of researchers have also been questioned.

Clearly, education with respect to how data is managed and guarantees of appropriate provisions of security and confidentiality mechanisms will be essentially in gaining the greater confidence from those whose electronic patient data may be used for research purposes.

3.6 Summary of requirements and process for real system deployment

Clearly a move to a digital health age needs consider a wider range of issues beyond the technology. Efforts in providing 'joined up healthcare' and 'joined up medical research'—in an environment in which we have yet to overcome questions about who owns the information we are working with and whose responsibility it is to ensure the relevant consent is obtained—will need to go hand-in-hand with efforts to develop a robust legal and ethical framework.

4 External dependencies

This section considers external dependencies and integration issues for UK-based health grids. Such dependencies arise out of requirements that extend beyond the boundaries and responsibilities of the system components being developed: those systems, for example, with which a national grid for breast screening and its local manifestations would have interfaces; the standards, initiatives, policies and laws required for, or affecting, integration, their status and future development; and any software or middleware required to facilitate the composition of the system.

Although many of the issues exposed will be generic, this section will address the general technical dependencies pertaining to a project to develop federated digital mammography support for the NHS Breast Screening Programme.

4.1 Standards, initiatives and policies

The e-GIF Standards Framework

The foundations of Government (and thus NHS) policy for informatics is the e-Government Interoperability Framework (e-GIF), which describes policies and standards for interoperability across the entire public sector. Its requirements include:

- the use of the browser as the key user interface,
- the adherence to open, international, Internet and Web-based standards across all public sector systems,
- the selection of XML as the key interface for access to, and manipulation of, information, and
- the recording of metadata, based upon the Dublin Core set [50], about the information content.

e-GIF is mandatory for exchanges of information and interactions between NHS organisations, between NHS and other organisations, and between the NHS and the citizen [51].

e-GIF is maintained by the Office of the e-Envoy and published in two parts. Part 1 contains the policy statement. Part 2 is the Technical Standards Catalogue, detailing standards covering interconnectivity, data integration, metadata, access and business areas. The document has taken recommendations from the NHSIA and establishes three healthcare standards: HL7 Version 3, the NHS Data Dictionary and SNOMED CT. The Department of Health (DoH) statement on e-GIF compliance is presented in [52].

e-GIF compliance is enforced through the Standards Enforcement in Procurement (STEP, currently at version 10 [53]). DICOM is endorsed in e-GIF as a standard for communications in medical imaging [54]. DICOM is discussed in more detail in Section 4.1.

STEP Version 10

STEP [53], which stands for Standards Enforcement in Procurement specifies standards and policies for procurement under the National Programme for IT (NPfIT [21]) and consists of two main parts:

- NHS IT standards and procurement policies: these state the NHS policies with regard to a number of widely-used IT standards and show how they are applied in procurement, and
- the STEP questionnaire: when used in IT procurements, this helps ensure that the goods and/or services procured conform to the appropriate IT standards.

STEP Version 10 is fully aligned with the National Programme for IT (NPfIT). STEP has now been expanded from the use of standards in NHS IT procurements to cover general strategy and policies for the use of IT standards. These policies are essential reading for those committed to develop systems (directly or indirectly) for NPfIT.

STEP approves HL7 Version 3 as the communications standard applicable for communications between organisations [55], particularly in communication with the NCRS, and indicates that specific messages will be approved. HL7 is discussed in more detail in Section 4.1.

SNOMED CT

SNOMED Clinical Terms (CT) [56] [57] is an international ontology that covers the whole of medicine with almost one million English terms, describing over 300 000 distinct concepts and 1.4 million relationships between them. The ontology was developed by combining The College of American Pathologists (CAP) SNOMED RT with the National Health Service Information Authority (NHSIA) Clinical Terms Version 3 (Read Codes).

NHS Data Dictionary

Having defined standards for patient demographics, for the terms, structure and presentation of health and healthcare related information, and for message exchange at many levels, the NHS Data Dictionary [58] describes how these standards are implemented within the NHS and provides additional, local standards required for the operation of the national service. These standards include support for the internal market, comparative data analysis, and the preparation of performance tables and reports for the Department of Health.

Thus the data dictionary contains a wide and varied collection of standards, from the Commissioning Data Standards, which support the communication of activities between healthcare providers and commissioners, through the National Cancer Waiting Times Monitoring Data Set, which supports the central electronic collection of patient level information to monitor waiting times for comparison with targets set by the National Cancer Plan, to the National Minimum Datasets for Cancer, which describe the SNOMED CT coded information that should be collected for all cancer patients.

BS7799

The British (International) Standard BS7799-1 [59](ISO 17799) is an established code of practice for information security management. It lists a comprehensive set of controls, comprising best practice in information security. It advocates the systematic determination of risk assessment, and the application of a risk control programme at all stages of development and operation.

Much of the code of practice pertains to controls implemented outside the information system: asset classification; operating procedures; management policies; employee education; physical and environmental security; recruitment, and other personnel issues. However, the design of the information system must support, and be sympathetic to, whatever controls are adopted.

A second document, BS7799-2, specifies how this code of practice should be applied, in conjunction with ISO 9001 (a standard for quality management systems) and ISO Guide 73 (a glossary for risk management). This specification takes the form of a six-page list of checks and requirements, followed by a ten-page summary of items arising from BS7799-1.

These documents contain many generalities and qualified statements: 'audits shall be planned carefully'; 'access shall be protected'; '*relevant* requirements shall be defined'; '*appropriate* procedures shall be implemented'; etc. The overall effect is to require only the existence of procedures, processes, and documentation: it is up to the organisation to establish procedures and processes that are appropriate and effective. Thus compliance is no absolute measure of the security of a system, but a reflection that the owners of the system understand the balance between risk and cost and hopefully have achieved a good balance between these factors. However, given the ease by which information can currently be obtained [60], simple controls outside of the information system clearly need to be established.

The most important elements of BS7799, implied by its links to ISO9000, are continual process improvement and accreditation. Organisations committed to BS7799 will manage security in an environment of continual process improvement, and thus the security of their systems will improve and react to new and changed threats. Accreditation provides an additional, external guarantee that the organisation is meeting its stated goals. However, where data is as private and sensitive as in health care, it seems essential that some minimum, absolute measurement of information security is also required. The work of Ross Anderson for the BMS (see, for example, [61] and [62]) is of relevance in this regard.

The NHS has now adopted BS7799-1 as part of its security guidance. It is also mandated by the Patient Information Advisory Group (PIAG) for use in the design of information systems that support clinical studies. Even if this were not the case, the fact that BS7799-1 and BS7799-2 represents good practice in information security makes it an essential aspect of technical developments in this domain.

NHS Strategy for Cryptographic Support, DTS, SUS and eSMTP

A related development is the NHS Strategy for Cryptographic Support Services [63], which pertains to the application of digital signatures to:

- authenticate electronic records and their authors,
- provide proof of a user's entitlement to use electronic information, and
- sign and seal electronic workflow documents.

In addition, the strategy concerns itself with the establishment of non-repudiation services to provide proof of origin, delivery and receipt.

The NHS is apparently committed to a national Public Key Infrastructure (PKI) for all employees with responsibility for electronic data, allowing both secure communication and non-repudiation: messages can be

encrypted using the public key of the recipient; documents or changes can be signed with the private key of the author (their provenance can then be confirmed by anyone with access to the public key).

There is a happy coincidence of specification in this area: an appropriate PKI infrastructure is an important element of grid infrastructures, it is specified in IHE use cases, and it is NHS policy. Thus it seems obvious that a national grid to support, for example, breast screening, together with other similar developments, would look forward to utilising an existing PKI infrastructure.

Unfortunately, however, the implementation of this policy has floundered. A successful implementation of PKI had allowed pathology laboratories to send encrypted, signed pathology reports to GP practices [64]. However, due to difficulties with the supplier [65], the service was closed on the 31st July 2004 [66] to be replaced by unencrypted communications through SSL [67] via the NHS Data Transfer Service (DTS).

Cryptographic standards are still listed as current by the NHS [68] and given the programme to issue NHS staff with smart cards, it seems that PKI is set to return in the future.

NHS Data Transfer Service (DTS)

The NHS Data Transfer Service (DTS) [69] is a strategic solution that provides secure Electronic Data Interchange (EDI) messaging facilities. The DTS is a centralised service for the transfer of application messages in a variety of syntaxes including EDIFACT, XML and Flat File. It is available over NHSnet only. The DTS is made up of two elements:

- the data transfer clients, which are part of the applications at the end sites, and
- the central data transfer server.

Clients transfer encrypted data to the central server. The central server then either downloads the data over an encrypted link to another client or transfers the data to the eSMTP messaging service, which forwards the data as an e-mail message. An X.400/eSMTP gateway is provided for exchanges with organisations, such as external trading partners, which continue to run EDI over X.400. The service also offers reverse IP-Domain lookup and 'strong authentication'. The DTS in itself is merely a transitional system for extended SMTP (eSMTP). Detailed information on all aspects of DTS is available from the DTS pages on the NHSIA web site [70].

The DTS is to be used for EDI previously undertaken over the X.400 messaging service, which was withdrawn at the end of 2003. The principal workflows supported by the DTS are for the following:

- NHAIS (Exeter Systems)(see section 4.2),
- NHS-Wide Clearing Service (NWCS),
- National Strategic Tracing Service (NSTS), and
- Pathology Messaging Implementation Project (PMIP).

The objective of standardisation is to ensure that all NHS organisations are able to perform application messaging with other organisations, both within the NHS and external trading partners, previously achieved using the X.400 messaging service. Standardisation also means that new or additional workflows could be supported if required. In addition, the DTS complies with e-GIF unlike the X.400 service.

NHS Secondary Uses Service (SUS)

The vision for the Secondary Uses Service (SUS) is recognised as ambitious and it will take time to design, deliver and implement. The SUS will not be achieved in one go, and so will need to be delivered as a number of smaller managed iterations.

The proposed scope of SUS is outlined in the National Programme for Information Technology (NPfIT) document, "Secondary Uses Service — Scoping Report". SUS phasing assumptions start from the premise that certain components of the SUS are non-negotiable, e.g., the explicit deliverables in the National Application Service Provider (NASP) contract schedules, whilst other service elements are still sufficiently embryonic as to require further design, consultation and refinement.

The initial priority for the SUS is to establish a robust and flexible infrastructure that can take forward responsibility for a national clearing service and achieve the following early goals.

- Creation of a robust technical infrastructure fully integrated into the NHS SPINE.
- Creation of a SUS logical data infrastructure based on an NHS unified data model.
- Improved security and access control via the SPINE Role-Based Access facility.
- Smooth migration from the existing NHS Wide Clearing Service (NWCS) to the new SUS clearing function.

- Support for Payment by Results (PbR) via the provision of core national functionality.
- Establishment of a national Pseudonymisation Service.
- Creation of interfaces to the Personal Demographics Service (PDS), Spine Directory Services (SDS) and the Transaction Messaging Service (TMS).
- Generic Data Quality Reporting Service available for all SUS users.
- Facilities to perform online queries or to extract data from SUS.
- Access to a range of new data items derived automatically from incoming data.

The initial release of the SUS replaces the existing NWCS services with NHSCRS-compliant solutions. In the first instance, these will run side-by-side with existing clearing service modules, which will enable a managed transition from existing standards (EDIFACT, UDF, DTS) to the new standards; XML (eXtensible Markup Language) and the TMS (Transaction Messaging Service). SUS release 2a is aiming to complete the transition from DTS and EDIFACT messaging to TMS and XML messaging. Data will be received by the SPINE and initially held in a secure database. From here, it is imported into the SUS as pseudonymised records where it is processed and made accessible via on-line reporting functions. During 2005, it is anticipated that the first series of Local Service Provider (LSP) solutions will begin coming on stream.

HL7

Health Level 7 (HL7) is a (not-for-profit) company formed by interested parties to harbour the development of standards for communication between differing hospital systems. It is intended as a comprehensive framework for defining data formats for exchange of information between systems, and this remit has naturally been extended to include formats for the representation of certain kinds of information.

Introduced in the late 1980s, HL7 Version 2 is a messaging standard that defines functions within the clinical enterprise and sets standards for the messaging between them. Version 2 messages are organised into segments, a MSH—message header—segment accompanies every HL7 message, communicating information such as the version of the HL7 standard being used, the type of message being sent, and the encoding characters used for field and segment separators. Other segments include the PID—patient identification—segment: segments are separated by the vertical bar and terminated with a carriage return.

HL7 Version 2 has been used to obtain a significant degree of integration of systems, particularly when consensus IHE implementations have been used. HL7 Version 3 is more of a toolkit than a standard: it can never achieved plug-and-play interoperability because there is no framework for semantics, and its flat structure makes it difficult to communicate text-based documents, limiting its application in the exchange of health records. Thus all integration is bespoke and, as a consequence, expensive.

Initial development of HL7 Version 3 began in 1991–2 with the revised process adopted in 1997. Version 3 uses a meta-modelling approach, allowing compliant systems some ability to comprehend messages and take appropriate action: the meta model—the RIM (Reference Information Model)—is the basis for the derivation of specific *message information models*—MIMs—for application areas appropriate to healthcare. Training and tools for the capture of usecases and workflows in UML have been provided from the outset, and there is little scope for optionality.

Unfortunately, the small number of core, abstract classes within the RIM do not form an appropriate basis for the language: the RIM is too abstract to be unambiguously interpreted by the working groups and committees of domain experts delegated to the task of developing message information models, yet not abstract enough to justify its designation as the seventh OSI level in healthcare communication. Design decisions taken by one committee have turned out to be inconsistent with those of another, and thus opportunities for re-use and abstraction have been lost. Additionally, design process is very slow: it may take 18 months for a new message class to be agreed, and the results may not prove satisfactory in application. As such, after more than a decade of work, HL7 Version 3 still only has draft status.

Of course, the HL7 community is aware of these problems, and is trying to deal with them.

A further problem from an NHS perspective, is the influence of its healthcare economic model on the development. A strong element of HL7 (and of DICOM and IHE in turn) is upon inter-enterprise charging and administration: working groups for health record and clinical data representation [71] are recent additions to the HL7 fold. Thus messages have been defined from this financial rather than clinical perspective, complicating implementation in the UK where the economic model is different.

The decision by the NHS to use this draft standard for inter-enterprise communications has risks, but with no standards based alternative, and given the NHS' financial resources and the installed base, it can shoulder some of the HL7 development process where it has a particular interest, such as GP-GP communication and the

National Care Records Service. In local implementations, where small budgets and tight timescales operate, version 2 through IHE is likely to predominate.

HL7 houses a number of special interest groups of interest to healthcare grids:

- **CCOW**, the Clinical Context Object Workgroup is developing standards for unifying the context of clinician interaction with health care systems. This allows a clinician to establish a secure context—at the level of the patient or a particular patient’s encounter with the health care system—for subsequent queries or operations.
- **Arden Syntax** supports the clinician to create and disseminate computerised *Medical Logic Modules (MLM)* among personnel, information systems, and institutions. Each module contains sufficient knowledge to make a single decision: typical module types could include contraindication alerts, disease management suggestions, data interpretations, treatment protocols, and diagnosis scores. Each MLM also includes sufficient infrastructure to support automatic linkage and assembly into a knowledge base.
- **Imaging Integration** ensures that the classes and attributes necessary to allow integration with DICOM oriented systems are included in the HL7 Reference Information Model, and facilitate the publication of an HL7/DICOM Implementation Guide.
- **Clinical Genomics** recommends enhancements/extensions to HL7 for exchange of information about clinical genomic orders and observations, collects, reviews, develops and documents clinical genomics use cases, and evaluates existing genomics standards formats such as BSML (Bioinformatics Sequence Markup Language), MAGE-ML (Microarray and GeneExpression Markup Language), LSID (Life Science Identifier).
- **Java** defines application programming interfaces (APIs) to HL7 version 3 artifacts for the Java platform, and promote the development of implementations to validate the API, to serve as reference implementations, and as tools.
- **Public Health** helps to assure that HL7 models and messages address the requirements of the many government and non-governmental agencies involved in population-based and public health surveillance and response activities, including event detection, outbreak investigation, health monitoring, disease/condition case reporting, environmental observations related to health issues, emergency coordination, and legal issues including chain of custody and isolation/quarantine.
- **Security** identifies the needs and technical means for security and evidence of accountability for information communicated according to the HL7 standards.
- **XML** produces recommendations on the application of XML across the entire HL7 remit, coordinating the XML representation of templates, documents and messages.

In summary, HL7 is a critical standard for health-grid type developments. Version 2 is likely to be found in general intra-trust computing scenarios, whereas version 3 will be found in specific inter-trust systems. Version 3, although far from perfect, is of far more interest to the computing researcher than version 2, and should be the focus of projects with a significant research element. Projects with a more pragmatic element are likely to require attention to both standards, complicating the project.

HL7v3 is gaining considerable momentum, and has developed into the focus of a wide range of related health computing issues of interest to the e-Scientist. e-Scientists and Computer Science researchers would find great benefit being represented in these SIGs, where they will find practical examples of many cutting and bleeding edge applications of computer science.

DICOM

DICOM [35] is a comprehensive standard for systems that acquire, transfer, track, process, store, and display images, image metadata and related reports. ‘Image’ in this circumstance includes not only x-ray, MRI and ultrasound, but also waveform data from cardiology and electrophysiology, and expanding to include microscopy, pathology and ophthalmology. ‘Storage’ includes on-line in PACS systems and offline on CD media. ‘Display’ includes provision to consistently apply annotations, to adjust brightness and contrast to achieve consistent display of grey scales, and to control the production of hard copies to paper and film.

Starting at the imaging modality—the instrument or scanner that outputs digital data—DICOM comprehensively specifies the operation of digital systems including: data models used to describe images; metadata definitions for the data models; actions that may be performed upon images; the encoding, semantics, communications protocols and message sequence required to perform these actions.

An image file is described according to an Information Object Definition (IOD), with defined IODs including Digital Mammography x-ray Image; Secondary Capture Image; Magnetic Resonance Image; Ultrasound Multi-frame Image; Mammography CAD Structured Report. An IOD may be 'normalised' or 'composite': Normalised IODs effectively restrict included metadata to that immediately related to the study/image; composite will include other relevant data. A good example of a data item excluded from a normalised IOD, but included in a composite IOD is the relevant patient name. Similar attributes of IODs are partitioned into modules, which are re-used between IODs. Typical modules include patient, device and series information.

IODs are combined with appropriate commands to create Service Classes (section PS 3.4), which include: storage; query/retrieve; print; and worklist management. A Service Class places requirements upon both the consumer and provider of the service and these requirements are detailed. Message streams between consumers and providers are constructed, encoded and exchanged according to standards specified in PS 3.5 and 3.7. Another noteworthy section is PS 3.6: the data dictionary.

Where HL7v2 and DICOM overlap, IHE normally chooses to recommend a DICOM implementation, which both reflects the success of DICOM, and the roots of IHE in the Radiology department. However, this is not to say that there are no problems with the DICOM standard: in organising attributes into modules and comprising IODs from reusable modules, attributes can be duplicated in two or more modules leading to the potential for error; and while better than HL7v2, a DICOM conformance statement still offers no guarantee of interoperability.

One of the immediate justifications for the implementation of grid middleware in healthcare scenarios is the handling of large files such as those generated by modalities. Thus in dealing with all the central issues of image management, DICOM occupies a critical place in many, if not most, health grid projects.

IHE

Integrating Health Enterprise [72] describes and organises sets of use cases into *profiles* accounting for the major integrative tasks that are expected to occur within a hospital. IHE is a prescriptive implementation of HL7 version 2 and DICOM which supports plug-and-play integration between systems without the development of complex, custom middleware. It specifies both which standards are to be used to achieve each use case within a profile, and the information model to accomplish the tasks successfully. As such, IHE develops recipes for obtaining a basic level of integration between healthcare information systems. Other standards are used as necessary: the time synchronisation use case in the technical framework specifies the use of NTP or SNTP time service queries; Kerberos Network Authentication Service version 5 is specified for security services.

IHE does not specify which systems will perform what tasks—it defines actors with roles which may be assigned to systems as required. Terms commonly used in the definitions include the following.

- **order placer:** acts on behalf of the clinician in placing an order for a radiology service.
- **accession number:** a unique identifier for an order.
- **order filler:** acts on behalf of the order placer in fulfilling the order.
- **requested procedure:** is the smallest unit of chargeable work in an order.
- **requested procedure identifier:** a unique identifier for a requested procedure
- **scheduled procedure steps:** must be performed to complete a requested procedure

IHE organises interactions into the following profiles:

- **Scheduled workflow** is the most important profile in IHE simply because it addresses the efficient management and delivery of radiology services. Consisting of over 40 transactions between 9 actors, the profile allows compliant systems to achieve significant functionality. While IHE tries to avoid specifying a physical association between actors, in the scheduled workflow, several actors are grouped and thus system boundaries are suggested. Thus the *Image Manager*, *Image Archive* and *Performed Procedure Step Manager* actors are grouped together, and are suggestive of how IHE views the role of the PACS server. This grouping, and particularly its cardinality, is a limitation of IHE which will be discussed later. The Performed Procedure Step Manager provides an important degree of control and validation to the operation of the department, a requested procedure can be tracked, images supplied can be compared with the images requested, the image manager and archive can prepare to receive files from any connected modalities, and images could be automatically routed to the image display if required. All components in a health grid that have any exposure to IHE will need to participate in a scheduled workflow.
- **Patient information reconciliation** most commonly deals with emergency admissions where the patient's identity is not clearly established. In this profile the PACS system will receive updated patient data and

reconcile this new information with any already held within the DICOM files it has stored. Clearly this is also important as names, addresses and other demographic details change, additionally we may wish to record a summary of the results of assessment and information about other relevant procedures such as physical examinations, biopsies, treatment, HRT etc, along with the screening history.

- **Consistent presentation of images** allows images and image annotations to be viewed correctly and consistently on different hardware. This is particularly important where measures of contrast are artificial, such as in MRI.
- **Presentation of grouped procedures** allows number of individual requested procedures, acquired in a single image capture session to be processed together. This is frequently required in MRI, where a single helical scan of the torso is taken, but the image is split into three sections to apply different brightness/contrast (*windows*), to be individually assessed by different specialist radiologists, and thus individually charged.
- **Access to radiology information** allows the management and delivery of DICOM images and structured reports other information systems and their users. A *report repository* actor is introduced, which could represent systems holding the patient health record. Potential uses of this profile are expected to be rolled up in scheduled workflow operations, however, in some potential architectures the profile could be used to provide support for image series exchange between PACS systems.
- **Key image note** allows a user to flag images as having particular significance, and annotate the selection with a comment. In the envisaged e-DiaMoND application, this could have significance in arbitration, open second reading and on-the-job training.
- **Simple image and numeric reports** simplifies the DICOM Structured Reporting standard to meet many of the reporting needs of the radiology department, facilitates the use of digital dictation and voice recognition, and assigns actors to the roles of report creation, management, storage and viewing to allow vendors to implement the function. In the envisaged e-DiaMoND systems, individual reports would be processed to derive further information ranging from 'all clear' letters, calls for assessment, status reports that coordinate these mail shots and the overall performance of the clinic. Adoption of reporting workflow is not yet widespread.
- **Post-processing workflow** allows the scheduling and monitoring of typical post-imaging work and in a digital breast screening environment provides a mechanism for the agglomeration of individual patient-only image series into a screening session, for image processing including the application of the SMF algorithm, and in support of computer aided detection.
- **Charge processing** supports the exchange of information related to charges among departmental systems, enterprise-wide information systems and billing systems.
- **Basic security** works within an institution's existing security policies and procedures to help protect the confidentiality of patient information. It provides for the consolidation of audit trail events on user activity across several systems interconnected in a secure manner. The profile is based upon Kerberos and as such has broad compatibility with the Globus toolkit.

In simple terms, one could imagine image acquisition in the screening process to be expressed in terms of the IHE profiles and use cases as follows:

- an *Order Placer* actor notifies the Department System Scheduler (RIS), the *Order Filler* actor, of forthcoming appointments. The order placer could be the main hospital appointments administration system, driven from GP practise systems, or autonomous within the screening department. The order will contain all of the patient demographic information required during image generation, securely retrieved from the HIS or RIS using the *Basic Security* profile.
- The *Order Filler* creates a unique identifier for the request, the accession number, and generates entries for *requested procedures* within the radiology department. There is no specification on the granularity of the requested procedure: the entire screening visit could be a single requested procedure, or one could be requested for each breast, or each view of each breast.
- Each requested procedure is comprised of *Scheduled Procedure Steps*, which describe the work done by the operator at the mammography machine. As the operator completes each step, they become *Performed Procedure Steps*. Images are generated by the mammography machine and transferred to the PACS system. The DICOM standard ensures that images that have not been transferred cannot be deleted.

- Images are received by the PACS and a *Post Processing Workflow* operates, queueing images in the PACS for scheduling as a *Post Processing Worklist* which is the digital equivalent of a screening roll. The image manager for the PACS locates and requests additional and appropriate medical history information from its own archives and from hospital records according to local and national policy, and offers post-processing worklists to appropriate radiologists.
- A Radiologist logs onto the screening workstation, which displays the worklists assigned to them. The Radiologist retrieves the image series corresponding to one of the worklist, and according to their preferences, examines the images. For each image a *simple image report* is compiled, recording as a DICOM Structured Report the conclusions of the screening session. The image is windowed and annotated as required by the policy of the local centre so that the diagnosis can be repeated using *consistent presentation of images*.
- The e-Diamond PACS information manager compares Radiologist annotations and generates arbitration sessions or assessment requests for approval as appropriate.
- The Senior Radiologist is notified of differences of opinion and can request expert arbitration by sending a worklist to a remote e-DiaMoND PACS for the attention of another radiologist through *access to radiology information*. Both PACS systems monitor the workflow using *post processing workflow*. A DICOM structured report would be appended to the image, maybe with some *key image notes* and further annotations describing how the conclusion was reached. The Senior Radiologist may use the example in continuous training.
- Another option open to the radiologist is to execute a post processing workflow that operates a centrally located and licensed Computer Aided Diagnosis service.
- The final post processing workflow, triggered by the completion of reading for a daily screening session communicates information to the ADT system to generate letters as required, and to deposit a record of the screening in each of the health records of the women attending the session.

IHE it might be regarded as a panacea to the future development of this and other projects, since can encapsulate the workflow of the local screening clinic and enable cooperation between hospital systems. However, there is an opposite point of view which argues that the fundamental strength of IHE is also its fatal flaw because although IHE offers easy and immediate interoperability between systems, based firmly in use cases, the goal of current and future research must be for autonomous interoperability relying upon underlying semantic models of the healthcare process: the HL7v3 approach which is plainly beyond the scope of IHE.

HL7 version 2 is not fully endorsed, having '*draft interim standard*' status [73], approved for intra-Trust communications where no alternative exists. However, successful PACS integration at present depends upon IHE, which is based on HL7 version 2 and DICOM. Given the short timescale for the rollout of National PACS services [74], many of the local projects will make a pragmatic decision to use IHE, and thus HL7v2 to achieve integration with other hospital systems.

Apart the previous limitation, the current standard also has some more difficulties that include the single PACS limitation and the bundling of IM and IA components. These can be partly mitigated by the engagement of e-Diamond and its industrial partners in the standards making process to ensure that sufficient functionality is supported to achieve integration. The alternative is the development of thicker layers of middleware to hide the lack of critical functionality.

Thus in a very real sense, IHE must be seen as an expediency: a stop-gap that addresses deficiencies in the current standard while the difficult problems of developing and implementing HL7v3 are solved.

4.2 Systems

In this section, we will examine those dependencies implied by the function of the local clinics and the NHS-BSP. These are use-case driven, and in the main imply integration with systems within the trust, the BSP and the NHS.

An efficient Breast Screening Programme will provide and depend upon a number of services at the local level, including: creation and management of appointments; access to daily appointment schedules; access to digital mammography instruments; processing, storage and archival of images; screening workflow management; visualisation of images; access to health records at the local and national level; integration of screening records into the local and national health records; supply of attendance registers; quality metrics; training support.

Integration will be best supported if the local breast screening system behaves as a regular component of a hospital network, providing and consuming relevant services in a standard fashion according to its expected workflows. The radiology function is currently divided into the following components: the image server or

Picture Archive and Communication System (PACS), the radiology workstation where images are displayed and annotated, a departmental scheduling, management and reporting system or Radiology Information System (RIS), and hospital based systems for managing health care and medical records, the Hospital Information System (HIS) and the Admit-Discharge-Transfer (ADT). We will examine each of these components in turn looking for gaps between standards, practices and the requirements of the NHS-BSP and health grid projects.

Picture Archive and Communication Systems (PACS)

The PACS system provides storage, archival and serving of images and associated data, including relevant patient information, image parameters and record provenance, from imaging systems including X-ray, ultrasound and MRI, although this is set to extend beyond the radiology department into pathology, dermatology, cardiology, ophthalmology and any other imaging applications. It also provides services related to the composition and execution of workflows, the reconciliation of the patient information stored in image annotations, image processing, image presentation, and reporting.

A PACS system has grid-like attributes in the storage and management of large volumes of information: in much the same way that grid middleware virtualises remote storage, if a PACS is informed of a scheduled examination, it will queue and prefetch previous, relevant imaging studies from offline storage to an accessible location in anticipation of a clinician's request. IHE [72] has had noteworthy success in integrating PACS with other systems from differing manufacturers. However, an IHE compliant PACS has specific shortcomings relevant to e-DiaMoND and other health grid applications:

- IHE does not support multiple PACS systems within a departmental or hospital network. This poses two problems: federation between PACS at different screening centres is the core goal of e-DiaMoND; screening activities within the local centre are likely to be implemented on a dedicated PACS system because digital mammogram images are considerably larger than those produced by other modalities. However, diagnostic and clinical workstations have a *Multiple sources* option in the *Access to radiology information* profile which could be re-purposed;
- IHE envisages a single Image Manager (IM) and Image Archive (IA) component per server. In combination with the previous limitation, we might expect a standard IHE PACS to offer poor scalability in scenarios requiring high user counts, complicated image processing, or large image archives; and
- the market for Hospital Information Systems is specialist and immature, thus off-the-shelf implementations of the standard make no separation between 'commodity' services (file serving, storage and archival), and 'value added' services related to its specific function, allowing the vendors to charge inflated prices for simple hardware upgrades.

While the failings of the standard could be addressed through engagement with the IHE community, or bypassed through the development of suitable middleware or the use of HL7v3, there are good arguments for the development of a bespoke PACS for the NHS-BSP that

- includes functionality to cooperate with other PACS and thus can be integrated into a grid;
- has a parallel architecture of multiple IM and IA components; that
- runs on commodity hardware such as storage-area network and blade server technologies.

Thus access to grid IA services should be regulated by the PACS offering important benefits for security as the number of users and processes with direct access to the grid layer could be restricted. However, engagement with standards bodies should accompany development so that the resulting system is an early demonstrator of new, standard functionality rather than a bespoke cul-de-sac.

Beyond the obvious screening workflow, the PACS must also support, at minimum, image retrieval for assessment, treatment and general medical record access, and metadata reporting to both the RIS the national system. Dependent upon local RIS functionality other functions may be required (see below). Thus the bespoke PACS will need to support a wide range of IHE use cases including: Scheduled Workflow; Patient Information Reconciliation; Reporting Workflow; Post-Processing Workflow; Consistent Presentation of Images; Access to Radiology Images; and Evidence Documents. These will be important both in the daily function of the screening department, with the care path through assessment, diagnosis and treatment, and in its integration with the wider hospital network.

The issues involved with the development of a custom PACS system are discussed in the following references [75]. Of great interest to developers of health grids is the Pan London PACS project [76]. Covering St George's, Epsom/St Helier, Kingston, Mayday, Queen Mary's and The Royal Marsden, this will provide a uniform PACS service allowing the exchange of images and radiology reports between London Trusts, which share important imaging services. The Pan-London PACS will need to make a significant contribution to the development of HL7v3 and its use within the NHS.

Radiology Information System

The Radiology Information System (RIS) provides the balance of information management and departmental scheduling functionality that is not included in the PACS. Booking, registration, information retrieval from HIS/ADT, examination scheduling and departmental management are all undisputed RIS territory [77]. However, the location of services that manage workflow, report writing and communication is not as clear because the capability to computerise radiology reporting predates the ability to provide a comprehensive PACS service: indeed many existing systems are entirely bespoke and would require considerable effort to retrofit to modern standards. Additionally the reasons for the purchase of RIS and PACS systems are different: RIS systems are implemented within the radiology department to provide internal automation and management, whereas PACS systems are acquired on behalf of the hospital to provide location independent access to radiology images. Thus there is considerable overlap in functionality between commercial products, although the separation of health record functionality between RIS and PACS is artificial and will become increasingly blurred as service based implementations are established.

Radiology management and reporting in the Breast Screening Program is diverse and not extensively computerised. Individual centres have their own policies regarding information required to support screening, operate their own workflows, and record the results of screening on forms of their own design using local terminology. Real comparisons and interoperability between centres will require considerable changes in working practise: it would be best if standardisation began now in preparation for a national system, rather than combining both changes into a single episode. Early agreement of standards would allow individual centres to progress towards computerisation of the BSP at a pace appropriate to local priorities and resources.

Current low levels of computerisation in the BSP imply that a successful RIS implementation could lead to great improvements in the efficiency of screening clinics. Given that the program is under pressure to manage an ever increasing work-load, it is likely that once digital mammography systems are taken up by local clinics, some local RIS provision will be present.

The NHAIS Exeter System

The NHAIS Exeter System [78] is a software suite used by all Health Authorities in England and Wales for the administration of cancer screening call/recall programmes and to deal with patient registration and contractor payments. A range of products and services augment the core software, including the ophthalmic payments system, management information tools, and electronic links between HAs and other NHS bodies. These are supported by a national network of NHAIS support staff. New NHAIS products and services are in development to support Primary Care Trusts in their new responsibilities. These changes will include the development of a standard reporting system to replace the range of systems currently in use. This system would enable reporting data to be stored digitally and in accordance with agreed standards and terminology. e-DiaMoND would again seek to integrate with this system to perform the function of recording screening decisions or structured reports. These reports would then be stored and made available as appropriate using grid technology. Again, integration with the Radiology reporting system would require e-DiaMoND to adhere to IHE standards.

The Screening/Diagnostic Workstation

The screening workstation is the visible front end of digital breast screening and provides the radiologist with their interface. Minimum requirements are for the display and navigation of DICOM images/image series, and the creation and display of structured reports and image annotations, making the screening workstation a specialised IHE diagnostic workstation.

IHE envisages the diagnostic workstation as a flexible, intelligent tool offering considerable freedom in interaction with images and data, offering support for the following profiles: scheduled workflow; post processing workflow; consistent presentation of images; access to radiology information; evidence documents. This picture is complicated by the ambiguity between RIS and PACS functionality in workflow and report creation.

This flexibility must be restricted in the screening clinic, where workflows are subject to strict constraints such as: notifications should also follow strict chronological order and all women attending a session together must be notified of the results of screening simultaneously to avoid undue distress; each local centre has policies on work allocation, second reading and arbitration; the programme requires that continual assessment exercises are completed. However, these restrictions could be equally be enforced within a customised PACS as within the workstation, concentrating bespoke development to a single component of the system.

Despite the existence of the majority of the standards needed for a general, software based IHE diagnostic workstation, an off-the-shelf product is not readily available: current implementations are not only proprietary, but normally directly integrated into the PACS [79]. Standard web technologies have significant limitations if deployed in a PACS role, notably 8-bit per channel RGB. Thus web based implementations are applet based and unsurprisingly tied into the single manufacturer who developed them. A software based Diagnostic Workstation

would be a great asset to the NHSBSP, indeed it would be a useful project in its own right, and there are existing open source initiatives that could be taken forward.

Manufacturers like Sectra have also developed workstations which specialise in mammography and workflow. One significant element which is missing is the ability to work across organisations, a challenge which e-DiaMoND would seek to work with the vendors on enabling.

Admission-Discharge-Transfer Systems

A necessary pre-condition for the effectiveness of any screening programme is a strong correlation between early diagnosis and an improved outcome. Thus it follows that processes which identify and bring the 'at risk' population to the screening clinic, and the mechanisms to integrate suspected cases into the subsequent care path are vitally important. Close coordination with authoritative and current sources of eligibility also helps to prevent distressing letters being received by relatives of the recently deceased.

In an IHE compliant network, general patient health record access is through the Hospital Information System, or HIS. The management admissions is often, but not exclusively, handled by a separate system called an Admit-Discharge-Transfer (ADT) System. The HIS/ADT system acts as the *Order Placer* actor in IHE workflows: a clinician will access one of these systems to place an order for a radiology service.

The screening workflow is more complicated than a routine request for an X-ray of a suspected fracture and this introduces complications into this simple picture. However, one might imagine the following workflow in IHE terminology:

- the set of eligible women within the catchment of the screening centre is identified;
- a draft screening schedule is agreed with the screening clinic;
- invitations for appointment against this schedule are created, mailed and managed in collaboration with the screening clinic;
- the final schedule together with relevant patient details is communicated to the RIS for *order filling*;
- further patient information may be requested by the RIS to automatically annotate images and reports, to perform *Patient Information Reconciliation*, or to provide key data to support the screening activity including: age; HRT status; previous assessment; prior treatment.

Once screening is complete,

- orders for radiology reports are generated and batched for reading
- batches are read and radiology reports generated
- completed screening reports are communicated to the health record;
- the radiology clinic requests discharge for women with clear reports, and places orders for assessment in cases requiring followup;
- the ADT system coordinates the mailing of all clear letters and invitations for assessment as required.

Clearly, this function requires access to reliable demographic and health information sourced from several systems including the HIS/RIS for repeat appointments, the local ADT system for referral from primary care, local cancer registries and ultimately to the single health record. The route to the latter could be central through the NHS-BSP grid, or distributed through the local HIS. Owing to the confused policy on IHE and HL7, the system will need to cope with multiple standards including IHE, DICOM, HL7 version 3 and e-GIF. More on these standards can be found in 4.1.

Screening clinics generally have poor IT support and many/most clinics still manage this process by hand. Since it is not clear that off-the-shelf ADT systems are capable of managing a screening programme, given the key role ADT plays in screening, and the commonality of this element of the screening workflow between different screening programs, there is a good argument for a dedicated screening ADT system, which given the commonality of requirements, could be implemented centrally in support of the NHS' total screening requirement.

4.3 Services

In this section we present some health services that are being into consideration for adopting them for future healthcare.

The National Care Records Service

Previous screening mammograms and related reports would be held within the national grid for Breast Screening. However, the results of symptomatic mammography, in assessment and treatment, the related pathology reports from biopsies, and records of any subsequent treatment for cancer, together with general information held in GP files relating to HRT and family history are all important background to the screening process and their availability may contribute to a reduction in false positives and interval cancers. Access to Cancer Registries and the Office of National Statistics is also important in identifying the characteristics of the eligible population. Facilitating access to the components of an individual's health record is the goal of the National Care Records Service (NCRS).

Electronic records are very well established in primary care, with many GP practices completely automated, but have yet to make as strong an impression within hospitals. Projects have demonstrated success [80], but most have now been halted pending the implementation of the NCRS. The spine will consist of a core medical record, containing basic demographic and health information, together with pointers to remote records in primary, secondary and tertiary care. The intention is that NHS staff can access basic details from the system's inception, augmenting them from other systems as more sources of information are connected to the spine. Eventually it is hoped relevant information can be shared between the Social Services and the NHS, preventing frequent and embarrassing dislocations of service provision for the vulnerable.

However the development of the NCRS is by no means straightforward. There seems to be a fundamental disjoint between the requirements for Caldicot oversight of medical record transfer on an individual, need-to-know basis and the universal access to envisaged by the service (or the pan-London PACS project [76]). Perhaps in recognition of this fact, the scope of the NCRS has recently changed from 'all 50m' to 'consenting' patients [81], which would at least allow the NCRS to operate legally. BT have been awarded a contract to deliver sections of the NCRS [82], but difficulties with the specification have arisen [83]. A body representing General Practitioners has recommended that their members boycott the system until wider consultation occurs, and pending concerns over security [84], although the security of GP records themselves has been brought under question [85]. Project risks are compounded by the early adoption of HL7v3 as the underlying communications standard: HL7v3 is work in progress, and again in common with the Pan-London PACS project, significant development of the standard will be required to achieve the specification.

Given remote access to information from primary, secondary and tertiary care, information from the health record could be required at several points during the screening workflow: by the radiographer prior to and during the screening session; by the workflow composing a worklist for the radiologist; or by on demand from the radiologist during a screening session. Different levels of access are implied by the roles: the radiographer's patient facing role would be served by general information that prompted exploration of the patient's general health; the PACS administrator maintaining workflows should be able to manipulate relevant information without being able to access the information itself; the radiologist may have full privileges to GP records but denied social care records.

Access to locally held medical history could be possible through IHE as soon as suitable systems are available. Wider access to the NCRS depends upon a number of conflicting political, ethical and technical factors which could have a profound effect on system architecture, and it is difficult to guess the eventual access path. Access to the service may have an element of local control, through a Trust gateway, or systems may be able to connect to the NCRS directly. The former might indicate the IHE based interfaces are required, the latter, HL7 version 3.

Diagnosis services

A number of academic and commercial institutes are developing algorithms to mark up mammograms, through the association of structured reports or annotations to the images themselves—DICOM presentation states—based upon similarity matches across mammography archives. Since a National BSP grid will operate open interfaces and standards throughout its fabric: the ability to meet this specification is a natural requirement of any third party supplier to the BSP.

The constitution of the BSP grid offers benefits in a single point of interface to each supplier, in licence control, usage metering, and crucially access policy to source data both from the programme and the health record, to ensure better control of disclosure of patient information. It also offers the service provider the widest array of data with which to make predictions and refine algorithms.

One can imagine a scheduled workflow submitting an image, single image series or a whole screening roll as a job to a service hosted within the grid fabric and receiving modified images through the *evidence creator* profile. The process could conceivably be started either by the radiologist at the screening workstation, or automatically by the PACS or RIS servers during the regular workflow. Service based architectures, and the re-alignment of Globus behind the Web Services Resource Framework (WSRF) will be beneficial for the integration of diagnosis services.

Research interface

Mammograms, combined with associated demographic and clinical information collected during screening, are a powerful research tool. Where consent has already been obtained, within the context of the clinical trial for instance, one can imagine two scenarios operating: a clinician calling up a breast screening history through the NCRS and appending it to an electronic on-study form; a gateway service on the national grid for Breast Screening allowing requests for documents and images from the trials office. Source data could be projected and restricted to match the protocol and presented to the user, and possibly supplied in the absence of consent where the information set is considered to be of low risk and subject to PIAG monitoring and approval.

Alternatively, where a specific hypothesis is being tested, a service within the national grid for breast screening could be created which would hide sensitive personal data from the triallist or epidemiologist, but would nevertheless allow sophisticated computations to be performed on the complete source data set and only the anonymous the results of those computations viewed - metering important healthcare statistics in a dashboard style fashion. How these documents are secured, stored, communicated and validated by their commissioners, and finally publicised after the primary publication of the research remain open questions, but clearly there is great potential to revolutionise clinical research with minimal risk to the individuals concerned.

Other difficulties in clinical research include the rapid recruitment of a statistically sound and significant study population, and the completion of the data collection required for analysis. One could imagine a research interface providing support for these goals: the research interface could allow the screening of patients for eligibility and prompt the responsible clinician to contact the patient for recruitment, or where patients have been previously recruited into a study, or otherwise when it is ethically sound, request the collection of extra data according to an approved research protocol. This would significantly reduce the cost and effort required to perform high quality clinical research. Workflows for the research interface will be complex and require considerable development to ensure all concerns are addressed.

4.4 Summary

Issues surrounding integration of computer systems in pursuit of better healthcare are not solved. While considerable progress is being made, solutions will require real advances in software engineering, architecture, security, information representation and management, standards development, distributed computing, and modelling. Thus health-grids are a particularly fertile area for both e-Science and general computer science research.

5 Security issues

Whereas other non-functional issues, such as manageability, dependability, and quality of performance are largely domain- and application-dependent, security issues pertaining to health grids are largely generic. This is primarily due to the fact that they are derived from British and European legislation. As such, we address these security issues here. We address *some* other non-functional requirements in Part III of this document.

5.1 Security

Security requirements are often characterised in terms of the acronym *CIA*—representing confidentiality, integrity and availability. A system such as e-DiaMoND will—inevitably—be concerned primarily with the first two aspects of security. In this respect, there are ten key areas (in no particular order) that need addressing. These are outlined below.

- **Anonymisation.** There is a clear need for adequate confidentiality mechanisms and minimal information flow when dealing with patient data.
- **Audit trails.** All actions need to be recorded. A key design decision taken within the e-DiaMoND project was that the database will always store ‘old’ versions of records as well as updated ones—nothing is ever deleted.
- **Physical security.** It is of clear importance that the equipment holding medical data is protected by appropriate physical security mechanisms. It should also be made as difficult as possible for unauthorized users to obtain access to network connections.
- **User authentication.** All requests for information need to be authenticated in some way.
- **Client and server authentication.** An approach that focuses primarily on authentication and access control, advocating the use of proxy certificates and global to local identity mapping will be appropriate in this respect.
- **Security breach detection.** Security breach detection is usually an after-the-event measure. It is, however, important to monitor and look for suspicious activity on the network. Alongside detecting a breach, it is also useful to have a limit on what each user can do.
- **Encrypted data movement.** Mutual authentication and encryption work well for local transfers. If the client and server are situated at remote sites, then VPN technology may be used to protect the transfer of data. The VPN creates an encrypted channel between the two participating sites, thereby preventing snooping by third parties.
- **Data integrity.** Not only should data be encrypted, but it should also be digitally signed. This signature will ensure that the data received is the same as that sent. By signing the results of queries sent to the database it will also be possible to be sure that any data output from the system can be trusted.
- **Availability.** The availability of the data when it is required is essential for the smooth running of a healthcare centre. The methods of ensuring this will include standard high-availability techniques, including redundancy; the availability of workstations is also of importance. This requires appropriate maintenance contracts to be in place, and the appropriate use of redundant hardware. A robust network infrastructure is essential within each healthcare centre. The monitoring of network traffic along with the health of switches and other network equipment is an important task. This will allow for the early detection of failures.
- **Access control.** The implementation of a sufficient fine-grained and flexible access control model will be essential for a rolled-out health grid.

Secure health grids

We focus exclusively upon *information security*: issues pertaining to, for example, audit trail capabilities and the NHS firewall—although obviously of extreme importance—are not considered here. (The wider e-Health security issues pertaining to systems such as e-DiaMoND are discussed in [86].) Although our views are necessarily informed by our experiences with e-DiaMoND, we consider that vision described in this section, which is based on [87], should be of interest to the wider UK e-Health community.

The common notion across all the definitions of the *grid* is that of a *virtual organisation*—many disparate logical and physical entities that span different administrative domains coming together to form a single logical

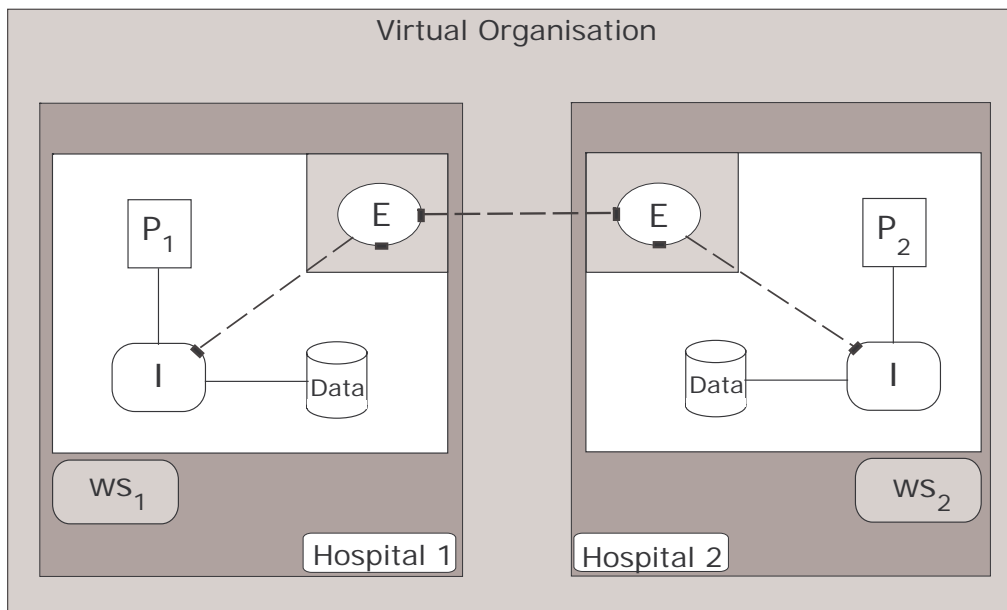


Figure 10: Abstract view of the e-DiaMoND architecture

entity. Any UK-based e-Science project concerned with healthcare has a particular virtual organisation in mind: the National Health Service.

UK-based e-Health projects are obliged to enforce the principles of the Data Protection Act of 1998 [88] and Caldicott Guardian [46], as stated in Chapter 3.

The NHS comprises a number of independent legal entities known as *hospital trusts*. As each trust retains the ownership of all data located at its site, coupled with the fact that each trust determines who can access its data (and under what circumstances), it is easy to conclude that the NHS genuinely does constitute a virtual organisation.

e-DiaMoND VO

The e-DiaMoND platform is concerned with the provision of functionality for several different drivers drawn from the breast care arena, with epidemiological studies and person-centric healthcare being prime examples.

With respect to the first of these examples, one might imagine a national repository of mammograms and related patient data being an invaluable resource for studies of cancer trends. Having sought appropriate approval, a hospital might permit a clinician from a different part of the country to run queries across a subset of its patient information. This permission might, perhaps, be given to a particular individual under certain conditions—perhaps only access to those mammograms associated with smokers over 70 who died between 1995 and 2000 will be granted. It is thus necessary to offer means both of expressing this policy and enforcing it in a fashion that does not impact significantly upon performance.

With respect to the second of these examples, consider a woman who lives in West London—where her hospital records are based—but travels daily to East London to work. It would be more appropriate and convenient for the woman to attend for routine screening in East London—during her lunch break—than it would in West London. Therefore, the West London hospital offering remote access to that woman's information to the East London hospital would offer benefits for that individual, and would be exactly the kind of benefit one would expect from a national database that supported the breast screening process. Again, specific secure access without a significant impact upon performance is required.

The core e-DiaMoND system consists of middleware and a virtualised medical image store to support the concept of a data grid. The virtualised medical image store comprises physical databases, with each being owned and managed by a Breast Care Unit (BCU). The e-DiaMoND grid is formed by participating BCUs coming together as a virtual organisation and uniting their individual databases as a single logical resource.

To motivate the use cases described in the following section we will use an abstract view of the e-DiaMoND system, in which, rather than considering mammography-specific use cases and BCUs, we consider hospitals and more generic realisations of these use cases.

At each node the services are grouped into internally and externally facing services and the virtualisation of

the data sources is assumed to take place at the service level. This abstract view of the e-DiaMoND architecture is presented in Figure 10. Here, we have two hospitals that come together to form a virtual organisation, both of which have externally facing services (E): it is these externally facing services that facilitate communication between sites. In addition, each hospital has its own locally owned database (Data), internal services (I), access control policies (P) and workstations (ws).

This architecture allows each hospital to retain full control of its data and to determine who can access it (and when): this is, of course, in accordance with the principles of the Caldicott Guardian. All user interactions with a hospital are made via the externally facing services, and it is only these externally facing services that are required to present a consistent grid interface.

5.2 Security use cases

In [89], five classes of threat to consider for IT-based healthcare systems were identified:

- insiders making innocent mistakes, causing accidental disclosure of confidential information;
- insiders who abuse access privileges;
- insiders who knowingly access information through spite or for profit;
- an unauthorised physical intruder gains access to information; and
- vengeful employees and outsiders.

Mindful of these classes, within e-DiaMoND, a threat analysis was conducted using the method described in [90]. The team identified assets, threats to those assets, and countermeasures to secure those assets from the threats. For each asset, the team identified the level of confidentiality, availability, and integrity. Risks were measured thus:

- Essential: the system should not be implemented without the relevant security measures in place.
- High: there is a high risk to the project if we do not deploy the relevant countermeasures.
- Medium: there is an average risk to the asset if we do not deploy countermeasures
- Low: there is very little risk to the asset so deploying security countermeasures to these assets is low priority.

We do not propose to report the results of this threat analysis exercise here (the interested reader is referred to [91] for relevant details). Rather, in this section we consider several security use cases that have informed our thinking within the e-DiaMoND project and will be relevant to other projects within the domain. We do not claim that the use cases are exhaustive—it would be impossible to provide such a collection of use cases in the space available—but we would claim that the use cases we present here are representative of the requirements of that health grids should satisfy. Furthermore, given the space available, it is necessary to present these use cases at a relatively high level of abstraction. It should be noted that these use cases are associated with an idealised health *data* grid and, as such, are complementary to the use cases of [92], which is focussed on *computational* grids.

Use case 1: distributed queries of patient data

This use case is illustrated in Figure 11.

A user wishes to query the data held on a subset of the hospitals that form the health grid. Each hospital is allowed to decide its own policy for data access. The user should receive the combined results containing only data that they are permitted to access.

In the first step, a request is sent from a workstation to an externally facing service at the local hospital. The external service then forwards the request both to an internal service at the local hospital and to a number of external services at other hospitals. These external services at the other hospitals then pass the request to their local services. When an internal service receives a request it reads the local policy, and uses it to decide if the user is authorised to access the data requested. At the local hospital the user will probably have significantly higher access than at remote hospitals. If access to the data is permitted the local service will retrieve it from the data source and return the data to the external service; if permission is not granted, a message will still be returned to the external service. The data is then returned to the external service at the local hospital which will combine it and return it to the user.

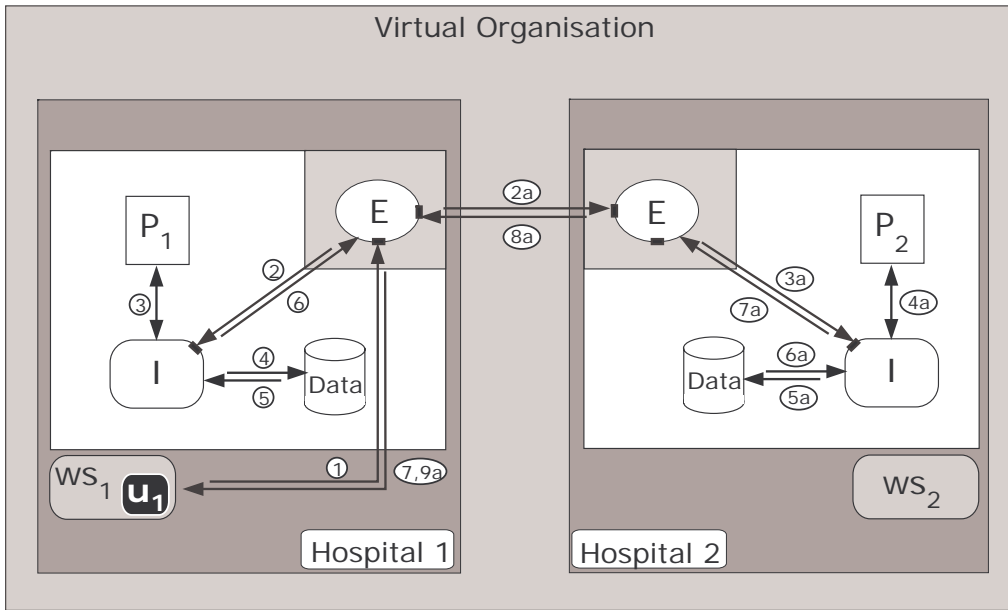


Figure 11: Use Case 1

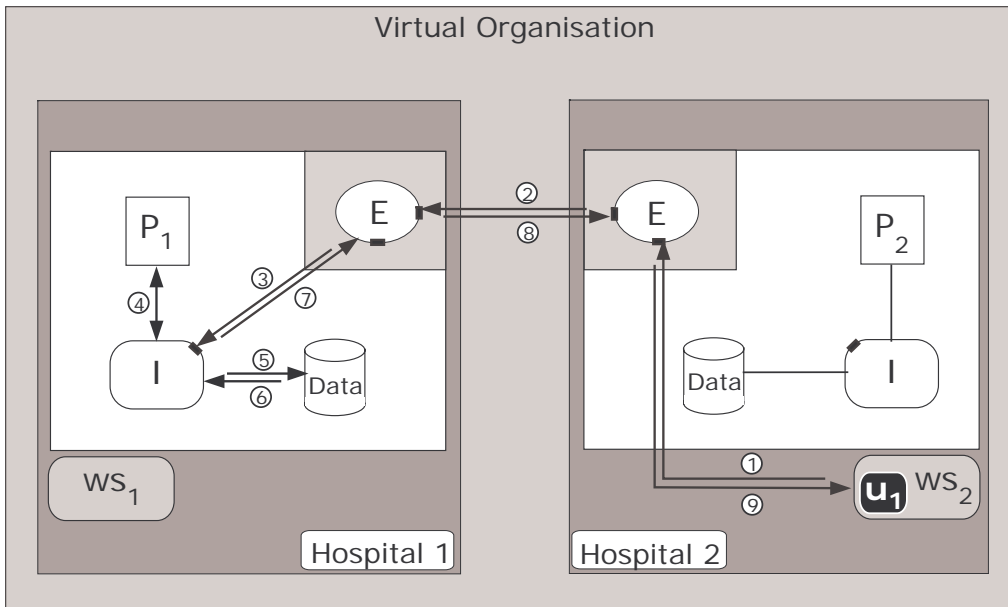


Figure 12: Use Case 2

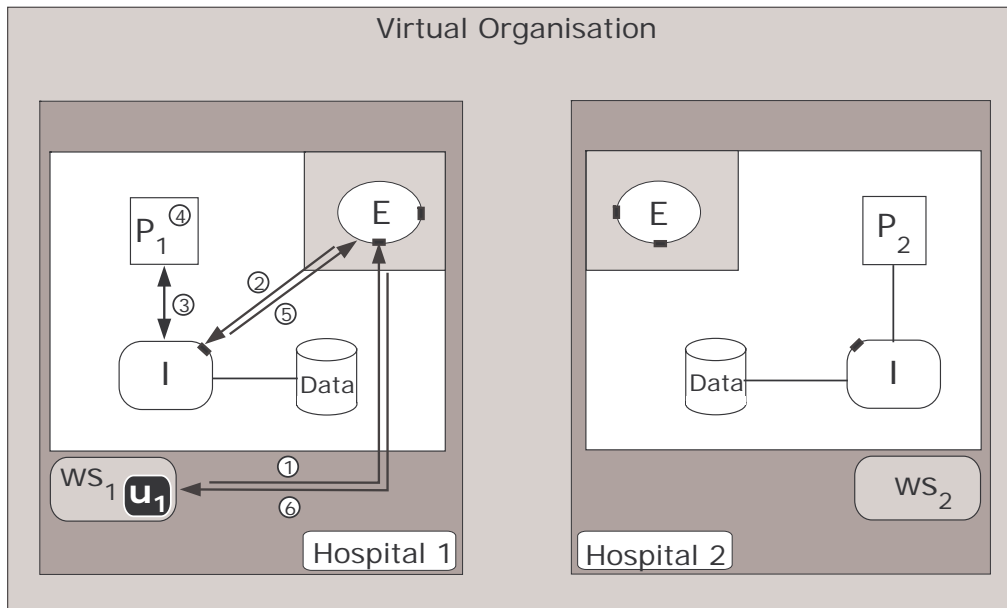


Figure 13: Use Case 3

Use case 2: working at a remote hospital

This use case is illustrated in Figure 12.

A doctor is working at a remote hospital, which is part of the health grid. The doctor should be able to access data from their home hospital, though their request may be subject to a policy that differs from the one used when they are at their home institution.

The first step is a request from the doctor being sent from a workstation at a remote hospital. Although this request is for data from the doctor's home hospital it is first sent to the external service, which is local to the workstation. The external service will then forward the request to the external service at the doctor's home hospital. When the internal service receives the request, it will know the request comes from a doctor who works at the local hospital and it will also know that the request came from a workstation at a remote hospital: as such, it may apply a different policy with respect to the user's request.

Use case 3: delegation of access permissions

This use case is illustrated in Figure 13.

A senior health professional would like to grant access to data to a colleague. This access should be temporary, and could be granted to either a named individual or a group of people.

The first step is a request sent from the workstation to the external service at the local hospital. This is then forwarded to the internal service in the same way as a request for data. The internal service will then check the current policy to see if the user is allowed to modify the policy in the manner requested, and, if the user is allowed to make the modification, the policy is changed and all subsequent access requests will be subject to the new policy.

Use case 4: external access

This use case is illustrated in Figure 14.

Either a health professional working from home or an individual patient wishing to see their own records should be able to access data in accordance with the local hospital's access control policy. The hospital would use a different policy for such external access than would be used for requests from a remote hospital. This use case differs from the others as the request comes from outside of the current virtual organisation. The exact mechanism for authenticating the user is not specified, but it is vital that the internal service is aware of the origin of the request.

Use case 5: modification of data

This use case is illustrated in Figure 15.

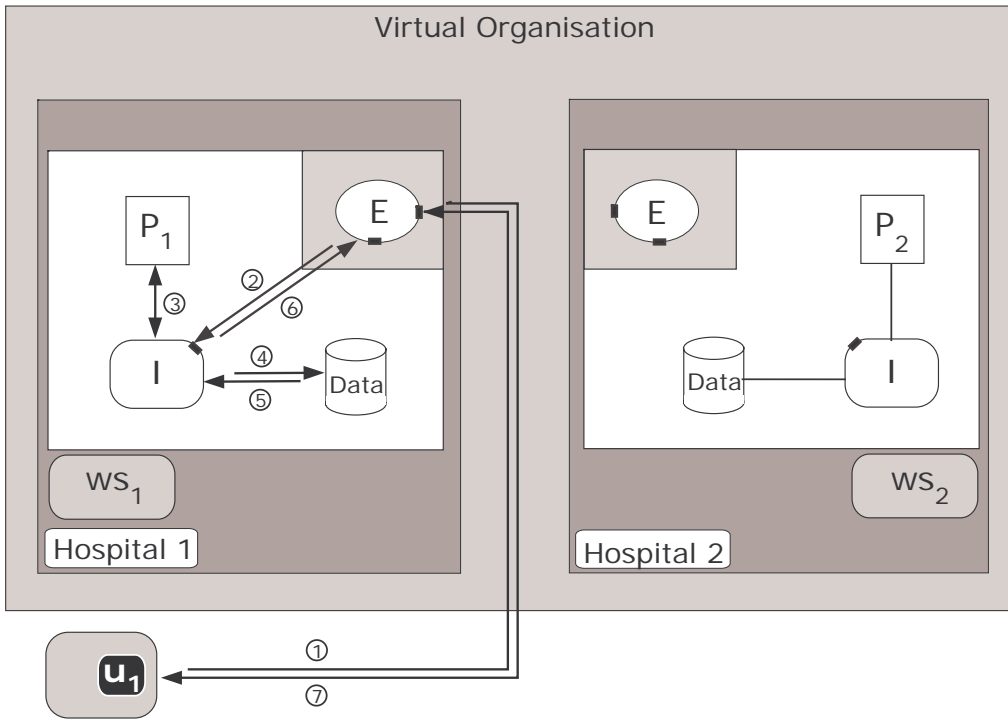


Figure 14: Use Case 4

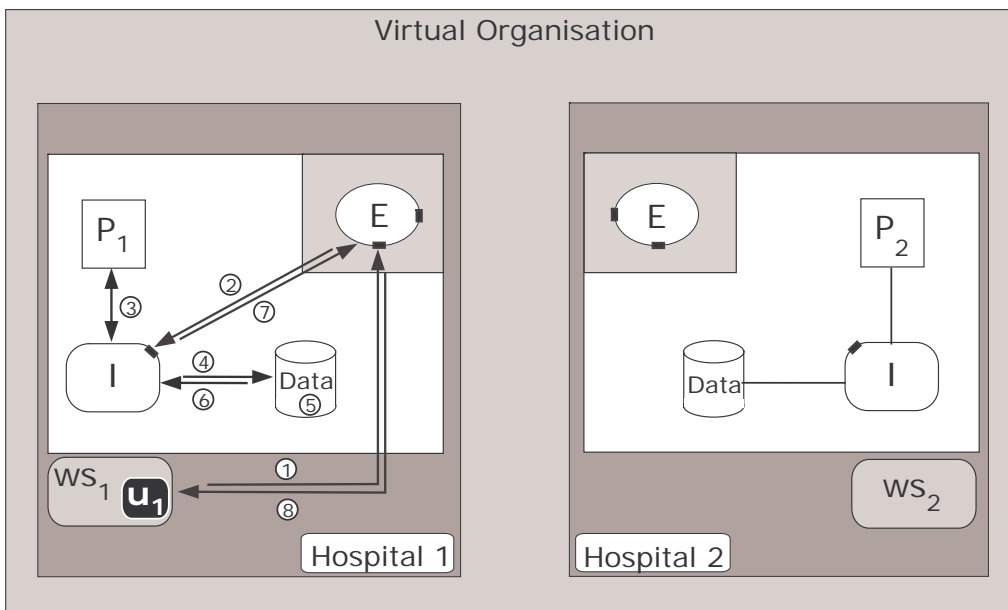


Figure 15: Use Case 5

Having made a clinical decision about a case, a doctor wishes to modify the data stored in the health grid. A doctor will only be able to modify data if a hospital's policy allows it. Each hospital is responsible for the data it stores and as such it should keep a record of all modifications made. This use case is similar to the delegation of access permissions described above, with the only difference being that the data—rather than the policy—is being changed.

Use case 6: transferring patient records

Our final use case draws together aspects of each of the previous use cases. We include this as it imposes an additional requirement that is not introduced by the more generic use cases.

In this use case a patient has moved and is now being treated at a new hospital. As the patient is likely to stay at the new hospital for some time, it would make sense to move their data. To be able to move the data it will first need to be read: this may involve a distributed query as data may already be present at other hospitals. The data will then need to be deleted from one hospital and copied to another—as the responsibility for it has transferred. This will involve the modification of data. Finally the access policies at both of the hospitals may need to be changed to reflect the change of ownership of the data.

If an error were to occur during this transfer process the system could be left in an unstable state. To prevent this, a notion of transaction management must be used to ensure that all the steps have occurred successfully before the modifications are committed.

Further use cases

Use case 6 combined several of the earlier use cases as well as imposing an additional requirement. From the point of view of security requirements, virtually all of the other use cases that we could provide are specializations of one, or subtle combinations of several, of the use cases detailed. For example, the combination of use cases 2 and 3 are representative of a situation in which a clinician wishes to update the access control policy from a remote hospital.

5.3 Technological issues

The e-DiaMoND prototype infrastructure has been developed using a number of currently available technologies, with the primary focus of the project being twofold: to establish the feasibility of grid-enabled healthcare and to identify the key issues that would have to be solved to deploy such a solution for real. Rather than concentrate on the architecture of the prototype system, in this section we consider the second of the above points, with a view to advancing the health grid security agenda. In particular, we explore what current technologies are available to implement the idealised health grid described in Section 5.1 that is also capable of supporting the use cases described in Section 5.2.

First and foremost an appropriate service-based infrastructure is needed. The current trend—which appears set to continue—is towards using web services for this purpose. It is, of course, possible to build a grid from the ground up using web services and this approach has been used successfully in many projects. It is also possible to use a toolkit such as the Globus Toolkit [39]—this has been the approach taken in the development of the e-DiaMoND prototype.

There has been much work at standardising grids with the Open Grid Service Architecture (OGSA) [93] being prominent in this regard. The current Globus Toolkit is built upon the Open Grid Service Infrastructure (OGSI) [34], although future versions of the toolkit will be based upon the emerging Web Service Resource Framework (WSRF) [94] standards.

In order to present a grid interface to a data source such as a relational database the e-DiaMoND project uses the OGSA-DAI (Open Grid Services Architecture — Data Access and Integration) distribution [37]. The OGSA-DAI distribution is also compatible with the OGSA-DQP (Open Grid Services Architecture - Distributed Query Processor) distribution [95] which is capable of performing joins between OGSA-DAI data sources.

Many of the differences between the use cases presented in Section 5.2 are related to the identity of the user and the location of the workstation that they are trying to access data from. To provide user credentials, X509 certificates [96] can be used. Similarly, it would be possible to use X509 certificates to identify the workstation that the request was sent from, although this introduces its own problems. Support for X509 certificates is built into the Grid Security Infrastructure (GSI) [97], which is part of the Globus Toolkit. An alternative approach to authentication is to use Kerberos [98], which differs most prominently from the use of X509 certificates in that it uses symmetric temporary session keys, with a trusted third party server issuing these session keys.

There are many mechanisms for the secure transfer of data between the hospitals, many of these are based on SSL (Secure Sockets Layer) [99, 100] which provides a secure channel. It is equally possible to encrypt the data before transmission which can then be passed through an insecure channel.

To describe access requests and access control policies, the eXtensible Access Control Markup Language (XACML) [101] could be used. This can be used in conjunction within the Security Assertion Markup Language (SAML) framework [102], which is a framework for exchanging authentication and authorization information.

This is, of course, just one possible framework for controlling access to data. An example of an alternative approach, using *role-based* access control, might be the utilisation of the PERMIS infrastructure [103].

5.4 A gap analysis

Having identified the key constraints associated with health grids, the use cases pertaining to information security and the currently available technological solutions, we are now in a position to present a gap analysis. In presenting this gap analysis we make a number of assumptions. The justification for making these assumptions is that without certain fundamental aspects of a secure distributed architecture being in place, the consideration of information security is relatively futile.

First, we assume the existence of an appropriate service-based infrastructure to facilitate the type of distributed healthcare grid with which we are concerned. Second, we assume that there is some method for issuing credentials. Third, we assume that appropriate mechanisms are in place to ensure the secure transfer of data between different members of the virtual organisation. Fourth, and finally, we assume that there is some means by which users can describe their requests. Specific technologies that might be used to realise these assumptions were discussed in Section 5.3.

Given our use cases and constraints—as well as the above assumptions—we have identified five gaps that must be addressed to ensure an appropriate level of information security for grid-enabled healthcare systems. We consider each in turn.

There is a need for each individual entity within the virtual organisation to be able to describe the access control policies associated with that site's data: the utilisation of the OASIS eXtensible Access Control Markup Language (XACML) [101] is a partial solution to this problem. However, the verbosity and complexity of XACML descriptions makes tool support essential for the language's widespread adoption. Unfortunately such tools for writing and processing XACML are still in their infancy.

If temporary access to data—as described in the delegation use case—is to be allowed, then either the policies must contain time-based information, or a secondary process would have to change the policies at the appropriate time.

The need to allow access to only that patient data that is absolutely necessary, is stated in the second and third principles of the Caldicott Guardian. It is also a requirement of the Data Protection Act (in accordance with the second principle).

If a doctor is performing a distributed query to find information about a patient, they may find that they are being denied access to certain data. If the doctor is trying to make a diagnosis it is important that they know that their access has been denied. Likewise an epidemiologist may draw false conclusions if they are silently denied access to certain data about certain patients. This would suggest informing the user when access has been denied.

However, there may be a case where access is denied because a patient does not wish certain aspects of their data to be used for anything but primary care. In this case it would be relatively simple to deduce much of the data from messages describing the denial of access: there is the potential for information flow.

The suitable description of denial of access—to avoid both the drawing of false conclusions and the potential for information flow—is a key issue.

If credentials are to be used to authenticate users in our vision of a health grid, then several problems need to be overcome. First, it is important that the credentials are portable: this is needed if doctors are to be able to move around their own hospital or to access data when visiting other hospitals. It would not be feasible for doctors to carry a computer with them at all times, so a more portable solution would be needed.

Another problem associated with credentials is propagation: external services need to be trusted to pass users' credentials to other services. One method commonly employed for this task is a proxy certificate; however, these have the problem that the user has to trust the intermediate services to use the proxy certificate as they intended. Other solutions may involve the user signing a more specific request, this however reduces the ability of an intermediary to make useful decisions for the user.

Finally, it must be possible to revoke a set of credentials: this is especially important if doctors are carrying their credentials around with them. Current systems often rely on the service actively requesting lists of revoked certificates, leaving a window of opportunity for people to misuse the credentials. Use of the Online Certificate Status Protocol (OCSP) [104] may reduce the window to a minimum but this has the drawback of requiring frequent communication.

If data is modified at a single hospital as in use cases 3 and 5 then it is possible to handle any transactional requirements within an internal service. For more complicated use cases, such as the transfer of patient records, transactional support is required that spans hospitals. As all communication between hospitals is handled by

the externally facing services, the transactions will need to take place between these services. Although there have been some proposed standards for transactions between web services, at the time of writing these are not commonly supported.

One of the requirements of the health grid is that it should be possible to give temporary access to data. It is important that people given temporary access do not make a copy of the data. In our idealised health grid, data would have a lifetime and would be deleted after use: if this were the case, it would then be following the fifth principle of the Data Protection Act.

Part III: The Future

6 Technology options

This section of the blueprint provides an overview of relevant web and grid standards and specifications, many of which have emerged since the beginning of the e-DiaMoND project. In particular, this section discusses their potential use in a health grid environment.

We start by providing an overview of web services.

6.1 Web services

A web service is a collection of protocols and standards used for exchanging data between applications. Specifically, the generic term *Web Services* is used to describe the means of integrating Web-based applications using XML [105], SOAP [106], WSDL [107], and UDDI [108].

Software applications written in various programming languages and running on various platforms can use web services to exchange data over computer networks like the Internet, in a manner similar to inter-process communication on a single computer, with this interoperability being due to the use of open standards.

OASIS [109] and the W3C [110] are the steering bodies responsible for the architecture and standardization of web services. To improve interoperability between web service implementations, the Web Services Interoperability (WS-I) organisation [111] has been developing a series of profiles to further define the standards involved.

WS-I [111] is an open industry organisation chartered to promote web services interoperability across platforms, operating systems and programming languages. Specifically, WS-I provide resources for web services developers to create interoperable web services and verify that their results are compliant with WS-I guidelines. Key WS-I deliverables include profiles, sample applications and testing tools.

Web services can provide very loose coupling between an application that uses the web service and the web service itself. This should allow either piece to change without negatively affecting the other. This flexibility may become increasingly important as software is built by assembling individual components into a complete application. Web services also provide interoperability between various software applications running on various platforms.

6.2 SOA

Service Oriented Architecture (SOA) is an architectural style, the goal of which is to achieve loose coupling among interacting software agents. A service is a unit of work undertaken by a service provider to achieve desired end results for a service consumer. Both provider and consumer are roles played by software agents on behalf of their owners.

The idea of SOA departs significantly from that of object-oriented programming, which strongly suggests that one should bind data and its processing together. The results of a service are usually the change of state for the consumer but can also be a change of state for the provider or for both.

Before it can be said that an architecture is service-oriented there are a few rules that have to be followed [112]:

1. The messages must be descriptive, rather than instructive, because the service provider is responsible for solving the problem.
2. Service providers will be unable to understand your request if your messages are not written in a format, structure, and vocabulary that is understood by all parties. Limiting the vocabulary and structure of messages is a necessity for any efficient communication. The more restricted a message is, the easier it is to understand the message, although it comes at the expense of reduced extensibility.
3. Extensibility is vitally important as the world is an ever-changing place and so is any environment in which a software system lives. Those changes demand corresponding changes in the software system, service consumers, providers, and the messages they exchange. If messages are not extensible, consumers and providers will be locked into one particular version of a service.
4. An SOA must have a mechanism that enables a consumer to discover a service provider under the context of a service sought by the consumer.

It is generally accepted that a web service is a SOA with at least the following additional constraints:

- interfaces must be based on Internet protocols such as HTTP, FTP, and SMTP; and
- except for binary data attachments, messages must be in XML.

There are two main styles of Web services: SOAP web services and REST web services.

6.3 SOAP

SOAP (Simple Object Access Protocol) is a lightweight protocol for exchange of information in a decentralised, distributed environment. It is an XML-based protocol that consists of three parts: an envelope that defines a framework for describing what is in a message and how to process it, a set of encoding rules for expressing instances of application-defined datatypes, and a convention for representing remote procedure calls and responses [106].

SOAP 1.1 was suggested in a note to W3C in May 2000 (by Developmentor, IBM, Lotus, Microsoft, Userland, etc), as a protocol for exchanging information in a distributed environment. The W3C SOAP 1.1 document was only a note available for discussion. The first working drafts were published on the 17th of December 2001 and SOAP 1.2 was released as a W3C Recommendation on the 24th of June 2003.

SOAP web services

A SOAP web service is the most common and marketed form of web service in the industry, with some simply collapsing 'web service' into SOAP and WSDL services. SOAP provides 'a message construct that can be exchanged over a variety of underlying protocols' according to the SOAP 1.2 Primer [113]. In other words, SOAP acts like an envelope that carries its contents. One benefit of SOAP is that it allows rich message exchange patterns ranging from traditional request-and-response to broadcasting and sophisticated message correlations. There are two flavours of SOAP web services: SOAP RPC and document-centric SOAP web service.

A SOAP RPC (Remote Procedure Call) web service breaks the second constraint required by a SOA. A SOAP RPC web service encodes RPC (remote procedure calls) in SOAP messages. In other words, SOAP RPC 'tunnels' new application-specific RPC interfaces through an underlying generic interface. Effectively, it prescribes both system behaviours and application semantics. Because system behaviours are very difficult to prescribe in a distributed environment, applications created with SOAP RPC are not interoperable by nature.

Faced with this difficulty, both WS-I basic profile and SOAP 1.2 have made the support of RPC optional. RPC also tends to be instructive rather than descriptive, which is against the spirit of SOA. Ironically, SOAP was originally designed just for RPC.

REST web services

Representational State Transfer (REST) is an 'architectural style' that basically exploits the existing technology and protocols of the Web, including HTTP and XML. The term REST was first introduced to describe the web architecture. A REST web service is an SOA based on the concept of 'resource'. A resource is anything that has a URI and may have zero or more representations. A REST web service requires the following additional constraints:

- interfaces are limited to HTTP;
- most messages are in XML, confined by a schema written in a schema language such as XML Schema from W3C;
- simple messages can be encoded with URL encoding; and
- service and service providers must be resources while a consumer can be a resource.

REST web services require little infrastructure support apart from standard HTTP and XML processing technologies, which are now well supported by most programming languages and platforms. REST web services are simple and effective because HTTP is the most widely available interface, and it is good enough for most applications. In many cases, the simplicity of HTTP simply outweighs the complexity of introducing an additional transport layer.

REST vs SOAP

REST is simpler to use than the better-known SOAP approach, which requires writing or using a provided server program (to serve data) and a client program (to request data). SOAP, however, offers potentially greater capability. For example, a syndicator that wanted to include up-to-date stock prices to subscribing Web sites might need to use SOAP, which allows a greater amount of program interaction between client and server.

REST is consistent with an information publishing approach that a number of Web log sites use to describe some aspects of their site content, called RDF Site Summary (RSS) [114]. RSS uses the Resource Description Framework (RDF), a standard way to describe a Web site or other Internet resource.

In addition, REST is an architectural style, while SOAP is a messaging protocol. REST applications can use SOAP. The W3C SOAP 1.2 recommendation was specifically designed to support RESTful applications [115].

SOAP attachments

Often, there is a need for a SOAP message to be transmitted together with attachments of various sorts like images, drawings, xml documents, etc. Such data are often in a particular binary format. The SOAP with Attachments [116], a technical note submitted in W3C in 2000, details usage of 'MIME multipart/related' media type and URI schemes for referencing the MIME parts. The methods described in the note treat the multipart MIME structure as essentially a part of the transfer protocol binding, i.e., on a par with the transfer protocol headers as far as the SOAP message is concerned. The multipart structure, though given a name (SOAP message package) is not an entity that can be unambiguously identified as such because there is no token explicitly expressing the intent to make it such an entity.

A 'SOAP message package' contains a primary SOAP 1.1 message. It may also contain additional entities that are not lexically within the SOAP message but are related in some manner. These entities may contain data in formats other than XML. The primary SOAP 1.1 message in a message package may reference the additional entities. Such additional entities are often informally referred to as 'attachments'.

Using SOAP with attachments enables solutions for integration of disparate systems and provides a viable foundation for emerging XML-based specifications.

XOP

On 2004, the XML Protocol Working Group released a first draft of XOP [117], XML-binary Optimised Packaging, and a revised draft of MTOM, the Message Transmission Optimisation Mechanism, that leverages XOP.

XOP is an alternate serialisation of XML that just happens to look like a MIME multipart/related package, with an XML document as the root part. That root part is very similar to the XML serialisation of the document, except that base64-encoded data is replaced by a reference to one of the MIME parts, which isn't base64 encoded. This results in avoiding the bulk and overhead in processing associated with encoding, the only way that binary data can be fitted directly into an XML world.

XOP can be used for any XML-based format; MTOM is just a description of how XOP is layered into the SOAP HTTP transport.

MTOM and XOP have much broader vendor support than any predecessor specification for XML-to-binary serialisation. MTOM and XOP describe how to produce optimised binary encodings of XML content within SOAP 1.2 payloads. MTOM and XOP preserve one of XML's great strengths: the transparency of the tagged, logical data structure that a particular document implements.

For any given XML document, MTOM and XOP preserve its logical transparency structure by encoding that structure in a text-based 'XML Information Set' manifest, while allowing any of the document's contents to be serialized to any binary encoding. In particular, these specifications support binary encoding of XML content as Multipurpose Internet Messaging Extensions Multipart/Related body parts and encapsulation of those parts — along with the associated XML Information Set manifest - within SOAP 1.2 envelopes. The specifications also describe how to encapsulate binary-encoded XML body parts directly within HTTP packets (in cases where SOAP doesn't enter the equation), thereby reducing the size of XML files for transmission and/or storage.

One limitation of MTOM and XOP is that they only can be used to define hop-specific encoding contracts between adjacent nodes within an XML/SOAP-message-handling transmission path. The specifications don't describe how to define global XML-encoding optimisation policies that apply across any arbitrary number of XML/SOAP-handling intermediary nodes - an important requirement.

At the time of writing MTOM and XOP aren't yet ratified W3C standards and few commercial implementations exist. It has been argued that companies that want to base their XML-optimisation strategy on these specifications might have to wait a few years before they are implemented broadly in commercial application platforms, middleware environments and development tools [118].

6.4 ebXML

Electronic Business using eXtensible Markup Language [119], commonly known as e-business XML, or ebXML, is a family of XML-based standards sponsored by OASIS and UN/CEFACT (United Nations Centre for Trade Facilitation and Electronic Business) whose mission is to provide an open, XML-based infrastructure that enables the global use of electronic business information in an interoperable, secure and consistent manner by all trading partners.

ebXML uses web service standards as a foundation layer. ebXML builds upon these to define agreed mechanisms for aspects such as the registration, description and discovery of services, and service contract (or quality of service) negotiations in a generic manner. ebXML is not itself a standard; rather, it is a container for several key specification standards administered by UN/CEFACT and OASIS. Key ebXML standards include [119] ebXML Messaging Services, ebXML Registry, ebXML Business Process Specification Schema and ebXML Collaboration Protocol Profile and Agreement, which are designed to support advanced integration requirements, whether in a private sector (e-business) or public sector (e-health, e-government) context.

ebXML Messaging Services is a standard under the e-Business XML umbrella which provides a secure and reliable SOAP / web services based transport protocol to the ebXML Architecture.

The International Organization for Standardization (ISO) has approved the following four ebXML specifications as the ISO 15000 standard, under the general title, Electronic business eXtensible markup language:

- ISO 15000-1: ebXML Collaborative Partner Profile Agreement.
- ISO 15000-2: ebXML Messaging Service Specification.
- ISO 15000-3: ebXML Registry Information Model.
- ISO 15000-4: ebXML Registry Services Specification

In late April 2004, Health Level 7 (HL7), expanded ebXML adoption in health care by releasing its new transport standard that includes support for ebXML messaging. However, other initiatives indicate greater use of ebXML in health care beyond message transport. The HL7 announcement approved ebXML messaging, and corresponding web services messaging profiles, as draft standards for trial use (DSTU) for 24 months. See [120] and [121] for further details.

ebSOA and web services

On April 2004, OASIS announced plans to advance an electronic business architecture that builds on ebXML and other web services technologies. The new OASIS Electronic Business Service Oriented Architecture (ebSOA) Technical Committee will use ebXML Technical Architecture v1.04 as a starting point for describing a service-oriented architecture and practical implementation techniques that take into account the ebXML OASIS Standards (recently approved as ISO 15000), as well as recent developments from other OASIS technical committees and standards bodies.

eBusiness Service Oriented Architectures provide an operating system that doesn't run on a single computer, but instead exists on multiple computer systems, with various specialised software providing a framework where composite applications based on XML services can be assembled.

SOA patterns will become the roadmap that companies can use to publish services into a stable and secure shared business processing environments, currently being built on the Internet. Companies will also subscribe to processes and information feeds based on XML messaging formats, which will be delivered over the Internet-based SOA frameworks.

The ebSOA Tech Committee is focused on providing architectural descriptions and usage guidelines for constructing these reliable Service Oriented Architecture frameworks.

OASIS and its member organisations will build momentum for ebXML by articulating how the standards fit into the emerging service-oriented architectures that users need to stay competitive in their IT efforts.

A modern service-oriented architecture is core to enabling enterprises to efficiently integrate and distribute business processes across multiple systems with reliability and security. The OASIS ebSOA Technical Committee will explore how various ebXML and web service components work together as a cohesive system. This work is essential to every enterprise.

The tension between web services and ebXML communities

ebXML was a predecessor of Web Services and gave the ideas of modern web services development.

According to OASIS analysts,

“when the ebXML architecture was first conceived, the term ‘web services’ hadn't even been coined, and now, many of the ebXML layers have been reconciled to embrace core World Wide Web Consortium (W3C) technology, such as WSDL and SOAP.”

Despite this, many will argue that ebXML was very hard to use, define and install. OASIS is trying to recreate a new generation of ebXML messaging standard that is much more reliable and flexible, and —importantly— by learns from the past's mistakes. The Web Services community though seems to be rather critical about this new standard, claiming that it is much more complicated to use as well as very difficult to install it. The arguments are many from both sides and details can be found in the following: [122], [123], [124], and [125].

Of course, there is the view from some web service researchers:

“unless web services technologies are applied within the context of standards such as ebXML, we will end up with simple, stateless web services, and not the complex and collaborative business transactions that organisations need.” [126]

6.5 Other web services standards and specifications

Web Services Notification (WSN)

The Web Services Notification (WSN) pattern [127] defines a set of specifications that standardise the way web services interact using the notification pattern. In the notification pattern, a web service disseminates information to a set of other web services, without having to have prior knowledge of these other web services. Characteristics of this pattern include the following.

- The web services that wish to consume information are registered with the web service that is capable of distributing it. As part of this registration process they may provide some indication of the nature of the information that they wish to receive.
- The distributing web service disseminates information by sending one-way messages to the web services that are registered to receive it. It is possible that more than one web service is registered to consume the same information. In such cases, each web service that is registered receives a separate copy of the information.
- The distributing web service may send any number of messages to each registered web service; it is not limited to sending just a single message.

WS-Addressing

WS-Addressing [128] is a W3C specification that provides transport-neutral mechanisms to address web services and messages. Specifically, this specification defines XML elements to identify web service endpoints and to secure end-to-end endpoint identification in messages. This specification enables messaging systems to support message transmission through networks that include processing nodes such as endpoint managers, firewalls, and gateways in a transport-neutral manner.

An endpoint reference represents the address of a web service deployed over a network endpoint. It is represented as an XML serialisation usually returned by a web service request to create a new resource. An endpoint reference might contain, besides the web service address, metadata such as service description and reference properties.

WS-Security

WS-Security [129] is an OASIS specification that describes enhancements to SOAP messaging to provide quality of protection through message integrity, message confidentiality, and single message authentication. These mechanisms can be used to accommodate a wide variety of security models and encryption technologies.

WS-Security also provides a general-purpose mechanism for associating security tokens with messages. No specific type of security token is required by WS-Security since it is designed to be extensible, and as a result, to support multiple security token formats.

Additionally, WS-Security describes how to encode binary security tokens. Specifically, the specification describes how to encode X.509 certificates and Kerberos tickets as well as how to include opaque encrypted keys. It also includes extensibility mechanisms that can be used to further describe the characteristics of the credentials that are included with a message.

WS-Transaction

WS-Transaction [130] is a WS-I specification that defines what constitutes a transaction and what will determine when it has completed successfully. Each transaction is part of an overall set of activities that constitute a business process that is performed by cooperating web services. WS-Coordination [130] is a companion specification that defines the context and exactly how information is exchanged during the business process.

Business Process Execution Language

Business Process Execution Language (BPEL) [131]—co-developed by BEA Systems, IBM, Microsoft, SAP and Siebel Systems—is now an OASIS working standard that provides a language for the formal specification of business processes and business interaction protocols. By doing so, it seeks to give web services the ability to support business transactions. The specification defines an interoperable integration model intended to facilitate the expansion of automated process integration in both the intra-corporate and the business-to-business spaces. The complementary specifications WS-Coordination and WS-Transaction orchestrate the choreography of web services while BPEL serves to articulate them.

Web Services Choreography Interface

BPEL actually competes with another prominent standards effort, the W3C's Web Services Choreography Interface (WSCI) specification [132]. WSCI—co-developed by BEA Systems, Intalio, SAP, and Sun Microsystems—is an XML-based interface description language that describes the flow of messages exchanged by web services in choreographed interactions. It seeks to describe the observable behaviour of web services in terms of dependencies among exchanged messages, featuring sequencing rules, correlation, exception handling and transactions.

6.6 Grid services

The basic principles of grid computing were formed in the early 1970s when computers were first linked by networks, and the idea of harnessing unused CPU cycles was born. With the rapid expansion of the Internet in the last decade the idea of a grid was raised again and many grid oriented organisations emerged.

OGSA (Open Grid Services Architecture) [33] [93] was the first attempt, initially driven by the Globus Grid Forum (GGF) community [133], to specify and document an architecture and a framework for grid services deployment. OGSA is a distributed interaction and computing architecture based on grid services, ensuring interoperability on heterogeneous systems so different types of systems can communicate and share information.

GGF is a community-initiated forum of thousands of individuals from industry and research leading the global standardisation effort for grid computing. GGF's primary objectives are to promote and support the development, deployment, and implementation of grid technologies and applications via the creation and documentation of 'best practices' in technical specifications, user experiences, and implementation guidelines.

The OGSi (Open Grid Services Infrastructure) specification version 1.0 (OGSi) [34], released in July 2003, referred to the base infrastructure on which OGSA was built. At its core was the 'Grid Service Specification', which defined the standard interfaces and behaviors of a grid service, building on a web service base. Specifically OGSi defined a set of conventions and extensions on the use of Web Service Definition Language and XML Schema to enable stateful web services.

The Globus Project [134], which started at Argonne National Laboratories and the University of Chicago, aimed to provide a sort of 'sum of services', or 'toolkit', approach to grid computing. In other words, the Globus toolkit was intended to include libraries, services, and functionality necessary for applications developers to build a stand-alone application that executed in a grid-like environment. With the Globus Toolkit 3 (GT3) [39], in the middle of 2003, the Globus Project offered an open source implementation of version 1 of the OGSi Specification.

Concurrently, the OGSA-DAI project [37] was concerned with constructing a middleware to assist with access and integration of data from separate data sources via the grid. The project was conceived by the UK Database Task Force who worked closely with the Global Grid Forum and the Globus team [134].

When the e-DiaMoND project started, the Globus Toolkit 3 was the state-of-the-art middleware for grid deployment. It was used and deployed from a huge global community and it was the first grid toolkit that tried to encapsulate various aspects of the grid. Moreover, it was the most complete in terms of security integration and support, while its architecture was also based on the web service model. As a result e-DiaMoND grid has been built based on GT3 and uses OGSA-DAI 4.0 as middleware for supporting the exposure of its data resources to the e-DiaMoND grid.

6.7 WS-RF

On the 20th of January 2004, the Globus Team and IBM in conjunction with HP announced the WS-Resource Framework [94], a new web services specification to integrate grid and web services standards. These documents were submitted to OASIS standards group [109] in March 2004.

The WS-Resource Framework is inspired by the work of the Global Grid Forum's Open Grid Services Infrastructure (OGSi) Working Group. Indeed, it can be viewed as a straightforward refactoring of the concepts and interfaces developed in the OGSi v1.0 specification in a manner that exploits recent developments in web services architecture. The WS-Resource Framework retains essentially all of the functional capabilities present in OGSi, while changing some of the syntax (e.g., to exploit WS-Addressing) and also adopting a different terminology in its presentation. In addition, the WS-Resource Framework partitions OGSi functionality into five distinct, composable specifications, namely WS-Resource [135], WS-ResourceProperties (WSRF-RP) [136], WS-ResourceLifetime (WSRF-RL) [137], WS-ServiceGroup (WSRF-SG) [138], and WS-BaseFaults (WSRF-BF) [139] specifications.

These specifications allow the programmer to declare and implement the association between a web service and one or more stateful resources. They describe the means by which a view of the state of the resource is defined and associated with a web services description, forming the overall type definition of a WS-Resource.

A WS-Resource is defined as an entity composed by a web service and a stateful resource. A stateful resource can be used in the web service message exchanges. WS-Resources can be created and destroyed and their state can be queried or modified via message exchanges. WS-Resource implements the ACID properties, which are familiar from transaction processing. Most of these properties are described in the Web Services Atomic Transaction specification [WS-AtomicTransaction]:

- Atomicity: stateful resource updates within a transactional unit are made in an all-or-nothing fashion.
- Consistency: stateful resources should always be in a consistent state even after failures.
- Isolation: updates to stateful resources should be isolated within a given transactional work unit.
- Durability: provides for the permanence of stateful resource updates made under the transactional unit of work.

A stateful service is a service that has access to, or manipulates, logical stateful resources through the propagation of execution context in headers on message exchanges. In general, a stateless service enhances reliability and scalability. For example, after a failure, a service can be restarted without concern of previous interactions. New service instances can be created or destroyed in response to the system load. Thus, stateless services are considered a good engineering practice by the web services community. However, there are situations where a stateful service—a service that manipulates stateful resources based on message exchanges—may be desirable. Such scenarios involve interoperability among services.

Globus Toolkit 4

Globus Toolkit 4 (GT4) (released in early 2005) features a new implementation of the Web Services Resource Framework (WSRF) and the Service Notification (WSN) standards. GT4 provides an API for building stateful web services targeted to distributed heterogeneous computing environments.

All the well-known GT3 protocols—WS-GRAM for resource management, RFT for data management and MDS for information services—have been redesigned to use WSRF.

6.8 Other relevant middleware

In addition to the afore mentioned standards and specifications there are additional, specialised in the domain of computational grid and job submission management, such as Condor-G [140] and UNICORE [141].

Condor-G

The Condor-G system leverages two distinct areas:

- security and resource access in multi-domain environments, as supported within the Globus Toolkit, and
- management of computation and harnessing of resources within a single administrative domain, embodied within the Condor system.

Condor-G combines the inter-domain resource management protocols of the Globus Toolkit and the intra-domain resource and job management methods of Condor to allow the user to harness multi-domain resources as if they all belong to one personal domain.

UNICORE

UNICORE (Uniform Interface to Computing Resources) offers a ready-to-run grid system including client and server software. Although UNICORE makes distributed computing and data resources available in a seamless and secure way through intranets and the Internet, comparing it with Globus, UNICORE does not have any Resource Discovery mechanisms and as so could not be described as a full grid system. UNICORE represents a very different approach, that of modelling the universe of work flows that users might conceivably wish to submit to a grid. Whereas in Globus resource discovery is conceptually separated from the language in which the actual work flows and scripts are submitted, in UNICORE they are almost exactly the same thing.

EGEE

The Enabling Grids for E-science (EGEE) project [142], funded by the European Commission, has been built on the EU Research Network GEANT [143] with the aim to exploit grid expertise generated by many EU, national and international grid projects to date. gLite [144] was born from the collaborative efforts of more than 11 different academic and industrial research centres as part of the EGEE Project, with the goal to provide a bleeding-edge, best-of-breed framework for building grid applications tapping into the power of distributed computing and storage resources across the Internet. At the time of writing gLite is still in its early stages of development.

gLite

In the gLite prototype, the AliEn [145] components are used to provide an initial implementation of components and services like a file and metadata catalogue, task queue, package manager and various user interfaces in the form of a command line prompt, an application API and corresponding Grid Access Service (GAS). AliEn (ALICE Environment) [145] is a lightweight grid framework which is built around Open Source components using the web services model. It has been initially developed by the ALICE collaboration [146] as a production environment for the simulation, reconstruction, and analysis of Physics data in a distributed way. The architecture of AliEn provided a blueprint and a starting point for developing the gLite architecture.

OMII

OMII_1 is a web service infrastructure for building grid applications to support collaborative computing, developed by the Open Middleware Infrastructure Institute (OMII) [147]. It focuses on providing an open source system that addresses the user requirements of combining ease of use within a secure environment. The OMII_1 base is a freely downloadable and has open source web service container with WS-Security [129] enhancements. By adding the OMII_1 services to this container, its capability is extended to provide a secure and accountable file and compute grid.

Legion

Another example of grid middleware that was initiated at about the same time as Globus project is the Legion project at the University of Virginia [148]. Legion took a different approach from Globus, in which the components, resources, and services of the grid are all represented by software objects that can all be addressed within a single federated namespace. The Legion approach allows applications developers to select and define system-level responsibility. So Globus can be characterised as a 'sum of services' architecture, while Legion is an integrated architecture. Specifically, Globus uses sets of pre-existing components that are grouped into composite toolkits. For example, there are services for scheduling, authentication, and remote job spawning. As designers add a new piece of functionality to Globus, they design an interface for that piece. There is no underlying common architecture. In contrast, Legion builds its higher level structure on top of a single unified object model. When designers wish to add new parts they need to plug their new objects into the common programming interface so that they can communicate with the already established object model [149].

Avaki

Currently one out of many commercial grid software is Avaki [150]. Avaki, according to its creators, is more focussed on creating a comprehensive software that will bring together computing, data, and application resources from multiple locations, administrative domains, and computing platforms in an environment that should be secure and easy to administer. Avaki 2.1 grid software enables wide-area access to processing power, data, and applications in a single, uniform operating environment. The software provides platform for building compute grid, data grid and application grid. Avaki is one organisation, among many others, that is looking forward to use web services for grid application integration.

6.9 Web services and the NHS

It is inevitable, as web services are becoming widely used in the NHS for communication between systems, that any deployed health grid would have to be web services-savvy (See [151] for details.) In this respect, web services provide an alternative to more traditional messaging technologies, although in many cases their usage is not as well understood as older messaging approaches. Nevertheless, web services form an integral part of the National Programme for IT (NPfIT), which, in this context, cover SOAP, WSDL and UDDI.

Where web services are required, they must conform to the standards specified in e-GIF, i.e., SOAP v1.2 and WSDL v1.1. Where there is a requirement for such services to be published to a publicly available service directory, they must conform to UDDI v3. Conformance to these standards will make communications within NHS systems using web services easier, faster to implement and cheaper. It will also help enable interoperability and conformance with e-GIF.

ebXML and NHS

Details of the use of ebXML within NPfIT, the National Programme for IT in the NHS, are given in [152]. An overview is provided here.

A central component of the NHS Care Records Service is the Transactional Messaging Service (TMS) Spine using the ebXML Messaging Service OASIS Standard and is likely to become one of the largest applications of this technical specification. Transaction Messaging Service (TMS) soon will replace DTS.

The Transaction and Messaging Service provides the communications infrastructure for the National Programme. It serves to interconnect regional network clusters managed by Local Service Providers (LSPs) and

national services such as systems for electronic booking and transmission of prescriptions. The technology framework used for TMS is based on a large number of advanced technical specifications and standards. This includes the ebXML Messaging Service OASIS Standard, as well as the Security Assertion Markup Language (SAML) OASIS Standard and other web services specifications.

Within the TMSSpine, ebXML is used to provide reliable messaging functionality. National services such as the Electronic Booking Service (Choose and Book) and Electronic Transmission of Prescriptions are accessed using pairs of XML request and response documents. These documents are transported within the NHS network as ebXML messages.

6.10 Summary

The domain of web and grid services is continuously evolving and as such it is a moving target for all the health (and not limited to these) projects that they want to use it for their integration. As a result, it should be noted that the aim of every health project should be the compatibility—to as great an extent as possible—with the generic principles of web and grid services architectures.

7 e-DiaMoND-specific issues

In the previous sections, we have considered issues pertaining to the development of generic health grids within the UK. In this section, we turn our attention to e-DiaMoND-specific issues. Specifically we consider both the applications associated with the initial project and the proposed additional applications for follow-on projects.

7.1 e-DiaMoND applications

The initial e-DiaMoND project has concerned itself with the high volume, rapid activity of radiologists within screening clinics and the prototyping of digital systems to support this activity. Extending this capability to assessment and symptomatic clinics is considered by radiologists to be of equal importance. The potential to enable the integration of case information throughout the process is a key driver in this regard, as is the enabling of clinicians to work more collaboratively through technology. Such an extension will require the project to extend the grid architecture to cooperate with the assessment data acquisition processes, as well as existing assessment hardware and applications. Furthermore, this will require close cooperation with the NHS to enable such a system to work over NHSnet.

7.2 The breast screening domain

The NHS Breast Screening Programme [4] is a virtual organisation (VO) comprised from over 90 individual screening clinics throughout the UK, coordinated from Sheffield. A primary goal of the e-DiaMoND project was to conduct research that facilitates proper architecture and specification of integrated computer support for the BSP VO.

Computerisation of breast screening has the potential to improve the accuracy of the NHSBSP [153]. It also offers compelling benefits for patients, local centres, the national programme and the NHS. Patients benefit from optimal x-ray doses, lower rates of recall through poorly exposed or developed images [154], and reductions in processing time which in turn offers the potential to reduce anxiety. Local centres, once completely digital, could additionally benefit from the reduction in overheads and delays; in reduced cost in the storage and management of images; in integration with trust systems including local health and care records and appointment management systems; and in closer management of and increased efficiency in the screening process.

Integration of local services into a national grid for Breast Screening offers the potential for the creation of new services including: load balancing through the routine exchange of screening sessions; standardised, enhanced, metered training and continual assessment; specialist diagnosis services such as expert second reading/arbitration and computer aided diagnosis; and efficient support for clinical research. Further the NHS could benefit from improved economics and public perceptions from the combination of the above factors.

The data acquisition issue

The current situation in the UK causes significant problems with introducing digital information environments into the Breast Screening Programme. Currently the screening clinics and mobile vans use conventional screening machines which produce film x-rays. These film x-rays are read on large scale and desktop light boxes in darkened rooms. Surprisingly, full field digital x-ray machines are used in several of the assessment clinics and in the US, screening is usually performed on digital equipment. It is likely to be many years before conventional x-ray machines are replaced in the UK clinics unless significant investment is made to enable this sooner.

For this reason, e-DiaMoND has had to deal with the complex process of processing data to create digital cases from patient hardcopy folders.

The UK also differs from many other countries including the US in that patient information is held locally at a clinic or at a surgery, whereas in other countries this data is owned by the patient.

7.3 Screening

Every Breast Care Unit, symptomatic clinic and private breast imaging centre which supports digital mammography could be connected to an e-DiaMoND-like system and every image taken at those centres could be stored on a local storage unit which will be connected to form a vast 'virtual database'. A national solution, with the inevitable massive cost savings, run by well-trained, well-supported IT professionals is an attractive solution. Unfortunately, the key requirements pertaining to both security and interoperability mean that no existing plain internet-based solution is appropriate. Grid computing, however, allows an opportunity for the seamless, flexible and secure distribution of data and additionally allows the sharing and optimisation of computing power. Grid's flexibility will be a critical element in the introduction of such a system.

The case for utilising grid technology to support the work of highly skilled breast radiologists working for the UK Breast Screening Programme was described in the e-DiaMoND proposal and this case has been enhanced as we have explored this area further in the project. The pressures on clinical staff have increased substantially over the last twelve months as the age range for screening has increased from 50-64 to 50-70,

as public awareness has encouraged more women to take the opportunity to be screened and as expectations have risen in terms of detection rates encouraging clinics to perform multiple reads to ensure that fewer cases are missed.

These additional pressures come at a time when there is a severe shortage of experienced radiologists, and radiographers are undertaking traditional tasks previously carried out by radiologists to relieve the burden. Moreover there is an uneven spread of radiologist across UK where in some centres radiologists are sparse and in some others they are numerous.

The ability to deploy a system which enables clinics to share the reading of cases with other clinics and to benefit from computer aided detection and feature extraction algorithms clearly has its attractions. For example NHS BSP has long had concerns about the consistency of image quality across the country and the adherence to the nationally set quality control guidelines. PERFOMS [155] and QA ensure this to some degree, but a digital solution could provide regular, actual performance feedback.

Variations in quality of service across the UK in detecting cancers is alarming and this 'postcode lottery' is indicated by the following statistics [5].

- In the West Midlands, 5.9% of women are recalled from screening to assessment whereas in the South and West it was 9.1% and in Scotland 11.5%
- In the West Midlands, 1.6 women per 1000 screened had a benign biopsy carried out, whereas in the South and West it was 3.3 and Scotland 3.7
- In the South Thames region some 5.8 cancers are detected per 1000 women screened, whereas in Northern Ireland the number is 4.3.

This form of the e-DiaMoND system will allow women to move around the country and still have past mammograms available. In the UK, with the increasing acceptability of digital mammography and the movement towards electronic patient records, this model is clearly viable. The model is, however, dependent on digital mammography.

A key aspect to e-DiaMoND is the ability for a centre to request additional readings for either a case, or a whole screening session. This might be necessary for covering staff shortages due to illness, holidays or the (temporary) need to 'speed' up screening and assessment in order to catch up, for some reason. The end result is that a clinic connected to the e-DiaMoND system should never need to single read a screening session. In short, this functionality allows for balancing resources to ensure that no single unit ever 'fails' and could include the use of readers working at home in a tele-radiology fashion. This should bring the time between screening and a letter giving the all-clear to a minimum (subject to the need for all people in screening rounds to get all-clear or recall letters at the same time).

The e-DiaMoND workstations are based around a flat-panel display, thus providing the potential for massively reducing space problems at the clinic or at home, but also removing issues around the need for dark rooms, as their higher performance in lighter environments is well-known. Prior to images being dispatched for storage on the e-DiaMoND system, automatic, objective quality control procedures take place which check positioning and image quality in order to reduce the number of women recalled to assessment due to technical recalls. The e-DiaMoND workstations are built for optimal use so that when not in use, not only do they provide a compute resource for epidemiological and other research work, they are also suitable for teaching.

As the e-DiaMoND system becomes accepted and reaches certain levels of critical mass, more revolutionary concepts such as automated resource balancing based on a pool of data can be investigated. These concepts could follow the route of some US experiments that have shown the benefit of a financial incentive for each session read: enabling those radiologists who have time and seek money to work longer hours, and those that have other duties (or perhaps families) to work shorter hours. Again, the key word is flexibility.

Tele-radiology

The NHS in the UK has made radical changes to radiology services by utilising technology to use resources more effectively. Many hospitals now require clinicians to provide round the clock cover by enabling them to access digital information from home. An example of this is where a patient has a serious head injury, has undergone a CAT scan and the hospital requires the scan results to be read by a radiologist during the night. By enabling this data to be read from home, a decision may be made by the clinician without having to travel to the hospital. For small images, this process is achievable by suitable dial-up services and adequate security. Obviously this would be more problematic for larger data sets.

Extending this principal, a hospital may consider utilising resources from other hospitals or from outside the UK to cover shortfalls. In order to achieve this, such an outsourcing concept would be underpinned by robust and secure infrastructure in terms of network connectivity and authentication mechanisms, as well as agreed contracts and service level agreements, including costing models, liability models, and quality of service.

Where resources are utilised from outside of the UK, the NHS may wish to consider agreeing standards of education and training for those who will be performing these services.

Although the use of remote readers has long been discussed in radiology, systems such as e-DiaMoND now offers the opportunity to make it finally happen. However, the use of remote readers and several aspects of resource balancing require the disassociation of the assessment decision with the screening decision. It is believed that it is this association that provides the feedback necessary to a reader to know that they are not only reading well but also to expand their knowledge by seeing a case the whole way through to surgery. An e-DiaMoND-like system, by being present all the way through to end of assessment, could ensure that feedback is available to even remote readers in the form of monthly reports on the basis of biopsy proven results. This level of 'real-time' feedback could complement the current PERFORMS testing system (and in service monitoring) and would enable training to be focussed and provided to the reader immediately, if needed.

A future e-DiaMoND screening system could be built around the arbitration process which is the gold standard. The administrator will request two reads which could be any combination of named local radiologists or remote readers (found by asking the community for resource information). As per typical arbitration rules, the readers are able to 'not recall', 'recall' and all marked as 'recall' will be submitted for arbitration, with the system automatically compiling the cases for arbitration. The local site then allocates arbitration time at which the remote reader can also participate, sharing imagery, audio, video and even cursor movements. As it is known that arbitration provides a training and social opportunity, other readers are allowed to participate locally and remotely. The arbitration decision is made by a third reader who is a local radiologist.

Another aspect of breast screening radiology is a sense of isolation for readers in smaller clinics, and the belief that there are discussions with peers about difficult cases that enables learning and increased knowledge. A national system could encourage a sense of community and discussion by allowing local centres to submit cases for a national training database (see below), and by allowing readers to view the whole community of readers who are currently active. Regional groupings of radiologists could be encouraged to meet to review interesting cases from that area — again, flagged by the readers during screening. Readers can engage a training opportunity into their practices by tapping into e-DiaMoND's vast database of images to read several interesting cases prior to each screening session ('a warm- up session').

7.4 Training

The training of mammography radiologists in the UK is performed through both formal education and via a process of mentoring within the clinics, by skilled clinicians. This process is believed to result in varied skill levels across the UK as clinics where fewer cancers are seen or there are fewer senior clinicians often are unable to train to the same level as larger clinics with higher throughput of cases. In discussions [156] [157] with clinical collaborators on the e-DiaMoND project, it was stated that a trainee radiologist could only ever become as good as their mentor, which is an indication that clinicians recognise this constraint.

Radiologist training in the UK

The training of UK radiologists begins with four years of training in general radiology, followed by a further year of training devoted to one sub-specialty for those who wish to declare a specialist interest; or training in a mixture of two or more sub-specialties. These twelve months will usually be undertaken in the fifth year of training, but may be scheduled in a modular fashion over the fourth and fifth years. An additional period of sub-specialty training may be needed for those dedicated to a single sub-specialty, e.g. neuroradiology. All major teaching hospitals provide the necessary courses. Entrants to these courses must possess a medical degree. Once qualified, radiologists who have chosen to specialise in mammography will typically take up a post at an NHSBSP centre. For the newly qualified screening radiologist, training then continues less formally, 'on the job', as he/she learns the particularities of the clinic's work practices [157] [156].

Thereafter, opportunities for further training of a more formal kind are generally pursued in ways that are compatible with clinical work, i.e., short courses run by designated training clinics, and study days. Although the current system is clearly generating high quality readers there are many aspects to it which would prosper from having a properly designed on-line training system. Such a system should enable more people to learn; trainees to learn faster and better; would remove the issues regarding handling films; would enable training cases to be easily shared; would enable trainees to see a full range of different signs; and would enable trainees to share the experiences of groups of experts rather than just a single mentor. In short, a grid-based training system should improve screening accuracy by providing trainee readers with an additional depth and breadth of film reading experience in conjunction with reliable feedback.

The data used in training is usually selected by clinical staff, for unusual features, from standard screening and assessment information. These cases are archived for training and utilised within that clinic for education purposes. This information is currently in hardcopy.

e-DiaMoND training

The e-DiaMoND project has designed and developed a prototype system to demonstrate the power of a grid-enabled training system for breast screening. This system utilises digitised cases, selected as cases of special interest, and demonstrates the benefits of making this information available across screening clinics to other resources. Collaborative training may also extend to real time education with radiologists utilising technology to discuss such cases in real time.

The need for initial training and then constant training of radiologists and other readers is clear from the evidence of training versus performance papers. In fact, several studies have shown a direct correlation between reading volumes and reading sensitivity and specificity, with one study concluding that accuracy is related to the logarithm of the number of films that the reader has read previously (for between 10 and 12,000 cases). This is a known relationship in other domains and is known as the Power Law.

A training system built on e-DiaMoND could use the vast database and the feedback about each reader's performance to deliver highly- sophisticated, focussed training for each individual. It is an engaging learning experience helping readers to broaden their experience both before and during practice but also reducing the teaching effort that is required from the experts. e-DiaMoND could hold a national training database, completely anonymized and separate from the main e-DiaMoND database. At any point in the screening or symptomatic process, the reader at screening or assessment can mark a case as being of 'national interest' and provide annotations to areas of interest. These cases could then be stored in a holding area on the training database prior to being assessed by a team of national experts who will review the results and produce a consensus view on the case with a consensus annotation. Once this is done, the case becomes live and training systems across the country can start using it. Cases can come from screening, assessment and symptomatic clinics, and it might be expected that all interval cancers to be entered.

A critical aspect of the training system will be the generation of test rollers for the reader to review in a screening scenario. The rollers could be the normal mix of normals and cancers, or could be a whole series of cases related to one sign, such as architectural distortion. The normal cases will be drawn randomly from the vast number of normals sitting in the main e-DiaMoND system. Drawing from such a source ensures a novel reading session each time. The training system is supported by the reader being able to draw on a rich and diverse set of material beyond the cases. Namely:

- Complete teaching files that show not only x-rays, but also ultrasound, images of palpation, magnification views, and videos showing the radiologist decision making process.
- Multimedia material that could be used to introduce new topics, such as videos of lectures by luminaries, videos of working film readers, journal and conference papers, NHSBP publications, references to textbooks.

The e-DiaMoND system could maintain a central record for each user, allowing the system to monitor the user's progress but also allowing the user to move workstations as they need to: in short, the grid gives the user flexibility to fit training into his/her busy day. By monitoring the test rollers, the e-DiaMoND system will build up knowledge of how difficult or hard certain cases are and then request further annotations on those that are causing most difficulty. Dialogue between the trainee readers and colleagues over such cases could become part of the case record.

All the training will be conducted on the e-DiaMoND soft-copy review workstation which is geared towards screening and therefore able to emulate screening in a training environment. Once again, the sense of community will be employed to allow the trainee via the workstation to feel part of the community and be able to communicate with other trainees, or experts on the system.

7.5 Epidemiology

e-DiaMoND provides an unparalleled opportunity to study and understand epidemiology issues in breast cancer. For the first time, not only will all the textual and numeric data pertaining to a significant number of digitised cases be available but also all the image information along with the computing resources to be able to truly investigate image properties and correlate them with the text, numbers and known breast cancer risk predictors.

The most obvious examples are related to breast density. Epidemiologists have long believed that breast density (the non-fat composition of the breast) is related to breast cancer risk and recent papers have started to establish this fact. Radiologists know that the breast becomes less dense with age as menopause causes the dense milk producing tissue to change to fat, but there has been no conclusive graph produced showing density versus age due to the sheer numbers required to get an 'average'. The e-DiaMoND database opens up such opportunities.

The use by e-DiaMoND of Standard Mammogram Form (SMF) ensures that radiologists are not required to provide subjective reading of breast density: these values are automatically generated along with volume of glandular tissue.

Grid technology also opens up the prospect of linking with other grids, to, for example, perform pan-European or pan-World studies: How does temperature affect breast cancer risk? What about diet? How about ethnicity? What about use of HRT?. Access to epidemiological researchers could be done on the basis of cross-referenced databases. Epidemiological databases have such highly sensitive and specific information in that they can never really be anonymised. Thus, a central e-DiaMoND system could hold the images and much 'base information', but the bulk of the text and numeric data will held and cross-referenced. Examples of future uses are MOG (Mammography Oestrogen and Growth Factors) Study, which has 8000- 9000 women for whom information including questionnaire data, blood samples, endocrinological, family history, and socio-economic data has been collected. The image component of this study alone is around 14.4TB.

7.6 Data mining

Over the last 5 years, there have been commercially available computer-aided detection (CADe) systems which are aimed at providing a second read in screening, in effect replacing the second reader. Currently, these systems have struggled to make an impact outside of the USA, possibly due to the intensive nature of screening programmes in countries such as the UK, but also due to the cost and support requirements. Within the e-DiaMoND system, a centralised CADe service could be run, with the local centres being able to choose which CADe system they desire, but also, what settings they require it to run with (regulatory issues currently preclude this, but the grid and the database will enable this to happen on a local basis).

The choice of CADe system will be made on the basis of publicly available performance graphs showing how each CADe system is operating over the e-DiaMoND database. Whereas CADe is currently targeted at providing a second read, within e-DiaMoND where arbitration and resource balancing takes place, CADe becomes a support tool for all readers.

Despite the above, unfortunately the development of true and novel applications of image processing has been hampered by the lack of vast, available databases of images. e-DiaMoND produces a gold-mine of information to be mined (and the compute resource to do so) by sophisticated image processing algorithms for many purposes, including:

- development of next generation computer-aided detection;
- building a comprehensive model of normal variation across the population of temporal changes and left-right differences, for detection and diagnosis purposes but also training;
- development of tools such as 'find-one-like-it' for use in assessment or arbitration, and teaching;
- development of tools such as 'find-one-like-it' for use in treatment planning; and
- development of temporal comparison for, amongst other things, the analysis of change for improved computer-aided detection.

The generation of such vast databases is always prone to different equipment in different places generating different images, be it higher contrast, darker images, or noisier. The e-DiaMoND system addresses this issue head-on by the use of standardisation via SMF.

7.7 Technical constraints

While the e-DiaMoND system aims to support fully the previous mentioned applications there are numerous technical constraints that impose some limitations or add complexity to various aspects of the applications. Here we analyse the most important of these.

Storage

Storage is a critical constraint to our initial prototype since the medical images should be stored unaltered and uncompressed at each site. These requirements are a consequence of legal and ethical constraints that govern the health domain. In general the use of low quality or modified medical images is prohibited. This is because the possibility of misinterpreting the image is thought to increase when the quality is lowered or when modifications mask important fine detail. Additional factors that increase the storage requirements are the number of the mammogram copies need to be stored and the length of time they are needed to be retained for.

The e-DiaMoND project team considers a multiple tiered approach to storage is the most appropriate solution to the storage needs of a full scale system. The use of high speed, expensive, disks for storing the currently

active data and lower performance, cheaper, disks for storing data that is migrating to or from long term slower storage. This suggested mix gives a balance between cost and performance. In considering the BSP, women are called for screening every three years. The lower speed disks could hold all the data for the current screening population as well as data migrating to or from permanent storage. The high speed disks holding the data of those currently undergoing screening or treatment.

Longer term storage for archive and recovery could be achieved in several ways. Traditionally backup and archive is usually performed using tape as the medium. This is still an entirely feasible option, utilising one of many tape library products available. The tapes could also be replicated periodically and taken off site to a secure storage location.

There are several manufacturers of tape library systems:

- IBM, especially 3494 Tape Library;
- Quantum, especially PX720 Tape Library;
- StorageTek, especially SL8500; and
- HP, especially HP StorageWorks ESL9000 Series Tape Libraries.

Generally tapes have a bad reputation for their long term stability and general reliability.

More recently large arrays of removable IDE hard disk drives have become available. This could be a useful addition to supply more flexible storage options before the tape level. Moreover the removable disks have higher access speed than the speed of the tapes.

In addition CD/DVD media is also available for a medium term storage option. Nevertheless, the useful life of CD/DVD is frequently questioned.

Since the number of women that are having a mammogram per year and the number of the images per breast taken in each visit vary significantly from time to time depend upon any new health regulations, it is difficult to give anything other than a rough estimate of the total storage requirement. The following assumptions could be used to provide an approximation of the storage need.

- Number of screening invites accepted = 1.6M ([5])
- Number of views per breast = 2
- Number of mammograms per case = 4
- Number of Mammograms per year = 6.4M

The size of each mammogram is dependant on the method in which the digital representation was generated. The most likely methods are the Full Field Digital machines and the digitisation of x-ray films. These methods have a variety of hardware available to perform this task. This fact leads to a wide range of possible file sizes for the final digital images. An indication of what the size ranges are is the case of a file which results approximately to 14MB from an x-ray digitisation at the resolution of 82μ . This size can be increased up to approximately 75MB for images generated by a Full Field Digital machine.

If we therefore assume that each image will be 75MB then the total storage required per year would be in the order of:

$$6.6M \times 75MB = 480TB \text{ (where } 1TB = 1000GB)$$

In addition to this large volume of data generated by new visits, the BCUs will have to digitise and store the existing medical images. Although this would be a one off process over a number of years, the result of this process will involve a great deal of physical effort in addition to the vast storage needs. From the above calculations it is obvious that storage needs of an e-DiaMoND system is a far from trivial logistical problem.

Network requirements

The network requirements of the e-DiaMoND project could be used to lead thinking on the requirements the NHS network. There is clearly a need, based on the volume of the data that will need to be transfer over the network, for a high speed (GBit) Server to workstation connectivity.

Moving case data, from the server to the workstation in an acceptable time for the clinician who wants to use this data, is a critical factor for the success of the e-DiaMoND and its adoption by the clinicians. If a single visit generated 300 MB of images —full field digital machine taking 4 images— then if we assume 50% utilisation of a GBit network it will take in the order of 5 seconds to transfer the data. This derives easily form the following calculation: 8 bits per byte, $300 * 8 \text{ Mbits} = 2.4 \text{ Gbits}$ of data. So 50% of GBit is 500Mb/s, means 5 seconds allows 2.5Gb of data. In this calculation we have ignored headers and other network overheads.

The conclusion of our estimations in terms of network capabilities of the system is that the links between the sites that are sharing data should be as fast as possible. Moreover it is of equal importance that the connectivity to the network to be by physically secure connection points, and all data is transmitted in a secure manner.

Software

The operating systems of our current implementation are built on IBM AIX 5.x on the majority of the servers. Microsoft Windows XP are used on the current workstations.

In terms of middleware for the prototype system Globus Toolkit 3/OGSA-DAI hosted in a tomcat axis container has been used. The decision to use GT3/OGSA-DAI for e-DiaMoND was partly based on the assumption that these would be the likely technologies that grids would be built upon for the medium term. This is no longer the case, with web services being seen as the way forward by IBM and Globus with their announcement of WSRF [94].

It is however difficult to proceed at present as there were no stable releases of WSRF. There is also the additional complication that the proposed WS standards that underpin WSRF are mainly in the early phases of consultation and not currently supported by available containers.

The other main software components being used by the project are: IBM DB2 that has been used as a RDMS for storing non textual data and IBM Content Manager being used as an image store. Content Manager requires Websphere and I4C (Information Integrator for Content) has been used to federate CM and DB2.

Scalability

Scalability is one of the primary goals of implementing e-DiaMoND's infrastructure.

Currently there is a problem utilising Content Manager's federation layer I4C which potentially leads to instabilities if any node is unavailable or has a slow responding time. The best practise to overcome these kind of problems is not to use proprietary federation software. The project has acknowledge that limitation and the need for a more flexible solution, and full deployment of e-DiaMoND will be based in web service based federation layer.

Furthermore the e-DiaMoND project use a Firewall/VPN to produce the illusion of a single subnet to the various machines participating in the e-DiaMoND grid. The additional firewall machines become redundant if the solution were to be deployed within the NHS firewall.

Cost

Cost is a problematic issue of developing new state-of-the-art technologies in all the research projects which are aiming for a real life full deployment. It is inevitable that the cost of deploying a system such as e-DiaMoND in all the 90+ BCU sites will be increased dramatically as the high profile requirements of the digitisation of the National Breast Screening Service are trying to be met.

The requirements for hardware increase dramatically the overall cost of the project. The balance between cost and performance is obviously a major consideration in the health arena, especially for a publicly funded organisation.

Another aspect of the overall cost of the project is the need of commitment by both the staff and management of the service as there will be a massive logistical task to convert the existing records into digital format. More importantly the scanning of x-rays will have to be done utilising suitable scanners, that give an appropriate spacial resolution as well as greyscale bit depth.

Another factor that will have an impact on the overall system cost is the availability of the system. To achieve high levels of availability different strategies depending on the locality of the system are required, as it is more difficult to ensure the availability of remote systems that are part of the grid than the availability of the local servers.

System administration

It is important to make the management of such 'complex' systems as easy as possible. This is partly to minimise the work load and partly to minimise the chance of mistakes being made. System administration it is especially important in the area of setting overall system security, including authentication and authorisation.

As a result, there will be a requirement for suitably qualified systems administrators to manage the overall system and to manage also the necessary of software patching as required.

Licensing

The licensing of the software need to be investigated as it is not yet clear enough whether it is allowable to run the applications on an arbitrary CPU in an arbitrary institution. Licensing constraint also depends on the number of the users that they are going to use each application and as such the definition of the user numbers is a necessity.

In addition, the meaning of 'site' license should be defined as this definition is quite vague when we are talking about grids. In the grid community it is also an open question what does the licensing term 'only run on 1 cpu' means in the context of grids as the software may be installed on a single machine within the grid, but many requests may come from anywhere within the grid.

Alternatively some software licences only allow for a single registered user to be active at any given time. This is workable within a small organisation with a limited number of people needing the facilities provided by the software. However it is more difficult in a VO where many users from distributed locations may need to utilise the software. Although more concurrent licences could be purchased, it is difficult to estimate the maximum number required. This would be especially true if the VO is dynamic.

The costs involved in acquiring the licences necessary is also worthy of consideration. This cost falls entirely on the member of the VO who provide access to this software. A method of charging will probably need to be developed to even out the impact of the upfront costs. Alternatively a user may have to register with the software supplier and each time he/she uses the software it is recorded. This would allow the charging of users directly by the software supplier. This is however unlikely and would have many issues that need resolving.

Monitor calibration

The transition of reading medical images from film and light boxes to utilising digital reading, requires great consideration regarding the selection and calibration of the equipment used. The ability to consistently read from the new flat panel monitors need to be considered for the choice of the equipment and research has to be undertaken to compare the effectiveness of various technologies currently on the market of conventional methods, namely CRT screens.

Taking breast screening as an example, the e-DiaMoND project utilised state-of-the-art T221 monitors (9 million pixel full colour) for its prototype applications. Through reviews with clinicians, it was evident that the display of a full resolution digital image could provide a clinician with a means of performing a reading using this new technology. However this process required particular attention to be paid in the scanning process for creating the digital image. Huge variances were seen between the scanners used, with the best results, in terms of optical density, achieved with a Multi-Array scanner. An additional consideration in moving to digital reading was the orientation of the screens. The preference from the clinical partners was to mimic the display orientation and layout used at present as this supported an easier transition for the users.

It is likely that a spacial resolution of better than $50\mu\text{m}$ and a greyscale depth of 12 bits or better will be needed to provide a suitable digital alternative to film. This however will only be verified by performing suitable research and large scale trials. It is therefore suggested that large scale research is performed to ascertain the real requirements for the effective spacial resolution necessary for the digital images to be equivalent to that of an x-ray. It is also necessary to demonstrate the necessary greyscale bit depth and what hardware is necessary to appropriately present this information to clinicians.

e-DiaMoND has also considered the requirements for maintaining calibration and display assessment when moving to digital reading. Digital monitors are hugely dependant on ambient conditions and without continual assessment, these screens will fail to provide the required results and thus will impact the ability of a clinician to make sound judgement on a case. This problem exists not just within the domain of Breast Screening, but across the NHS in general where there has been a move towards digital reading with the introduction of PACS systems.

The e-DiaMoND approach was to develop a display assessment test-bed and associated auto-calibration software to ensure a suitable performance fit to DICOM GSDF (Gray Scale Display Function) Standard calibration. This capability enables departments to continually monitor and assess the effectiveness of the digital screens within the working environment and warn departments when to take corrective action.

References

- [1] J. M. Brady, D. J. Gavaghan, A. C. Simpson, M. Mulet-Parada, and R. P. Highnam. e-DiaMoND: A Grid-enabled federated database of annotated mammograms. In F. Berman, G. C. Fox, and A. J. G. Hey, editors, *Grid Computing: Making the Global Infrastructure a Reality*, pages 923–943. Wiley Series, 2003.
- [2] HealthGrid. www.healthgrid.org.
- [3] HealthGrid White paper. <http://whitepaper.healthgrid.org>.
- [4] NHS Breast Screening Programme. <http://www.cancerscreening.nhs.uk/breastscreen/index.html>.
- [5] NHS Breast Screening Programme, Annual Review 2004. <http://www.cancerscreening.nhs.uk/breastscreen/publications/nhsbsp-annualreview2004.pdf>, 2004.
- [6] Two weeks for breast cancer screening, says Labour. <http://www.timesonline.co.uk/article/0,,19809-1574190,00.html>.
- [7] J. Esteve, A. Kricke, J. Ferlay, and D. M. Parkin. Facts and Figures of Cancer in the European Community. Technical report, International Agency for Research on Cancer, Lyon, France, 1993.
- [8] J. L. Young, C. L. Percy, and A. J. Asire. Surveillance, Epidemiology, End Results: Incidence and Mortality Data 1973–1977, 1981.
- [9] V. L. Shavers and M. L. Brown. Racial and Ethnic Disparities in the Receipt of Cancer Treatment. *Journal of the National Cancer Institute*, 94(5):334–357, 2002.
- [10] IARC. 7th handbook on cancer prevention, 2002.
- [11] P. B. Dean. Overview of breast cancer screening. In M. Doi, M. L. Giger, R. M. Nishikawa, and R. A. Schmidt, editors, *3rd International Workshop on Digital Mammography*, volume 1119 of *Excerpta Medica International Congress Series*, pages 19–26. Elsevier Science, 1996.
- [12] P. Forrest. Breast Cancer Screening. Report to the Health Ministers of England, Wales, Scotland and Northern Ireland. HMSO, London, UK, 1986.
- [13] R. G. Blanks, S. M. Moss, C.E. McGahan, M. J. Quinn, and P. J. Babb. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. *British Medical Journal*, 321:665–669, 2000.
- [14] C. B. Woodman, A. G. Threlfall, C. R. M. Boggis, and P. Prior. Is the three year breast screening interval too long? Occurrence of interval cancers in NHS Breast Screening Programme’s North West region. *British Medical Journal*, 310:224–226, 1995.
- [15] K Johnston and J Brown. Two view mammography at incident screens: cost effectiveness analysis of policy options. *British Medical Journal*, 319:1097–1102, 1999.
- [16] L. W. Bassett, B. Shayestehfar, and I. Hirbawai. Obtaining previous mammograms for comparison: usefulness and costs. *American Journal of Roentgenology*, 163:1083–1086, 1994.
- [17] A. J. G. Hey and A. E. Trefethen. The UK e-science programme and the grid. In *Proceedings of Computational Science (ICCS 2002), Part I*, pages 3–21. Springer-Verlag Lecture Notes in Computer Science, volume 2329, 2002.
- [18] UK e-Science Programme. <http://www.rcuk.ac.uk/escience>.
- [19] ESLEA, Exploitation of Switched Lightpaths for eScience Applications. <http://www.mb-ng.net/eslea>.
- [20] e DiaMoND consortium. e-DiaMoND: vision and critical success factors.
- [21] National Programme for IT in the NHS. <http://www.connectingforhealth.nhs.uk/>.
- [22] S. R. Amedolia, J. M. Brady, R. McClatchey, M. Mulet-Parada, M. Odeh, and T. Solomonides. MammoGrid: Large-Scale Distributed Mammogram Analysis. In *Proceedings of the XVIIIth Medical Informatics Conference (MIE 2003), St Malo, France*, volume 95 of *Studies in Health Technology and Informatics*, pages 194–199, 2003.
- [23] National Digital Mammography Archive. <http://www.i3archive.com>.

- [24] Axiopex data management and data sharing. <http://www.axiopex.org>.
- [25] CancerGrid. <http://www.cancergrid.org>.
- [26] CLEF: integrating information for the clinical e-Scientist. <http://www.clinical-esience.org>.
- [27] eFamily project. <http://www.efamily.org.uk>.
- [28] Equator Research Collaboration. <http://www.equator.ac.uk>.
- [29] Integrative Biology. <http://www.integrativebiology.ac.uk>.
- [30] Information eXtraction from Images (IXI). http://www1.imperial.ac.uk/medicine/about/divisions/cs/imagesci/imagephys_0/esience_0/default.html.
- [31] A. Knox. e-DiaMoND: architecture overview and functional model, 2003.
- [32] D. J. Power, E. A. Politou, M. A. Slaymaker, S. Harris, and A. C. Simpson. An approach to the storage of DICOM files for grid-enabled medical imaging databases. In *Proceedings of the 2004 ACM Symposium on Applied Computing*, pages 272–279, 2004.
- [33] I. Foster, C. Kesselman, J. Nick, and S Tuecke. The Physiology of the Grid: an Open Grid Services Architecture for Distributed Systems Integration. <http://www.globus.org/research/papers/ogsa.pdf>, 2002.
- [34] Open Grid Services Infrastructure (OGSI) Version 1.0. http://www-unix.globus.org/toolkit/draft-ggf-ogsi-gridservice-33_2003-06-27.pdf, June 2003.
- [35] DICOM Standard. <http://medical.nema.org/>.
- [36] The science of SMF. http://www.ctimi.com/portals/ctimi/cti_mirada/tech_licensing/smf_tools/mirada_smf_quantify.html?nv=nav_cti_mirada.
- [37] Open Grid Services Architecture–Data Access and Integration, OGSA-DAI. <http://www.ogsadai.org>, September 2004.
- [38] D. J. Power and E. A. Politou. e-DiaMoND data flow definition, 2003.
- [39] Globus Toolkit. <http://www-unix.globus.org/Globus>.
- [40] The NHS from 1998 to the Present. <http://www.nhs.uk/england/aboutTheNHS/history/1998toPresent.cmsx>.
- [41] NHS England. <http://www.nhs.uk/England/AboutTheNhs/Default.cmsx>.
- [42] NHS Wales. <http://www.wales.nhs.uk/page.cfm?pid=3331>.
- [43] NHS Scotland. <http://www.show.scot.nhs.uk/index.aspx>.
- [44] NHS Northern Ireland. <http://www.healthandcareni.co.uk>.
- [45] Data Protection Act 1998. <http://www.hmso.gov.uk/acts/acts1998/19980029.htm>, 1998. The Stationery Office Limited, London.
- [46] Medical ethics and law: confidentiality, data protection, Caldicott principles, computer use and patient records. <http://www.addenbrookes.org.uk/advice/medethlaw/confidential1.html>, 2003.
- [47] Human Rights Act 1998. <http://www.hmso.gov.uk/acts/acts1998/19980042.htm>, 1998. The Stationery Office Limited, London.
- [48] UK Biobank. <http://www.ukbiobank.ac.uk>.
- [49] NHS Information Authority in conjunction with The Consumers' Association and Health Which? Share with Care! Technical report, NHS, October 2002.
- [50] Dublin Core Metadata Initiative. <http://dublincore.org/>.
- [51] e-Government Interoperability Framework. <http://www.govtalk.gov.uk/schemasstandards/egif.asp>.
- [52] The e-GIF and the NHS — a policy statement. http://www.dh.gov.uk/PolicyAndGuidance/InformationPolicy/NationalITProgramme/NationalITProgrammeArticle/fs/en?CONTENT_ID=4054851&chk=TAOTak.

- [53] STEP — Standards Enforcement in Procurement. <http://www.nhsia.nhs.uk/step/pages/default.asp>.
- [54] Medical Imaging Strategy. <http://www.nhsia.nhs.uk/step/pages/usequest/p-mi.aspx>.
- [55] HL7 Strategy and Procurement Policy. <http://www.nhsia.nhs.uk/step/pages/usequest/p-hl7.aspx>.
- [56] SNOMED International. <http://www.snomed.org>.
- [57] SNOMED Clinical Terms. http://www.nhsia.nhs.uk/snomed/pages/ct_snomed.asp.
- [58] NHS Data Dictionary. <http://www.nhsia.nhs.uk/datastandards/pages/ddm/index.asp>.
- [59] BS7799. <http://www.standardsdirect.org/iso17799.htm>.
- [60] NHS patient privacy. http://www.theregister.co.uk/2003/02/11/nhs_patient_privacy_what_patient/.
- [61] Ross Anderson. Security in Clinical Information Systems. <http://www.cl.cam.ac.uk/users/rja14/policy11/policy11.htm>.
- [62] Security of Medical Information Systems. <http://www.cl.cam.ac.uk/users/rja14/#Med>.
- [63] Strategy for cryptographic support services in the NHS. http://www.nhsia.nhs.uk/nhid/pages/resource_informatics/pr/Cryptography%20strategy.pdf.
- [64] National Standard Pathology Reports Messaging. http://www.nhsia.nhs.uk/pathology/pages/secu_encryp.asp.
- [65] NHS PKI project in sick bay. http://www.theregister.co.uk/2002/09/19/nhs_pki_project_in_sick/.
- [66] PMIP PKI Closure. http://www.nhsia.nhs.uk/pathology/pages/pki_update150704.asp.
- [67] Arrangements for the transfer to Unencrypted Messages due to PKI Closure. http://www.nhsia.nhs.uk/pathology/pages/documents/Lab_guidelines_PKI_closure_v1-01.rtf.
- [68] Information Security Standards for NHS use. http://www.nhsia.nhs.uk/security/pages/security_standards.asp.
- [69] NHS Data Transfer Service (DTS) Strategy and Procurement Policy. <http://www.nhsia.nhs.uk/step/pages/usequest/p-dts.aspx>.
- [70] NHS Data Transfer Service (DTS). http://www.nhsia.nhs.uk/step/pages/usequest/r-links.aspx#nhsia_dts.
- [71] Health Level 7 and CDISC sign associate charter agreement. http://www.cdisc.org/news/news_03_16_2001.html.
- [72] IHE, Integrating the Healthcare Enterprise. <http://www.rsna.org/IHE/index.shtml>.
- [73] IHE Strategy. <http://www.nhsia.nhs.uk/step/pages/usequest/p-ihe.aspx>.
- [74] X-rays to be stored on computer (BBC). <http://news.bbc.co.uk/1/hi/health/3700381.stm>.
- [75] U. Schrader, E. Kotter, E. Pelikan, A. Zaiss, U. Timmermann, and R. Klar. Critical success factors for a hospital-wide PACS. *Proceedings of the 1997 Annual Fall Symposium*, pages 439-43, 1997.
- [76] PanLondon PACs. <http://www.krha.nhs.uk/PACS/Index.htm>.
- [77] N. Haramati. PACS and RIS: Approaches to Integration. *Journal of Healthcare Information Management*, 14(3):69, 2000.
- [78] The NHAIS (Exeter) System. <http://www.nhsia.nhs.uk/nhais/pages/default.asp>.
- [79] DICOM, Workstations and PACS. http://www.dclunie.com/papers/SPIE_20040217_Workstation.pdf.
- [80] JCIS in Addenbrookes. http://www.addenbrookes.org.uk/news/news2004/mar/achieve_040304.html.
- [81] UK Health department overhauls patient records advisers. http://www.infosecurity-magazine.com/news/040709_NHS.html.
- [82] BT and Accenture win £2.7bn NHS IT contracts. <http://www.computerweekly.com/articles/article.asp?liArticleID=1271>.
- [83] NHS wrestles with erratic demand for IT. <http://www.computerweekly.com/articles/article.asp?liArticleID=130912>.
- [84] GPs vote to boycott patient record database. <http://www.computerweekly.com/articles/article.asp?liArticleID=131577>.
- [85] NHS data security shambles. http://www.theregister.co.uk/2004/06/25/letters_2506/.

- [86] M. A. Slaymaker, E. Politou, D. J. Power, S. Lloyd, and A. C. Simpson. Security aspects of grid-enabled digital mammography. *Methods of Information in Medicine (to appear)*, 2004.
- [87] D. J. Power, E. A. Politou, M. A. Slaymaker, and A. C. Simpson. Towards secure grid-enabled healthcare. *Software Practice & Experience*, 2005.
- [88] Data Protection Act 1998. <http://www.hmso.gov.uk/acts/acts1998/19980029.htm>, 1998.
- [89] Computer Science and Telecommunications Board. *Trust in Cyberspace*. National Research Council, 1997.
- [90] I. Flechais, M. A. Sasse, and S. M. V. Hailes. Bringing security home: a process for developing secure and usable systems. In *ACM/SIGSAC New Security Paradigms Workshop*, August 2003.
- [91] M. A. Slaymaker, E. A. Politou, D. J. Power, S. Lloyd, and A. C. Simpson. e-DiaMoND: risk analysis. In *Proceedings of HealthGrid '04*, January 2004.
- [92] M. Humphrey and M. Thompson. Security Implications of Typical Grid Computing Usage Scenarios. In *Proceedings of the 10th International Symposium on High Performance Distributed Computing (HPDC)*, 2001.
- [93] Open Grid Services Architecture. <http://www.globus.org/ogsa/>.
- [94] OASIS WSRF TC. http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsrf, October 2004.
- [95] Open Grid Services Architecture - Distributed Query Processor, OGSA-DQP. <http://www.ogsadai.org/dqp/>, September 2004.
- [96] Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile. <http://www.ietf.org/rfc/rfc3280.txt>, April 2002.
- [97] Grid Security Infrastructure Documentation. <http://www-unix.globus.org/toolkit/docs/3.2/gsi/index.html>, October 2004.
- [98] Kerberos: The Network Authentication Protocol. <http://web.mit.edu/kerberos/www/>, October 2004.
- [99] The SSL Protocol Version 3.0. <http://www21.ocn.ne.jp/~k-west/SSLandTLS/draft302.txt>, November 1996.
- [100] The TLS Protocol Version 1.0. <http://www.ietf.org/rfc/rfc2246.txt>, January 1999.
- [101] S. Godik and T. Moses. eXtensible Access Control Markup Language (XACML) version 1.1, committee specification. <http://www.oasis-open.org>, August 2003.
- [102] OASIS Security Services TC. http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security, October 2004.
- [103] D. W. Chadwick. RBAC policies in XML for X.509 Based Privilege Management. In *Proceedings of SEC 2002*, 2002.
- [104] M. Meyers. RFC 2560 - X.509 Internet Public Key Infrastructure Online Certificate Status Protocol - OCSP. <http://www.faqs.org/rfcs/rfc2560.html>, June 1999.
- [105] Extensible Markup Language (XML). <http://www.w3.org/XML/>. W3C.
- [106] M. Gudgin, M. Hadley, N. Mendelsohn, J. J. Moreau, and H. Nielsen. SOAP Version 1.2 Part 1: Messaging Framework. <http://www.w3.org/TR/soap12-part1/>, June 2003.
- [107] E. Christensen, F. Curbera, G. Meredith, and S. Weerawarana. Web Services Description Language (WSDL) 1.1. <http://www.w3.org/TR/wsdl>, March 2001.
- [108] T. Bellwood, L. Clément, and C. von Riegen. UDDI Spec Technical Committee Specification, Version 3.0.1. Technical report, OASIS, October 2003.
- [109] OASIS: Organization for the Advancement of Structured Information Standards. <http://www.oasis-open.org>.
- [110] W3C. <http://www.w3.org>.
- [111] WS-I: Web Services Interoperability Organisation. <http://www.ws-i.org>.

- [112] Hao He. What is Service-Oriented Architecture? *O'Reilly WebServices.XML.com*, September 2003.
- [113] SOAP Version 1.2 Part 0: Primer. <http://www.w3.org/TR/2003/REC-soap12-part0-20030624/>.
- [114] RDF Rich Site (RSS). <http://www.oasis-open.org/cover/rss.html>.
- [115] http://searchwebservices.techtarget.com/ateQuestionNResponse/0,289625,sid26_cid921266_tax301569,00.html.
- [116] SOAP Messages with Attachments. <http://www.w3.org/TR/SOAP-attachments>.
- [117] XOP: XML-binary Optimized Packaging. <http://www.w3.org/TR/2005/REC-xop10-20050125/>.
- [118] Taming the XML beast. <http://www.networkworld.com/columnists/2005/011005kobielus.html>.
- [119] ebXML specifications. <http://www.ebxml.org/specs/>.
- [120] Equator Research Collaboration. <http://edodds.blogs.com/conmergence/ltebsoagt/index.html>.
- [121] Health Level Seven. <http://www.hl7.org/Press/20040427b.asp>.
- [122] <http://lists.ebxml.org/archives/ebxml-dev/200403/msg00046.html>.
- [123] <http://lists.xml.org/archives/xml-dev/200407/msg00135.html>.
- [124] David Webber. The Benefits of ebXML for e-Business. <http://www.idealliance.org/proceedings/xml04/papers/44/webber>
- [125] National Programme for IT in England. http://www.sunshine-healthcare.org/docs/IT_in_England.pdf.
- [126] M. Clark, P. Fletcher, J. J. Hanson, R. Irani, M. Waterhouse, and J. Thelin. *Web Services Business Strategies and Architectures*. Wrox Press, August 2002.
- [127] Web Services Notification. http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsn.
- [128] WS-Addressing. <http://www.w3.org/Submission/ws-addressing/>.
- [129] A. Nadalin, C. Kaler, P. Hallam-Baker, and Monzillo R. Web Services Security:SOAP Message Security 1.0 (WS-Security 2004). <http://docs.oasis-open.org/wss/2004/01/oasis-200401-wss-soap-message-security-1.0.pdf>, March 2004. OASIS.
- [130] Web Services Transactions specifications. <http://www-128.ibm.com/developerworks/library/specification/ws-tx/>.
- [131] Business Process Execution Language. http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsbpel.
- [132] Web Services Choreography Interface. <http://www.w3.org/TR/wsci/>.
- [133] Global Grid Forum. <http://www.ggf.org>.
- [134] Globus project. <http://www.globus.org>.
- [135] Web Services Resource. <http://docs.oasis-open.org/wsrf/2004/11/wsrf-WS-Resource-1.2-draft-02.pdf>.
- [136] Web Services Resource Properties 1.2. <http://docs.oasis-open.org/wsrf/2004/06/wsrf-WS-ResourceProperties-1.2-draft-04.pdf>.
- [137] Web Services Resource Lifetime 1.2. <http://docs.oasis-open.org/wsrf/2004/06/wsrf-WS-ResourceLifetime-1.2-draft-03.pdf>.
- [138] Web Services Service Group 1.2. <http://docs.oasis-open.org/wsrf/2004/06/wsrf-WS-ServiceGroup-1.2-draft-02.pdf>.
- [139] Web Services Base Faults 1.2. <http://docs.oasis-open.org/wsrf/2004/06/wsrf-WS-BaseFaults-1.2-draft-02.pdf>.
- [140] CondorG. <http://www.cs.wisc.edu/condor/condorg>.
- [141] UNICORE. <http://unicore.sourceforge.net/index.html>.
- [142] EGEE. <http://public.eu-egee.org/intro/>.

- [143] GEANT. <http://www.geant.net>.
- [144] gLite. <http://glite.web.cern.ch/glite/>.
- [145] Alien Grid. <http://alien.cern.ch>.
- [146] The ALICE Portal. <http://pcaliweb02.cern.ch/NewAlicePortal/en/index.html>.
- [147] OMMI. <http://www.omii.ac.uk>.
- [148] Legion project. <http://legion.virginia.edu/index.html>.
- [149] A. Grimshaw, W. A. Wulf, and the Legion team. The Legion Vision of a Worldwide Virtual Computer. *Communications of the ACM*, 40(1):39–45, January 1997.
- [150] Avaki. <http://www.avaki.com/>.
- [151] Web Services Strategy and Procurement Policy. <http://www.nhsia.nhs.uk/step/pages/usequest/p-wss.aspx>.
- [152] Case Study: UK National Health Service NPfIT Uses ebXML Messaging. http://www.ebxml.org/case_studies/NHS-ebMSG-casestudy-041206.pdf.
- [153] The future of screening (BBC). <http://news.bbc.co.uk/2/hi/health/2570787.stm>.
- [154] J. M. Lewin, R. E. Hendrick, C. J. D'Orsi, P. K. Isaacs, L. J. Moss, A. Karellas, G. A. Sisney, C. C. Kuni, and G. R. Cutter. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. *Radiology*, 218:873–880, 2001.
- [155] Personal Performance in Mammographic Screening (PERFORMS). <http://ibs.derby.ac.uk/performs/introduction.shtml>.
- [156] D. C. Anderson, J. Campos, M. Hartswood, M. Jirotko, L. Khoo, R. Procter, R. Slack, L. Smart, J. Soutter, P. Taylor, and L. Wilkinson. The role of computer based training in mammography. In *Royal College of Radiologists Breast Group Annual Scientific Meeting*, November 2003.
- [157] J. Soutter, J. Campos, M. Hartswood, M. Jirotko, R. Procter, R. Slack, and P. Taylor. Grid-based mammography training. *Hospital Radiologist*, 5(6), 2003.