Designing for e-Health: Recurring Scenarios in Developing Grid-based Medical Imaging Systems

John Geddes[a], Clare Mackay[a], Sharon Lloyd[b], Andrew Simpson[b], David Power[b], Douglas Russell[b], Marina Jirotka[b], Mila Katzarova[b], Martin Rossor[c], Nick Fox[c], Jonathon Fletcher[c], Derek Hill[d], Kate McLeish[d], Yu Chen[d], Joseph V Hajnal[e], Stephen Lawrie[f], Dominic Job[f], Andrew McIntosh[f], Joanna Wardlaw[g], Peter Sandercock[g], Jeb Palmer[g], Dave Perry[g], Rob Procter[h], Jenny Ure[h,1], Mark Hartswood[h], Roger Slack[h], Alex Voss[h], Kate Ho[h], Philip Bath[i], Wim Clarke[i], Graham Watson[i]

[a]Department of Psychiatry, University of Oxford, [b]Computing Laboratory, University of Oxford, [c]Institute of Neurology, University College London, [d]Centre for Medical Image Computing (MedIC), University College London, [e]Imaging Sciences Department, Imperial College London, [f]Department of Psychiatry, University of Edinburgh, [g]Department of Clinical NeuroSciences, University of Edinburgh, [h]School of Informatics, University of Edinburgh, [i]Institute of Neuroscience, University of Nottingham

Abstract. *The paper draws on a number of Grid projects, particularly on the experience of NeuroGrid,, a UK project in the Neurosciences tasked with developing a Grid-based collaborative research environment to support the sharing of digital images and patient data across multiple distributed sites. It outlines recurrent socio-technical issues, highlighting the challenges of scaling up technological networks in advance of the regulatory networks which normally regulate their use in practice.*

Keywords. E-Health, medical imaging, neuroscience, problem scenarios in distributed data-sharing, socio-technical system design

1. Introduction

There is an increasing drive within the UK to integrate healthcare data and services. The vision of 'joined-up' healthcare envisages services being delivered to patients through flexible – and perhaps virtual – organisational structures formed around networks of healthcare professionals working within, and across, multiple service units and administrative domains. Similarly, translational medical research focuses on reducing the turn-around time in the cycle that leads from identification of possible causes of illness (for example, particular genetic and/or environmental factors) to the

[1] Corresponding Author: Jenny Ure, School of Informatics, University of Edinburgh, Jenny.Ure@ed.ac.uk

investigation of disease mechanisms and the development of treatments, through to clinical trials and practice [1]. The realisation of this agenda is constrained by a range of recurrent issues and problem scenarios that have been given priority as one of the e-Science Grand Challenges. We discuss some of these issues in relation to the development of two Grid projects using distributed digital imaging and patient data across distributed sites.

1.1. Health Services

Healthcare services and research infrastructure in the UK and Europe are in a process of transition. The vision of translational and evidence-based medicine depends on a seamless infrastructure from lab-based research results to clinical applications. The reality however, is a patchwork of disjoint technical, professional and administrative architectures, a diversity of criteria and clinical protocols for data acquisition, a range of coding standards and differing guidelines for clinical practice and trial management. Furthermore, e-Health initiatives to streamline services, such as electronic patient records, are already generating debate over issues associated with the cost, benefits, quality and dependability of these services, the potential implications for patient confidentiality, and the potential risks in clinical applications. The collective consequence of these factors is that even modest levels of system and information integration have proved difficult to achieve in practice in healthcare [2], [3].

2. NeuroGrid

NeuroGrid[2] is a three-year, £2.1M project funded through the UK Medical Research Council to develop a Grid-based collaborative research environment for imaging in large scale studies for neuropsychiatric disorders in the UK. It will be developed around three component clinical exemplars in stroke, dementia and psychosis, and complex services for quantitative and qualitative image analysis. This project, which started in March 2005, has a project team distributed across 11 sites in the UK, bringing together the work of clinicians, clinical researchers and e-scientists at Oxford, Edinburgh, Nottingham and London, using a Grid-based architecture to address the different needs of each node.

2.1. Objectives

The project aims to enable rapid, reliable and secure data sharing through interoperable databases, with access control management and authentication mechanisms. It will also provide a toolset to facilitate image registration and analysis, normalization, anonymisation, real-time acquisition and error trapping, to improve reliable diagnosis, to compensate for scanner differences and to allow quality and consistency checks before the patient leaves the imaging suite.

The exemplar teams will use the infrastructure to address issues specific to their own domain of interest, as well as generic issues in the design of distributed Grid-based systems, and the aggregation and use of data in multi-site clinical trials.

[2] www.neurogrid.ac.uk

The requirements of the three clinical exemplars groups (stroke, psychosis and dementia) use the potential of the Grid in very different ways – from the creation of enhanced datasets for rare conditions, to the use of Grid-enabled tools for image acquisition, archiving and analysis, and to the analysis of variance in technical and human processes associated with data collection, curation, processing and uploading. This provides a range of opportunities for evaluating the potential of Grid-based applications in the neurosciences, as well as more generally in e-Health and eScience [4]. Many of the issues addressed in the paper have been mirrored in other UK HealthGrid projects, most notably eDiaMoND[3] [5, 6, 7, 8] which was a flagship pilot UK e-Science project on medical imaging in the context of breast screening, funded through EPSRC/DTI and IBM SUR grants to build a grid-enabled, federated database of annotated, digitised mammograms and patient information intended to aid research, into and detection and treatment of breast cancer.

3. Recurring Scenarios

We will discuss a range of issues which appear to be significant hurdles in the vision of e-Health, including translational research and large-scale clinical trials, using a number of prototype grid-enabled applications to exemplify recurrent problem scenarios. Many of the issues are arguably evident in other distributed networked systems in e-Business and e-Learning, for example, where scaling up of technical architectures has not been matched by a corresponding alignment of the local coordination and governance structures in heterogeneous and distributed local communities. Although we will draw on other projects, the focus is on those issues which have been most prominent in NeuroGrid in the first year:

- Aggregating data collected in different ways, for different purposes, from very diverse and distributed contexts
- Representing this data in ways which are meaningful and useful to communities with very different aims and frames of reference
- Managing clinical trials and associated ethical permissions and protocols across multiple communities, and for multiple purposes
- Aligning local aims and requirements with collective ones
- Aligning technical and human networks to advantage
- Integrating the technical work of system building, with the socio-political work of generating collective structures and agreements for the governance of the new risks and opportunities generated

4. Issues in Grid-based Medical Imaging

Radiological imaging in large–scale clinical trials promises substantial benefits in the diagnosis and assessment of specific treatment effects on pathological processes. The Grid offers a mechanism for further extending the size of datasets available for analysis, and for enhancing the speed and quality of analysis that can be performed on

[3] www.ediamond.ox.ac.uk

them. Researchers use innovative imaging techniques to detect features that can refine a diagnosis, classify cases, track normal or often subtle patho-physiological changes over time and improve understanding of the structural correlates of clinical features. Some of the variance is attributable to a complex variety of procedures involved in image acquisition, transfer and storage, and it is crucial, but difficult, for true disease-related effects to be separated from those which are artifacts of the process. There are two basic approaches to the extraction of detailed information from imaging data invoking different sets of challenges:

- Automated and computationally intensive image analysis algorithms for quantification and localization of signal differences have particular value in longitudinal imaging studies of change over time. This is particularly useful in identifying changes associated with the onset of psychosis, dementia or Alzheimer's disease, but particularly challenging in the harmonization of technical processing – as in the use of different scanners for example.
- Assessment by healthcare professionals, as in large randomised controlled trials or observational studies, uses imaging to distinguish between different underlying causes (e.g., stroke or psychosis can both be associated with similar behavioural presentation), to assess severity, progression or response to treatment, and may require collection, storage and dissemination of data from hundreds of centres. A particular challenge here is the intra and inter-site variance across raters.

Imaging research is traditionally often carried out in small studies in single research centres, where much of the knowledge about provenance, reliability and use is grounded in shared local knowledge, aims and contexts. Researchers and clinicians share an intimate understanding of the potential risks of combining local datasets for clinical purposes, based on a knowledge of the protocols and processes that could have contributed to the outcomes – which scanner, which control group, which protocol, etc, Scaling up technical systems has, in practice, been easier that scaling up the socio-technical and socio-political processes governing the collection, analysis, representation and use of data outwith it's context of origin[9].

As with the introduction of networked technology in education, new possibilities and new responsibilities associated with governance and use in practice have led to reconsideration of the nature of the processes and purposes of e-Health and e-Science systems, and the roles and responsibilities of the stake-holding entities within this [10]. The realization of a sustainable and reliable system will depend on bridging the gap between the vision of seamless integration and the more disjoint reality on the ground highlighted in the recent Healthgrid White Paper [11].

5. Data Quality Issues

The large scale aggregation of diverse datasets offers both potential benefits and risks, particularly if the outputs are to be used with patients in a clinical context. Thus aggregating data is a key issue for e-Health, yet data is not independent of context in which it is generated. Within small communities of practice a degree of shared and updated knowledge and experience allows judicious use of resources whose provenance is known and whose weaknesses are often already transparent. The same is

not true of aggregated data from multiple sources, where the process of deriving and coding may vary in both explicit and less obvious ways, even within communities of practice.

One approach is to make early use of prototypes to provide a 'sandpit' for promoting both technical and inter-community dialogue and engagement, and start the process of identifying, sharing and updating knowledge of emerging issues. The approach in Neuro Grid has been to focus early on trials with known datasets to generate an awareness of the types of variance that can arise and ways in which it might be minimized, harmonized, or made transparent to users given the ethical implications of use in the clinical domain. This will include technical differences between scanners, differences in use of protocols or in data input, differences in rating of images where these are not automated and differences in the administration of psychiatric tests such as the PANNS[4] test.

5.1. Data Collection: the Dementia and Stroke Exemplars

Multi-site clinical trials add additional complexity with the need to coordinate such issues as naming conventions for files, patient clinical trial ID management and acquisition parameters. The NeuroGrid dementia exemplar group involves researchers from the Institute of Neurology in London, and from University College London, who aim to use the Grid infrastructure to collect a new dataset and to develop methods of measuring image quality whilst the patient is still in the scanner, such that adjustments can be made in real time while the patient is still available, thus cutting the cost of and delay in re-scanning and optimizing the reliability of the dataset. Data being collected includes baseline demographic data (age, gender, but not any identifiable data), digital scans for each of the time-steps and outcome information about these cases, associated with each timestep. Data curation involves documenting the acquisition, processing, archiving, retrieval, aggregation and use of this medical data.

Working across sites and databases has highlighted how differences and mismatches occur, for example, in matching patient data to images or in labeling sequences of scans, and in some cases where staff fill in forms incorrectly. While there are regulatory requirements for good clinical practice and elaborations on these,[5] the ways in which a particular trial can tackle these problems remains to be worked out, and there is considerable uncertainty as to how regulatory requirements can be effective when translated into practice. Aggregating multiple datasets, in the e-Health context thus has implications for both accuracy and clinical diagnosis and treatment. In practice, a number of small scale responses are beginning to emerge in different nodes

- One group has adopted the use of tablet PCs for clinical staff to input data, using a wireless link to the relevant database, so that mismatches these can be rectified at the point of input, using the functionality of Microsoft Infopath to highlight mismatches when cross referenced to the database.
- The Stroke exemplar group in Edinburgh and Nottingham are developing error-management software that uses multiple measures of triangulating in on

[4] The Positive and Negative Syndrome Scale, (PANSS), is a 30 item assessment of positive and negative psychiatric symptoms

[5] EU Clinical Trials Directive, Directive 2001/20/EC

patient data to query mismatches between images and patient records from the multiple acquisition sites.

- The Psychosis exemplar group have generated harmonisation software for differences between scanners and Grid-enabling these algorithms will allow sharing of this across sites. Studies will also be done on the interpretation and use of clinical tests to identify a measure of the variance that can be expected as a result of differences with and between clinicians in the diagnosis of psychosis.

Part of the benefit of an early prototype is the opportunity to run trials to identify the parameters of variation across sites under different constraints and conditions, and use test data sets to evaluate the quality, validity and reliability of aggregated datasets. The ability to distinguish clinically significant differences in images from those that are artefactual is critical for the success of NeuroGrid, and early testing of the prototype will allow early engagement with this issue. The psychosis exemplar group in Edinburgh and Oxford will be testing software for harmonization across scanners at different sites. The dementia exemplar group will be evaluating the quality and speed of processing using a Grid-enabled toolkit, and the stroke exemplar group in Edinburgh and Nottingham, will utilise existing datasets to test Grid-enabled software for use with images from CT and MR scans, as well as gathering measures of variance in the rating of CT Scans within and between sites.

As problems have arisen, it has become increasingly clear that many are common to other e-Health and e-Science projects, and a range of emerging solutions and practical workarounds is being shared through an informal brokerage between active players within the Grid community.

5.2. Variance across nodes

As indicated earlier, 'joined up' systems face a range of known and unknown or unanticipated sources of variance in technical equipment, data acquisition, processing and curation, and also in human rating of images and of patient symptoms. Within NeuroGrid, the psychosis exemplar provided numerous opportunities to observe what Duguid and Brown [12] have called the 'social life of information' but has also provided interesting insights into the ways in which the technical and the human contribution of variance in data often becomes evident only in discussion with known datasets in real trials;

5.3. The role of informal dialogue

The use of ethnographic studies of clinical research practice has been a key part of our approach to understanding NeuroGrid requirements. Our findings suggest that awareness of data quality issues often only comes as a result of real community interaction within a co-located community [13]. This is hard to emulate in transient virtual organizations of the kind envisaged in eHealth and eScience, yet incidental observations suggest that many key observations on data quality were dependent on informed exchanges, from different experts, where knowledge from different domains came into play in relation to a specific problem.

Informal conversations between researchers in one example generated an awareness that the same protocol on the same image set had resulted in different

outcomes. Further discussions narrowed this down to differences in interpretation of a protocol, where tracing inside, or outside of a line resulted in volume differences. Discussion of a known dataset, in a known context, appears to help foreground anomalies, and improve data quality in ways that are hard to scale up. It also became evident from similar face-to-face discussions, again focusing multiple specialists on a shared problem, that aggregating data from sites with different demographic profiles was another source of variance, since brain shape is known to vary across ethnic groups, adding another dimension of variation within aggregated scan sets.

5.4. Involving stake-holding users

There is a push to improve data quality throughout the UK National Health Service (NHS) and, specifically, to improve the quality of data for auditing. Auditors routinely access various source of data, then combine and triangulate them to improve the quality of data that they extract. Similarly, researchers make use of data extraction forms designed specifically to capture the data needed for epidemiological studies and research nurses exercise considerable skill in ensuring that the data they gather is fit for the intended purposes. On the scale of aggregation entailed in Grid-based systems, there is arguably a need for a wider awareness of the issues in aggregating from multiple sources and an emphasis on strategies that can be adopted at different stages in acquisition, mining and use, so as to safeguard quality and reliability for use in clinical contexts by frontline staff. One interesting development in this regard is the potential for more active engagement of patients themselves as stakeholders in the use and updating of their medical records [14]. The leverage of end-user communities as stakeholders in maintaining the accuracy or currency of the process is one which is associated with real benefits and cost savings in e-Business [15] and may have some application in the context of medical informatics. In terms of system design, the work of Reddy et al [2], and Dourish and Bellotti [16] suggest that clinical staff using e-Health systems can make sense of, and coordinate work better if the system affords some degree of transparency about the activities of other users, and provides a context for coordinating information and planning across a distributed group.

5.5. Making sense of distributed data

The potential volume of data that can be aggregated via HealthGrids not only has implications for curation and quality but also for its interpretation by both humans and machines. Nonaka [17] highlights the importance of early articulation of shared frames of reference and situated contexts for envisaging and structuring the process collectively, by providing real or virtual opportunities for dialogue and exchange. In the more distributed context of the Grid, linking social and technical networks on an exceptionally large scale, there is increasing interest in the use of metadata and ontologies to formalise some elements of these shared frames of reference in human and machine readable form [18, 19]. Part of the motivation for this is that it affords automation of resource discovery and analysis, but the question remains as to whether formal descriptions can be sufficiently rich and expressive to model relationships between data providers and users. In this there is a trade-off between the benefits of share-ability and knowledge discovery across multiple datasets on the one hand, and the setting in stone of concepts and relationships which are constantly evolving. Our ability to anticipate the sorts of uses which might be made of data in the future, or other

ontologies with which they may be related, is time de-limited. As in many other contexts, there is a trade-off between speed, accuracy, validity and usability for particular purposes. As with the aggregation of multiple data sets discussed earlier, there are also aggregations of artefactual differences whose implications may be invisible to the user, but represent a potential risk in clinical use.

As diverse medical datasets come online in related domains and at different scales, the alignment of ontologies becomes a challenge. In the context of neuroscience, for example, there are datasets at different levels of granularity as well as in different modalities. The work of Sporns [20] highlights the extent to which imaging can be done at very different levels of granularity, and that the value of much of the research now ongoing will be in the integration of cross referenced data that can elucidate the structure and dynamics of the brain at very different levels of granularity, such as:

- MR images of structural changes in the brain using CT, PET or SPECT scans
- diffusion tensor imaging studies on the micro-structural development of white matter in the brain underpinning activation patterns detected in MR imaging
- genetic datasets associated with susceptibility to these disorders

The Human Brain project [21] addressed this issue early on in the context of collaboration with multiple groups, generating a reference ontology based on a Foundational Model of Anatomy (FMA) which allows diverse datasets, at different levels of granularity, to be aligned in a meaningful context for different purposes.

5.6. Aligning Competing Requirements

Many of the most intractable issues in integrated systems reflect the locally grounded nature of coordination and governance structures. Ethics and Security requirements were among the most recurrent issues encountered in NeuroGrid and eDiaMoND, and are one of a wide range of areas where there has been a tension between the requirements of distributed local groups.

5.6.1. Security vs access requirements

Common to all NeuroGrid exemplars is the need to determine secure and effective ways to aggregate and manage clinical trials data. The data takes the form of medical images and coded or descriptive information from patients who have consented to take part in trials and whose records contain material that is often highly sensitive in nature. The retrieval and access of this data requires new architectures to support the secure sharing of the data, both records and medical images. In the case of NeuroGrid, this also includes issues of anonymisation of faces in brain scans, given the potential in some formats, for reconstruction of facial volumes.

Within NeuroGrid, the exemplar groups need to run algorithms on other datasets that they do not own, and retrieve the results of this analysis; however, they do not receive the *original* data. In the case of scans of patients at risk of early-onset psychosis, direct access to the images is regarded as too sensitive and the solution agreed is to provide parametric statistical mappings of the original image data on which algorithms could be run, rather than the original. This adds some complexity to the workflows and the design as a whole, but aligns the competing requirements of the different stake-holding groups in a way which could be replicated to resolve this issue

elsewhere. Given the long term aims of translational medicine as a sustainable enterprise and the participation of commercial partners in clinical trials, both the architecture and the perception of security in Grid systems remains a critical issue [22].

5.6.2. Ethical Requirements

Issues such as ethical consent, IPR, and the development and implementation of shared protocols and administrative processes, challenge the local structures and *in situ* realization of coordination in distributed communities. Scaling up these less tangible architectures is a design issue of a more socio-political nature which has implications for how and if the e-Science vision can be implemented. While distributed, networked projects increasingly acknowledge the impact of human factors, the extent to which they can impede project realisation and the extent to which project work revolves around them is often under-estimated at the outset. By way of example, a recurrent barrier to the vision of e-Health is the difficulty of achieving agreement on ethical consent for use and/or re-use of patient data: neither NeuroGrid nor eDiaMoND are exceptions to this.

The eDiaMoND project was required to demonstrate the use of a grid-enabled digital mammography system. To prove the concept, it was necessary to consider the use of real data in real breast screening units, hospitals and research environments. This entailed managing an intricate arrangement of policies governing the use of patient data (e.g., research ethics review). In addition to delays, constraints and complications, data generated from research and re-used for subsequent clinical work does not have clear ownership. In addition, there are often constraints on linkage between research and clinical infrastructures including links between healthcare service and university networks [7].

The vision of translational research is to quicken the process between bench science and the delivery of healthcare to patients. In practice, however, transient virtual collaborations of the kind envisaged in e-Science lack either the formal infrastructure of contractual agreements evident in business supply chains, or the established norms and agreements that are generated in well established communities of practice. It may be that technical infrastructures scale up more easily than the socio-political and administrative infrastructures of the communities in which they must be embedded and used.

5.6.3. Aligning technical vs user criteria and requirements

Aligning requirements between distributed exemplar groups within a Grid project is one challenge, however, it is also the case that the stakeholders have competing aims and criteria. As the scale and scope of systems in the extended enterprise has grown, the difficulties of aligning aims and understanding across interdependent communities have become more critical, the interdependence of social and technical knowledge has become more apparent, and the tension between local and global requirements has become more problematic.

A recent overview of system design in business contexts [6] suggests that technologists' criteria for success are early closure on requirements, and adherence to time and cost constraints, with a robust design, while business managers criteria were, conversely, in favour of an evolving process that met a range of changing needs in a flexible way, and were not concerned about the cost, timescale or the design issues from a technological point of view. It is easy to see in this context how outcomes

satisfactory to one team might not meet the criteria of other stake-holding groups. This pattern was also evident in the eDiaMoND project, where the very different criteria, and aims of the technical and user communities significantly shaped the way this played out. The approach of the NeuroGrid team is to foster, where possible, a collaborative and participatory approach to design [23, 24] based on evolution from a very early prototype, around which system design could evolve in stages, from the basic need to share images which is core to all the exemplar groups.

6. Conclusions

We have discussed a range of scenarios that can be found across the HealthGrid community, ranging from the issue of aggregating heterogeneous, distributed datasets to the issues of scaling up local processes, protocols and coordination and consent structures. The most intractable of these have their roots in the coupled, socio-technical nature of infrastructural systems, and the difficulties inherent in scaling up information and communication networks in the absence of a corresponding architecture for coordination at a social, organisational, professional and political level.

6.1. Working up socio-technical arrangements

The concept of the collaboratory is central to the e-Science vision, yet there has been limited concern with the generation of the community and coordination infrastructures which will coordinate and sustain it. The experience of virtual business organisations in the context of the business supply chain suggest that the explicit management of the socio-technical whole is central to the success (or the failure) of collaboration. The e-Health vision – particularly in relation to translational medicine – embodies much of the supply chain concept and appears to be facing some of the same socio-technical challenges [8] [1]. NeuroGrid, like eDiaMoND, brings together disparate groups of clinicians, technologists and researchers with no prior working experience of large scale collaborative research or with the other project members. The technical work of system building is paralleled by the need to facilitate the generation of new structures and agreements for the governance of the new risks and opportunities generated when data is aggregated in this way.

The creation of real and virtual ‘shared spaces’ [17] on NeuroGrid included an early prototype as a ‘sandpit’ for engagement in areas of shared professional concern, as means of supporting this new hybrid community develop its own rules of engagement, and start making collective sense of local knowledge and requirements in relation to common project goals. Common challenges are coming into focus across the exemplar groups, and further collaboration will be encouraged through the use of workshops and special interest activity to resolve common issues in areas such as data quality, security, data ownership, confidentiality, IPR, ethics, and the management of clinical trials.

6.2. Dealing with Data Quality

In organic communities, the process of structuring collaboration, coordination and control structures happens as a matter of course, played out in shared contexts where

aims, terms and frames of reference are already well established. NeuroGrid is employing a simple early prototype to generate engagement and dialogue between partners, to enable earlier discussion of requirements for more complex services, compute capability and workflows, as well as data quality and configurational issues.In addition to ameliorating the recurring issue of requirements 'creep' late in the design process, it allows the disparate groups to discuss issues and possible actions in relation to a shared context.

Given that, in reality, many Grid-based collaborations are transient, and often led by funding considerations rather than a clear consonance of aims across participating groups, system design and management will increasingly rely on the creation of coordinating infrastructures – social, legal, ethical and professional. The recurring nature of problem scenarios in HealthGrid projects suggests that community building strategies such as early prototyping will be increasingly central to the realisation of the e-Health vision, and that further research is needed to both (a) identify the ways in which some of this may be integrated into the process of co-designing such systems, and (b) share strategies for designing technical information and communication systems more effectively around human ones [25].

Virtual organizations (VOs) such as these require a strategy for the negotiation of shared terms, processes, costs, risks and benefits, as well as the definition of those which are to remain local and the nature of the alignment between the two [26]. Collaboration across communities of interest depends heavily on finding practical ways of ensuring early engagement and dialogue, in areas of shared concern, so that the negotiation of diverse aims and requirements can inform the design process as early as possible [27].

7. Acknowledgements

The authors wish to thank other members of the NeuroGrid consortium for their input into, and comments on the work described in this paper, and in particular the role of the scientific collaborators in determining the requirements for this project. The authors wish to acknowledge the support provided by the funders of the NeuroGrid project: The MRC (ref no: GO600623 ID number 77729) and also the support of the UK e-Science programme.

References

[1] UK Clinical Research Collaboration Report www.ukcrc.org

[2] Hartswood, M., Procter R., Rouncefield, M., Slack R., Soutter, J. and Voss, A. (2003). 'Repairing' the Machine: A case study of evaluating computer-aided tools in breast screening. In Kuutti et al (Eds) *Proceedings of the Eighth European Conference on Computer-supported Cooperative Works*. pp 375-394.

[3] Ellingsen, G. and Monteiro, E. (2001). A patchwork planet: The heterogeneity of electronic patient record systems in hospitals. In Proceedings of the Information Systems Research Seminar in Scandinavia, Sweden, August.

[4] Geddes J., et al. (2005). NeuroGrid: Using Grid Technology to Advance Neuroscience, 18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05), pp. 570-572.

[5] Brady, M., Gavaghan, D., Simpson, A., Highnam, R. and Mulet, M. (2002). eDiaMoND: a grid-enabled federated database of annotated mammograms. Chapter 39 of Grid Computing: Making the Global Infrastructure a Reality, Wiley.

[6] Lloyd, A., Ure, J., Cranmore, A., Dewar, R. and Pooley, R. (2002). Designing Enterprise Systems. In Jardim-Goncalves R. Roy R. and Steiger Garcao A. (eds.) *Advances in Concurrent Engineering*. Swets and Zeitlinger, Lisse.

[7] Hartswood, M., Jirotka, M., Procter, R., Slack, R., Voss, A. & Lloyd, S. (2005). Working IT out in e-Science: Experiences of Requirements Capture in a HealthGrid Project. In Solomonides, T., McClatchy R., Breton V., Legre, Y. & Norager, S. (eds.) *From Grid to HealthGrid*. IOS Press ISSN 0926-9630.

[8] Jirotka, M., Procter, R., Hartswood, M., Slack, R., Coopmans, C., Hinds, C. and Voss, A. (2005). Collaboration and Trust in Healthcare Innovation: the eDiaMoND Case Study. Journal of Computer-Supported Cooperative Work, 14(4), p. 369-389.

[9] Reddy, M., Pratt, W., Dourish, P. and Shabot, M.M. (2003). Sociotechnical Requirements Analysis in Clinical Systems. *Methods Inf Med 2003; 42: pp 437-44*.

[10] Buetow, K. H. (2005). Cyberinfrastructure: Empowering a 'Third Way' in Biomedical Research. Science 308, May 6. pp 821-824.

[11] Breton, V., Dean, K. and Solmonides, T. (2005). The HealthGrid White paper, In Solomonides, T., McClatchy R., Breton V., Legre, Y. & Norager, S. (eds.) *From Grid to HealthGrid*. IOS Press ISSN 0926-9630.

[12] Duguid, P. & Brown, J.S. (2000). *The Social Life of Information*, Boston, MA: Harvard Business School Press.

[13] Hartswood, M, Procter, R., Rouncefield, M. and Slack, R. (2003). Making a Case in Medical Work: Implications for the Electronic Medical Record. *Journal of Computer-Supported Cooperative Work*, 12(3), p. 241-66.

[14] Pagliari, C., Donnan, P., Morrison, J., Ricketts, I., Gregor, P. and Sullivan, F. (2005). Adoption and perception of electronic clinical communications in Scotland. Informatics in Primary Care. 13 (2).

[15] Sawhney, M. & Parikh, D. (2001). Where Value Lives in a Networked world, *Harvard Business Review*, January, pp175-198.

[16] Dourish, P. and Bellotti, V. (1992). Awareness and Coordination in Shared Workspaces. In *Proc. ACM Conference on Computer Supported Work, pp107-114.*

[17] Nonaka, I. and Nishiguchi, T. (eds.) (2001). Knowledge Emergence: Social technical and Evolutionary Dimensions of Knowledge Creation, Oxford University Press, Oxford.

[18] De Roure, D., Jennings, N.R. and Shadbolt, N.R. (2001). Research Agenda for the Semantic Grid: a future e-Science Infrastructure. Technical Report for the e-Science Core Programme.

[19] Bechhofer, S.K., Rector, A.L. and Goble, C.A. (2003). Building Ontologies in DAML+*OIL*. *Comparative and Functional Genomics*, Volume 4, pp 133–141. John Wiley ISSN 15316912.

[20] Sporns, O., Tononi, G. and Kötter, R. (2005). The Human Connectome: A Structural Description of the Human Brain. PLoS Comput Biol 1(4): e42.

[21] Rosse, C., Kumar, A., Mejino Jr., J. L.V. Cook, D.L., Detwiler L.T. and Smith, B. (2005). A Strategy for Improving and Integrating Biomedical Ontologies in: *Proceedings of AMIA Symposium 2005*, Washington D.C., 639-643

[22] Breton, V., Dean, K. and Solmonides, T. (2005). The HealthGrid White paper, In Solomonides, T., McClatchy R., Breton V., Legre, Y. & Norager, S. (eds.) *From Grid to HealthGrid*. IOS Press ISSN 0926-9630.

[23] Buscher, M., Shapiro, D., Hartswood, M., Procter, R., Slack, R., Voss, A. and Mogensen, P. (2002). Promises, Premises and Risks: Sharing Responsibilities, Working Up Trust and Sustaining Commitment in Participatory Design Projects. In *Proceedings of the Participatory Design Conference*, Malmo, June, pp 183-92.

[24] Hartswood, M., Procter, R., Rouchy, P., Rouncefield, M, Slack, R. and Voss, A. (2002). Co-realisation: Towards a Principled Synthesis of Ethnomethodology and Participatory Design. In Berg, M., Henriksen, D., Pors J. and Winthereik, B. (eds.), special issue on Challenging Practice: Reflections on the Appropriateness of Fieldwork as Research Method in Information Systems Research, Scandinavian Journal of Information Systems, 14(2), p. 9-30.

[25] Bijker, W.E., Hughes, T.P. and Pinch, T.F. (1989). The social construction of technological systems: New Directions. In Bijker, W. and Law, J. (eds) *Shaping Technology, Building Society: Studies in Sociotechnical Change.* MIT Press, Cambridge, Mass.

[26] Von Krogh, F.G., Nonaka, I. and Nishiguchi, T. (2000). *Knowledge Creation*, Macmillan, London.

[27] Wenger, E. and Snyder, W. (2002). Communities of practice: the organizational frontier, *Harvard business review*, Jan/Feb, 139-145