The design of usable access to a large open microscopy environment

Case for support part 1: Previous track record

The team which will undertake this action research project offers an excellent combination of expertise and infrastructure to carry out the work to an international standard. **Applied Computing** at Dundee, founded in 1980 by Alan F Newell, has a 25 year track record of successful interdisciplinary human computer interaction research. Applied Computing is a pioneer of techniques for eliciting system design requirements for complex circumstances, including the application of computer vision to support independent living; electronic prescribing safety in high stress environments in healthcare; the parameters for acceptability of technological interventions, and optimising complex interfaces to enable interaction. The group has wide experience of interaction design for novel settings, including, for example, space systems, automobile control, electronic prescribing and other healthcare applications, communication systems and cognitive aids for older and disabled people, including ICU patients, word processing for dyslexics, and agent computing. The 50 staff and 20 full time research students taken together provide a rich multidisciplinary mix in a research environment which fosters collaboration and sharing of expertise. Disciplines represented include computer science, psychology, design, mathematics, ethnography and rehabilitation engineering. Applied Computing also has strong research and teaching links with Duncan of Jordanstone College of Art and Design's Schools of Design, and Television and Imaging. Since 2001 Applied Computing has won over £6m in funding from public and private sector, and charitable sources to further its research.

Jason Swedlow's lab has been in operation since 1998 at the Wellcome Trust Biocentre in Dundee and initiated (with Peter Sorger of MIT) the development of the Open Microscopy Environment that forms the exemplar and technical base for the research proposed here. The OME is an ideal project to work on from a usability perspective as it is technically mature but has been developed so far primarily with a view to providing backend and middleware tools to support biological image informatics. This work continues and is supported by a team of developers in Swedlow's lab who are responsible for progressing the functionality. So, the OME has strong technical/development support and these staff will work closely with the proposed research. The growing number of scientists using the OME, both within Swedlow's lab and more widely, provide a ready pool of scientist/users to supply input for ethnographic study and feedback on design stages.

Peter Gregor, the principal investigator, is Head of Applied Computing and is Director of the Digital Media Access Group (www.dmag.org.uk), a leading UK accessibility and usability consultancy and research group with clients ranging from the University of Oxford to Yahoo! in the US. With a background in experimental psychology, he has researched in the computing field for over 15 years and has published widely on interaction related issues. Within the broad sphere of interactive system design, his research interests include novel interaction modalities, applications and ICT support for older people, interactive systems for healthcare (for use by staff and patients) and the use of computers to interview. He currently holds research funding from government, charities and industry in excess of £1m. He teaches Human Computer Interaction, Interactive System Design and Introductory Java Programming, and has recently initiated a groundbreaking undergraduate degree programme in Interactive Media Design, a collaboration between Applied Computing and the University of Dundee's Design School. Publications can be viewed at http://www.computing.dundee.ac.uk/research/pubsearchresults.asp?12

Jason Swedlow established his laboratory in 1998 at the Wellcome Trust Biocentre, University of Dundee, as a Principal Investigator and Wellcome Trust Career Development Fellow, becoming a lecturer in 2000. His lab focuses on studies of mitotic and interphase chromosome structure and dynamics. Almost all of this work is based on using high- resolution fluorescence imaging to study cell mechanics. His important achievements include: first demonstration of the dynamics of a chromosomal enzyme in living cells (PhD thesis, published in Cell, 1993); first quantitative comparison of different modes of microscopy (published in PNAS, 2002); first biochemical analysis of an important mitotic regulator, Aurora B (published in JBC, 2001); establishment of an open source software development project called the Open Microscopy Environment (OME) for developing tools

for the management of image data (http://openmicroscopy.org); first linkage of the activity of the Aurora B kinase to a critical regulator of microtubule dynamics, MCAK (published in Dev Cell, 2004).

Gaining a PhD in Biophysics in 1994, Jason Swedlow was a postdoctoral fellow with Dr. T. J. Mitchison at UCSF and then Harvard Medical School from 1994 and 1998, supported by a Damon Runyon Walter Winchell Cancer Research Fund Fellowship from 1995 to 1997. In 2002, he was awarded a 5 year Wellcome Trust Senior Research Fellowship. He participates as Faculty in the Analytical and Quantitative Microscopy Course at the Marine Biological Laboratory and is co-organiser of the Immunocytochemistry, in situ Hybridization and Live Cell Imaging Course at Cold Spring Harbor Laboratory.

Ian Ricketts holds a personal chair in Assistive Systems & Healthcare Computing at the University of Dundee, Scotland, UK. He received his PhD from Dundee in 1977 and subsequently worked for Reuters Ltd., in London on software development for share price reporting and then with NCR Ltd in Dundee on the software for their early Auto-Teller Machines. He established the Computer Vision Research Group in Dundee in 1985 and combines this research interest with the application of computers to deliver improvements in health care, involving the analysis and design of large, complex, safety critical systems. Professor Ricketts has been a grant holder in over 88 research projects amounting to over £4m and has published his research widely. He has also produced a project management text 'Managing your software project', now published in many languages. Professor Ricketts is a member of the IEEE, the IEE, a Fellow of the British Computer Society, a Fellow of the Royal Society for the encouragement of Arts, Manufactures & Commerce and a non-executive director of Calico Jack Ltd. Publications can be viewed at http://www.computing.dundee.ac.uk/research/pubsearchresults.asp?23

Catriona Macaulay, who will be responsible for implementation of the project's ethnographic approach, first began exploring the potential of ethnography as a tool for systems design in her PhD, which explored the concept of information seeking through a one year field study at a UK national daily newspaper. Her extensive field experience was exploited on an EC ESPRIT funded project in 1999-2001, where as research fellow she developed field studies of household technologies. This work led to a series of collaborations with Professor David Benyon of Napier University exploring the role of ethnography in Scenario Based Design and led to the development of a method, FLEX, which provides a route from field studies to software engineering within a Scenario Based Design framework. In 2002 Dr Macaulay was funded by a leading UK based international textiles company to explore blue-sky opportunities for developing wearable products for first responders to emergency situations. She has also been commissioned by a UK learning environments developer to explore from a user centred design perspective techniques for visualising very large medical taxonomies.

Case for support part 2: Proposed research

Background

A major unmet challenge in e-science is the management of the burgeoning quantities of data and analysis involved in biological microscopy. The transition of a microscope's output from an "image", recorded on paper or film, to digitally recorded "data" has created new demands for storage, analysis and visualisation that are not adequately met in any available software package. With effective methods for management and visualisation, the potential for this data to contribute to the advance of scientific knowledge quickly is boundless; without these methods, the data and analysis carried out by individual scientists is likely to be repetitive and isolated, and tracking analysis that has been carried out will become increasingly difficult to manage. Solving this image informatics problem is critical: use of microscopy for large scale "screening by imaging," in which cells or organisms are treated with libraries of small molecules, banks of small inhibitory RNAs, etc. to identify chemicals or genes that affect a particular biological process by virtue of a change in cellular behaviour or appearance is an increasingly common technique in industry and academia—data is actively being acquired, with users demanding solutions for large-scale data management.

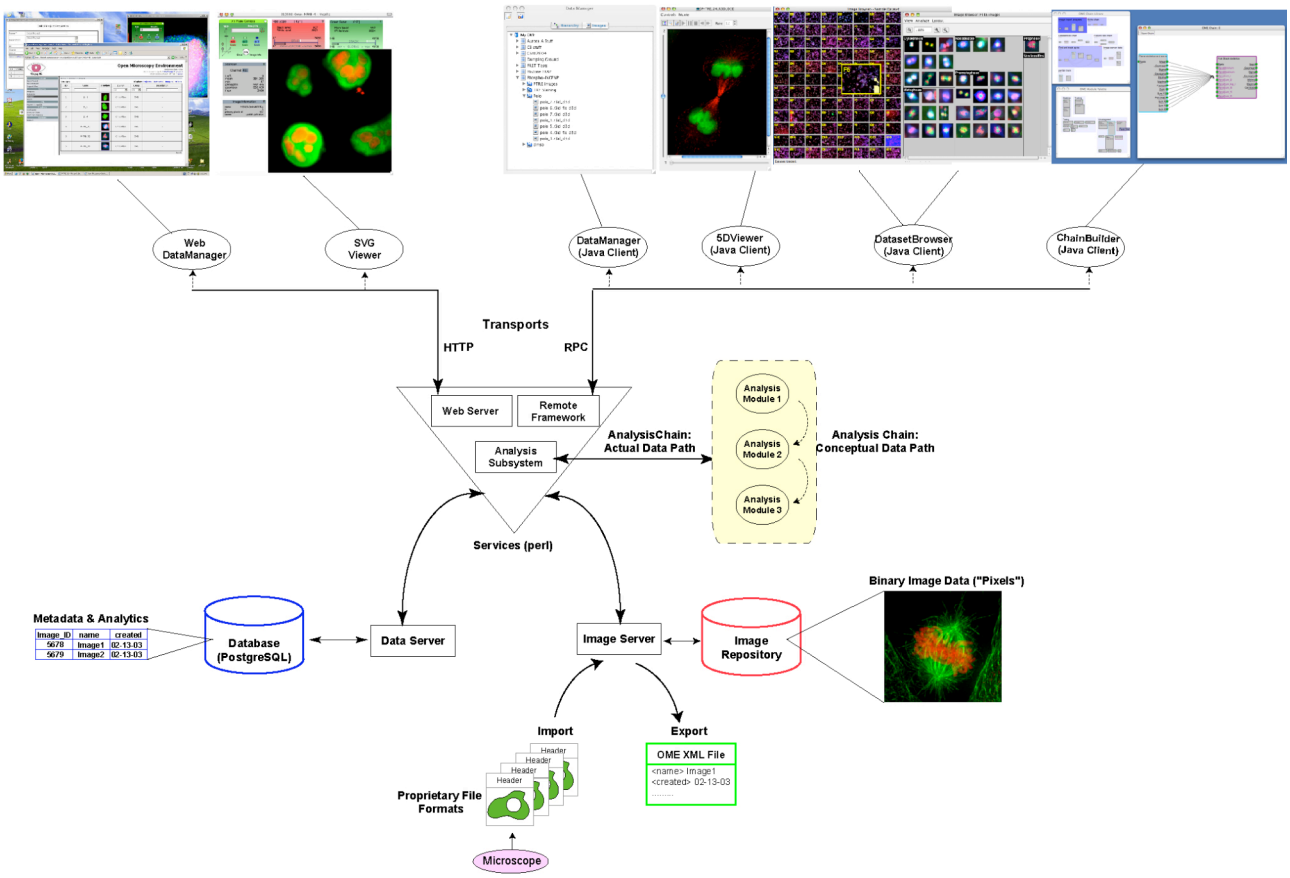

Figure 1. OME 2.4 Architecture. Diagram shows the relationships between the different OME servers, the Services layer, and the various user interfaces. For a full explanation see http://opemicroscopy.org/getting-started/ and http://openmicroscopy.org/system-overview/.

The development of the Open[1] Microscopy Environment (OME) project (Swedlow et al., 2003) is the first collaborative attempt to develop and deliver adaptable data management software tools for biological microscopy. OME provides data storage and management tools for large sets of microscope images, along with analytic results, and metadata describing the experiment, data acquisition, user, etc. (Figure 1). The OME project is an e-science solution for the storage and analysis of optical microscope image data. OME aims to automate image analysis, modelling, and mining of large sets of

images and specifies a flexible data model, a relational database, and an XML[1]-encoded file standard that is usable by potentially any software tool.

The OME project has focused on the implementation of backend and middleware tools to support biological image informatics and has only recently begun to develop user-oriented tools, which are still immature and have not benefited from more mature methods for HCI design. Difficulties are related in some cases to the design of the system at both top and lower levels, but more often they are related to a mismatch between how users think about their work and how the system offers support. At the time of writing there are two interfaces in operation, a web-based application for basic manipulations and a Java-based system, Shoola, which is essentially a framework of several semi-autonomous "agents" that provide different functionality and views into OME's rich data model (http://openmicroscopy.org/getting-started/). The OME Java Client is cross-platform and gets all of its information from a user-specified remote OME server and can be run any time the user is connected to the internet. Both access systems are freely available to the scientific community and are open source. Both are recognised as proof-of-concept and upon use, are immediately seen to suffer from a lack of HCI design application.

There is therefore a significant opportunity to investigate how scientists carry out their work and how that relates to e-science support for that process. There is a mature and in-demand e-science application in the form of OME which is technically stable and has been available since 2003; there is a well equipped pool of scientists working in the field internationally; it is becoming increasingly apparent that to derive maximum benefit from the OME's technological support possibilities for cutting edge research, the richness and complexity of the functionality of the system is generating it's own usability challenges (as noted generically by the e-Science Usability Task Force[2]). Recent evaluation studies by the OME team have identified a range of significant usability problems. An investigation of techniques for eliciting genuine requirements and for the staged introduction of usability improvements is proposed which will lead both to an improved working environment and to generic knowledge on the processes involved in the successful implementation of usable e-science applications.

Data visualisation & ethnography

Difficulties which arise in the use of large, complex systems often derive from the need for an overall, comprehensible, scheme for representing the application area visually. If a useful visualisation can be developed, then it serves to make all the detailed level usability questions more easily answerable.

Representing very large data sets visually is a complex problem and there is no universally accepted solution (Graham et al 2002). The problem is further complicated due to the need to show context and correlations between nodes (Chen and Yu, 2000), and in the case of OME to represent time based sequences of images. Approaches to representing very large data sets in 2D include tree maps (Johnson and Schneiderman, 1991), where each rectangle represents a node and subdivisions represent the node's children. Tree-maps can represent many levels but cannot display cross-references. Cone-trees (Robertson et al, 1991) are a common way of attempting to solve this problem by representing hierarchies in a 3-D visual format. This brings its own problem with occlusion of nodes and with hierarchies in the region of 40,000 nodes it is easy to imagine how potentially confusing the screen becomes. Self-organising maps have been developed which use clustering to create an analogy between visual proximity and meaning (Graham, 2001). More novel visualisations include landscape creation and 3D flyovers, and visualisations to deal with dynamic data (Bentley et al, 1992). The exploration of the problems of representing very large textual datasets has now been extended to include very large image and other media sets (Rodden et al, 2001).

A key point to note about all this work however is that it has been developed largely to respond simply to the challenge of the data set in question, not the context of use of that data set. It has been suggested that the central question for IT research should not be what kind of technology should we create, but rather what we want to do with it (Harper, 1998). As with many complex and highly skilled areas of work, the complexities of the contexts within which people will use OME make it very difficult for developers to generate appropriate requirements, and hence effective solutions. It is with

[1] XML = eXtensible Markup Language
[2] http://www.cs.notts.ac.uk/~tar/UTF.pdf

this problem in mind that the HCI and interaction design communities have adopted ethnography as a tool for generating requirements and undertaking design (Hughes et al, 1997). The existing system has been developed with little input from the HCI and interaction design traditions, or from the growing movement within these domains that recognises design systems for rich and complex work requires a rich and complex understanding of that work (Suchman, 1995).

In recognising the need to tackle context of use in a richer sense than previous 'task based approaches' had, the HCI design community also recognised the need to develop design methodologies which could incorporate the results of such a shift. Participatory design (PD) and its Scandinavian cousin co-operative design were developed to involve users more closely in systems design, and entail a shift in emphasis away from cognition/task-based approaches to design and towards understanding the behaviours of groups of users interacting in complex ways within organisational settings (Kyng, 1995). A number of techniques have been adopted and/or developed to facilitate the PD/co-operative approaches. Key are scenarios of use - narratives describing what people do/might do in the future when engaged in particular activities (Carroll, 1995; Mack, 1995). Scenarios are usually developed from in-depth ethnographic studies (Nardi, 1995) – capturing user's stories – and further developed or brainstormed by users and designers in design workshops or in brief sessions with workplace managers (Muller & Carr, 1996).

In recent years we have seen a considerable growth in the adoption of ethnographic techniques within design (Macaulay, 1998). Interest in ethnographic techniques in design can be traced to a growing concern about the gap between users and designers, and the perceived relationship between this gap and systems failures (Berg, 1998). More recently an extension of scenario based design – the FLEX method – has been made to explicitly link ethnographically-informed scenarios to software development through managing the process of abstracting from field study to use case (Benyon and Macaulay, 2004). FLEX has also been shown to aid communication in design projects by offering a lingua franca which can be developed from field study, through user participation workshops, to final implementation.

So, to respond to the growing e-Science usability challenge, we have an operative e-science domain and application; a body of knowledge and expertise on the challenges of information science and visualisation, and a proven approach to tackling complex and novel design domains. The project proposed here is an investigation of methods for improving usability – both from the requirements gathering and the implementation perspectives - of a large, functionally rich e-science application, and an implementation of the resulting findings in the OME as an exemplar.

Programme & methodology

The overall objective of this research is to contribute to the usability of e-science by investigating ways of improving access to manipulate, process and organise large and complex datasets. The Open Microscopy Environment (OME) will be used as a working exemplar. The proposed work will combine ethnography, interaction design, and technical implementation in an action research environment with a live project. Close collaboration with existing development staff will produce an improved OME environment based on full knowledge of how the scientists work and what their requirements for e-science support are, as well as producing extensive generic practice based evidence which will be applicable in other e-science domains.

Specific objectives of the research are as follows, in chronological order:

1. Using ethnographic techniques, to study the ways in which investigators perceive and interact with a complex e-science system (the OME);
2. To analyse the usability of the existing interfaces to the OME using appropriate usability and user experience evaluation tools;
3. On the basis of the findings from 1 & 2 above, to develop an iterative user experience improvement plan, adopting the best available techniques from human computer interaction studies and other fields; to apply this plan and measure the outcomes;
4. To disseminate process and outcome findings as widely as possible

This proposed work seeks to develop on the theme emerging in the e-science movement[3] that recognises that design success relies upon a rich understanding of the design problem being tackled. We will adopt a scenario-based approach to the project, undertaking an ethnographic study of the context of use of OME, and developing in cooperation with OME end-users a range of scenarios which can be abstracted to produce use cases for the development team.

Work plan

Months 1-6

- Usability review of existing OME environments, using expert review and evaluation during active use, with data logging, focus groups and suggestion boxes;
- Begin ethnographic studies of workspace, and scientific and technical staff roles; prepare ground for participatory design work (Kujala, 2003);
- Monitor process effectiveness, using interview and attitude measurement;
- Progress report for EPSRC;

Months 7-12

- Review state of the art in visualisation & interaction design (as it has progressed from proposal date);
- Scenario and use case development, based on modified FLEX approaches (see above). for identified parts of the systems which represent issues of more general feature complexity concern;
- Scenario and use case development for the highest level (i.e. metaphorical) view of the system, again using modified FLEX methodology;
- Publish initial findings in appropriate journals;
- Progress report for EPSRC

Months 13-24

- Implementation as simulations, or, where appropriate, developments of existing functionality for comparative evaluation;
- Iterative cycle of the above, leading into analysis of the process and development of the methods;
- Publish findings in appropriate journals;
- Participatory methodology dissemination workshop targeted at e-Science community. This workshop will refine and validate aspects of the methodologies;
- Progress reports as required;

Months 25-36

- Implementation and evaluation of refined version of OME;
- Publish findings in appropriate journals;
- Methodology dissemination workshop targeted at e-Science and wider design communities;
- Final report to EPSRC;
- International dissemination via appropriate journals and conferences
- The work plan for action research outlined above offers numerous benefits; it tackles the problem of designing for highly complex work contexts, it maximises the opportunity for uncovering insights into the context of biological imaging work where none currently exists beyond highly abstracted task descriptions, and it offers an opportunity to develop a method which has already shown promise.

Project management

Active development on OME itself is currently underway at six different sites. Significant effort has been expended in installing a series of project management tools, almost all of which are run from and sited at Dundee, to coordinate all work on the project. These include a CVS server (http://cvs.openmicroscopy.org.uk), and an internal documentation and proposal site (http://cvs.openmicroscopy.org.uk/tiki), and a series of automated testing and release tools.

[3] as expressed, for example, by the "Designing for Usability in eScience: A suite of tutorials for Researchers Interested in Improving Usableness and Usefulness of their eScience tools" workshop mounted recently by the e-Science Institute in Edinburgh

The proposers' record of published work, milestone achievement, and production of functional releases of OME (http://openmicroscopy.org/progress/) demonstrates the effectiveness of our methods. Overall management for the project will be carried out by the applicants, Peter Gregor and Jason Swedlow, with input from Ian Ricketts, who has published widely on project management (Ricketts, 1998). Ethnography input will be managed and coordinated by Catriona Macaulay. The team will work closely with the OME software development staff in post, as well as with scientists using OME.

Relevance to beneficiaries

This research will lead directly to better techniques for investigating requirements for large, complex e-Science applications; this advance will lead to applications which will facilitate cooperation and rapid results dissemination to scientific communities. The beneficiaries of this research will most directly be those involved with the Open Microscopy Environment; however, the research will also be of significant benefit to the wider scientific community wishing to embrace the advantages of e-science, by presenting the narrative concerning an investigation of the ways in which scientists interact with electronic representations of complex data sets and analyses, and also through the wider applicability of the interaction principles and practices which will be derived through this process.

The underlying approach to this work adopts an "action research" model which involves the researchers working in parallel with, and interacting with, a live international project, using pragmatic structured interventions to test improvements to the means of interaction with the system as a means of deriving research data to contribute to the next stage in the iterative investigative cycle. Details of the approach itself will be published as it progresses, with commentary on the relative success of different aspects of the interventions. It is expected that this experience will have much wider applicability in the usability and human computer interaction communities.

The researchers will collaborate closely with both the software developers and scientific staff to ensure maximum participation and benefit to these groups. The development of method itself will be of relevance to the wider design communities.

Dissemination and exploitation

The work will be disseminated to the target community of scientists via papers in appropriate high impact journals. In particular the proposers will target *Science*, to update Swedlow et al (2003) for the wider scientific community and *Design Issues*, International Journal of Human Computer Studies, and *Information Visualization* to disseminate the methodological investigation and findings. Conferences to be targeted will include *ACM CHI*, *DUX (Designing User Experience)*, *Doors of Perception*, *SIGGRAPH* and *IEEE Visualization*. As can be seen from the work plan, the project will also disseminate through 2 workshops which will be targeted at the e-science and wider design communities in years 2 and 3.

It is also important that the work should be disseminated to industry as e-science application development crosses the academic/industrial boundary. This will be focused around the production of a report targeted an industry audience, reporting on usability challenges and the experience of the project in meeting them, using OME as an example.

Justification of resources

This application has not been costed factoring in the new pay structures.

The project will employ one ethnographer for 18 months and an interaction design specialist for the full three year period. The initial stages of the project require the input of an ethnographer to establish how the scientists working with OME go about their work, including interactions with peers, computing systems and the wider scientific community. The work will lead to insights which translate this analysis into the broader forum of design for e-science, as well as informing the development of the interaction design and helping to create the groundwork for participatory design. The interaction designer will be responsible initially for collaborating with the ethnographer to relate this work to the design on the OME system. The interaction designer will also be required to carry out a full usability evaluation of the OME system at the start, then, in collaboration with the development staff (already in post), to work on the incremental investigation and evaluation of alternative design strategies for the interface. This person will also be primarily responsible for implementing the dissemination strategy. The PhD student requested will undertake research into the relationship of ethnography to software

design, using the OME environment, and its investigation by the staff team, as source for producing an analysis of the process itself.

The University of Dundee will provide access to all the staff, equipment and resources required to effectively carry out this action research, including access to developer time to assist with implementation of new design concepts. The project will also be closely linked to the new Interaction Design Lab at Dundee, including its video production unit and accommodation for staff and postgraduate students.

Equipment required is modest, with the request for one portable workstation for each of the 3 researchers. Apple equipment is specified as this is what is primarily in use in the existing research environment.

References

Bentley, R., Rodden, T., Sawyer, P., and Sommerville, I. 1992. A Prototyping Environment for Dynamic Data Visualisation

Benyon, D. and Macaulay, C. (2004) A Scenario-based Design Method for Human-Centred Interaction Design. In. 'Scenarios and Use Cases. Stories through the System Life-Cycle', Alexander I. and Maiden, N. (eds). Wiley. pp 211-234

Berg, M. (1998). The Politics of Technology: On Bringing Social Theory into Technological Design. Science, Technology and Human Values, 32.

Carroll, J. M. (1995). The Scenario Perspective on System Development. In J. M. Carroll (Ed.), Scenario-Based Design: Envisioning Work and Technology in System Development (p. 1-17). New York: John Wiley & Sons, Inc.

Chen, C. and Y. Yu, (2000) Empirical Studies of Information Visualization : a meta-analysis. Int. Journal on Human Computer Studies (IJHCS), Special Issue on Information Visualisation Evaluation, 53(5) : p. 631-866.

Graham, M. (2001). Visualising Multiple Overlapping Classification Hierarchies (PhD). Dept. of Computing, Napier University

Graham, M., Watson, M. and Kennedy, J. (2002). Novel visualisation techniques for working with multiple, overlapping classification hierarchies. Taxon, Vol. 51, Issue 2, pp. 351-358.

Hughes, J. A., O'Brien, J., Rodden, T., Rouncefield, M., and Blythin, S. 1997. Designing with ethnography: a presentation framework for design. In Proceedings of the Conference on Designing interactive Systems: Processes, Practices, Methods, and Techniques. Amsterdam.

Johnson, B. and Shneiderman, B. (1991). Tree-Maps: a space-filling approach to the visualization of hierarchical information structures. In IEEE Visualization,
Proceedings of the 2nd conference on Visualization '91, San Diego CA.

Kyng, M. (1995). Creating Contexts for Design. In J. M. Carroll (Ed.), Scenario-Based Design: Envisioning Work and Technology in System Development (p. 85-107). New York: John Wiley & Sons, Inc.

Macaulay, C., Benyon, D. and Crerar, C. (2000) Ethnography, Theory and Design: From Intuition to Insight. In International Journal of Human-Computer Studies, Vol.53, pp35-60.

Mack, R. L. (1995). Scenarios as Engines of Design. In J. M. Carroll (Ed.), Scenario-Based Design: Envisioning Work and Technology in System Development (p. 361-386). New York: John Wiley & Sons, Inc.

Muller, M. J., & Carr, R. (1996). Using the CARD and PICTIVE Participatory Design Methods for Collaborative Analysis. In D. Wixon & J. Ramey (Eds.), Field Methods Casebook for Software Design (p. 17-34). New York: John Wiley and Sons, Inc.

Nardi, B. A. (1995). Some Reflections on Scenarios. In J. M. Carroll (Ed.), Scenario-Based Design: Envisioning Work and Technology in System Development (p. 387-399). New York: John Wiley & Sons, Inc.

Ricketts, I.W. 1998. Managing your software project, London; Springer

Robertson, G. G., Mackinlay, J. D., and Card, S. K. 1991. Cone Trees: animated 3D visualizations of hierarchical information. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology, New Orleans, Louisiana.

Swedlow J. R., Goldberg I, Brauner E, Sorger P.K. 2003. Image informatics and quantitative analysis in biological imaging. Science 300:100-102