The Ubiquitous Lab – Or enhancing the molecular biology research experience

Juan David Hincapié Ramos

IT University of Copenhagen
Denmark
jdhr@itu.dk
www.itu.dk/people/jdhr

Abstract. This PhD research aims to implement solutions from the Pervasive Computing field for the molecular biology laboratory. As a special requirement, the applications have to use the grid infrastructure provided by the MiniGrid framework. The proposed solutions include a grid awareness system, an intelligent laboratory bench system, and an augmented paper system. Methods from ethnography and interaction design are used to guide the solution design and evaluation processes. The expected contribution to the field is to coin the concept of grid awareness, to create a smart laboratory bench and asses its impact in the researchers' work, and to explore the usage of grid computing in electronic laboratory books and augmented paper interfaces.

Key words: User interfaces, Distributed Artificial Intelligence, Life and Medical Sciences, Pervasive computing, Location-dependent and sensitive.

1 Introduction

Biology researchers have a great deal of tools and technologies available for carrying out their experiments. They work with DNA, physical equipment, interactive simulation environments and online tools. However, there is a great gap between their work in the lab and outside the lab: in the lab work is still very physical while work outside the lab occur almost 100% in the digital world.

In the lab they work with pipettes and tubes, with vortex vibrators and microwaves. They record data in physical notebooks and annotate the details of a running experiment in post it notes. Their measuring instruments print out the results. And all the data conveys to the laboratory book. The activities outside the lab, at the office space of each researcher, include analyzing data, simulating genetic activity, collaborating with colleagues, writing documents, building spreadsheets, and preparing slides.

This PhD research aims to deploy a grid infrastructure into the laboratory, and build Pervasive Computing systems that make use of the grid computing power. The Pervasive Computing field offers us many technologies for collaborating (e.g.: share experiment results with other researchers), capturing and accessing data (e.g.:

automatic capture, storage and identification of experiments results) and using digital/physical interfaces (e.g.: digital laboratory books), just to mention a few of them.

The rest of this document discusses the research in detail. First, the document presents the problem domain, the specific problems to target, and the proposed solutions. Second, the document presents related work. Third, the document presents the approach to execute this PhD research. Fourth, the document presents the current state of the research, with special focus on the Grid Awareness MDE project. Finally, the document presents the contributions to the Pervasive Computing field.

2 Problem Domain and Specific Problems

This section presents the domain for this PhD research. The domain includes concepts from Pervasive Computing and Grid Computing. The section also presents the specific problems that the PhD research explores and the projects to do it.

The first conceptual framework for this research is Pervasive Computing. Pervasive Computing is a way of thinking about computer systems and how we interact with them, where the human environment is taken into account, and the computers vanish in the background. Their disappearance is not physical but psychological: the technologies will be available so that we will be able to use them without thinking, focusing in the task to accomplish rather than in the interaction.

To achieve this level of integration and transparency, Pervasive Computing makes use of ethnographic tools and qualitative research for understanding the users' environment and reality. They give account of the reality in a descriptive way, while keeping a high level of objectivity. This description is the input for identifying the opportunities where Pervasive Computing systems can have great impact.

The technologies used within the Pervasive Computing research are related with localization, activity recognition, multi-display environments, context-awareness, capture and access, digital/physical interfaces and many others.

The second conceptual framework for this work is Grid Computing. According to [1], Grid Computing *was originally designed as a large network of computer systems able to offer an environment where computing and storage resources are shared on-demand.* Grid Computing researchers have long questioned the existence of a distinct "grid problem", because of the multiple applications of the grid, which range from resource sharing to artificial intelligence. However, this work sticks to the original idea of resource and computing power sharing on demand.

A grid platform defines a set of protocols and an API. First, the platform defines a protocol addressing the problem of interconnecting the computers that make the grid. Second, the platform defines [2] *the basic mechanisms by which… users and resources negotiate, establish, manage, and exploit sharing relationships.* Finally, the API provides the programming interface to create grid applications.

Pervasive Computing researchers have made used of grid computing platforms for many reasons. First, small, network enabled, and processing limited devices are common components of Pervasive Computing systems; therefore there is an opportunity for using the advanced computational power of a Grid Computing

platform. Second, as those devices usually have different execution environments, developing software for the grid lifts the burden of developing for each environment. Finally, integrating a Pervasive Computing system to the grid is easier than making use of mobile agents, nomadic software, or similar ones.

Research within the Grid Computing field has explored P2P architecture for what researchers have called MiniGrid platforms [4]. Compared to the traditional grid architecture, the advantage of the P2P architecture is that it scales well with the number of users, has not a single point of failure, creates local networks, and uses ad hoc configuration. However, P2P systems are typically used to share resources between users – for example media files – not to distribute computational tasks. A MiniGrid merges the benefits from Grid Computing and P2P architectures, creating local, collaborative grids for P2P distribution of computational tasks.

This PhD research aims to deploy the MiniGrid infrastructure into the laboratory, and build Pervasive Computing systems that make use of the distributed computing power. For Pervasive Computing the MiniGrid infrastructure offers the same benefits of a grid (connectivity, cross platform, easy of development), plus the advantage of using only local networks and ad hoc configuration.

To build Pervasive Computing systems that use a MiniGrid infrastructure posses several problems. The following are the three problems that this PhD research will address:

- MiniGrid Deployment: Heterogeneous computing devices coexist in the laboratory. These devices can be standard computers and laptops but also PDAs, embedded systems, and other appliances; some of them able to make use of the MiniGrid. This situation brings up the problem of getting the MiniGrid to run across heterogeneous devices.
- MiniGrid Integration: Integration to the internet is a common trend for all technologies. However, the MiniGrid is designed to work in a local network environment. How can the MiniGrid integrate to the Internet? How does this integration affect the services offered by the grid? How is the application development affected?
- Grid Awareness: The existence of a grid does not guarantee the users will use it. However, the processes in the grid have a social counterpart as they are based in devices and the devices have owners. The information of what is going on in the grid has the potential to yield interesting social phenomena that can have implications for the grid usage.

In order to propose solutions to these problems, the PhD research includes the following projects to explore solutions to each problem, by developing Pervasive Computing technologies:

- Interactive Workbench: This project provides a multi-touch surface for the lab bench and some related multi-touch applications relevant for the execution of experiments. This project covers the MiniGrid deployment problems.
- ipLabBook: The Intelligent Paper Laboratory Book uses digital pen technology and previous research in the field of electronic notebooks for building a paper laboratory book with digital features, and Internet integration through the MiniGrid. This project covers the MiniGrid integration problems.

- Grid Awareness MDE: This system gathers data about the functioning of the MiniGrid and makes it publicly available to the researchers, via calm technology in both common areas and their work computers. Later, a study on the impact of this information on the lab work and the researchers' participation in the mini grid is conducted. This project covers the grid awareness problem.

3 Related Work

The first attempt to bring a Pervasive Computing system to the biology research lab was the LabScape project [35]. This project provides a smart environment for researchers to follow the experiment execution in the cell biology lab. It proposes the usage of pervasive hardware (tablet PCs in different places of the lab), a software solution for moving data and state around the different instances (after authentication with RFID tags), and formal representations of the work. The representations allow other researchers to understand the work, and software programs to process them.

The ButteflyNet project [36] is slightly different because it brings Pervasive Computing technologies out of the lab, to the field biologists, providing them with mobile capture and access. The mobile capture feature implies designing a system that can work for long periods of time without recharge, robust enough for the outdoor conditions and with graceful degradation in case of accidents. The accessing of the captured data should be provided both on site and later on for analysis. The ButterflyNet system complies with these requirements by providing a paper-and-digital-pen solution that allows writing information in the field and linking it with a set of rich information like pictures, tags, GPS location, etc. After returning from the field, the scientist uploads the data in his digital pen, camera, RFID tags and GPS recorder to a computer, and the ButteflyNet application associates all the material in a digital notebook.

Finally the SmartTea project [37] deals concretely with the integration of Pervasive Technology and Grid Computing in a chemistry research lab. The SmartTea project provides a digital replacement for the laboratory notebook. In a similar fashion than the LabScape project, SmartTea focuses on the experiment creation, storage and execution. These processes are supported on a self defined ontology "that encompasses their major phases: planning the ingredients, planning the procedural steps, and recording the experiment". Particular to this project is the publishing capabilities of the implemented solution through the grid and the scientist working with a personal tablet PC.

4 Project Approach

This PhD research develops two base components and three projects. The base components are an extensive literature review, and an ethnographic study of the lab. So far the research has carried out these two base components, so that they are extensive and deep enough to support the development of the proposed projects.

Moreover, according to the necessities of the design processes, more fieldwork will be performed and specialized literature will be covered.

5 Current Status

The research currently has a first version of the ethnographic studies and broad amount of literature has been covered. The efforts are now focused in the design and development of the Grid Awareness MDE project.

5.1 Literature Review

The literary review has two core fields: Interaction Design (ID) and Pervasive Computing (PC). Starting with some general readings about ID [5, 6, 7], the focus moved towards Ethnography [8, 9, 16] and Qualitative Research [5, 10], and later into the usage of Personas [11, 12, 13, 14, 15], Scenarios [5, 12] and Prototypes [17, 18] in the design process. After a general introduction to PC [23, 24, 25, 26, 27, 28], the literature review has focused on subjects like paper interfaces and digital pens [30, 31, 32, 33], calm technology [29], persuasive technology [19], activity recognition [20, 21, 22], and localization [34].

5.2 Ethnographic Fieldwork

A series of ethnographic fieldworks with basic ethnographic tools have taken place at the Nucleic Acid Technology (NAT) Lab of the Århus University. These fieldworks make a descriptive analysis of the work at the lab. Two fieldwork sessions have been done, and more are to come for the evaluation phase of the projects.

The first fieldwork focuses on observation with note taking, unstructured interviews and video recordings. This time I was introduced to the personnel working at the lab and got acquainted with the environment and behavioral rules for the place. I also got to know the different roles participating in the research, learnt about the research interests of the lab, and followed a researcher as he went about doing experiments. Finally I made place-based observations and open interviews to most of the researchers regardless of their roles. From all this material I wrote a first ethnographic report, which is more descriptive than analytical.

The second fieldwork builds on top of the information originally captured, and therefore uses semi structured interviews to both clarify certain aspects from the first fieldwork, and elicit information about some issues of interest for us as technology designers. I gathered information about technology uses and opportunities like e.g. information storage and sharing, information services, portable and wearable devices, activity recognition, etc. From these materials I am writing a more analytical description of the work in the lab, as a complement to the first report.

5.3 The Grid Awareness MDE system

Another PhD research is developing the MiniGrid framework, and the first version is ready to be deployed. The MiniGrid will be used to run bioinformatics algorithms for molecular analysis and simulations; both tasks have high computational needs. This deployment gives us the opportunity to explore issues related to the Grid Awareness problem.

The CLC Bio Workbench is a bioinformatics desktop application for molecular analysis and simulations. The MiniGrid executes tasks submitted by the CLC Bio Workbench. During the fieldwork we have learnt that despite being widely accessible, researchers do not use the CLC Bio workbench. They still prefer using multiple other competing applications because they already know them and do not have to learn to use the new tool. This situation is particularly challenging since all the researchers have received training in using the CLC Bio Workbench. Moreover, the P2P nature of the MiniGrid leads us to the problem of acquiring a peer base big enough for the MiniGrid to offer a significant advantage in terms of processing power. However, peers are managed by people, and this people need to be approached in order to have them join the MiniGrid. For this purpose I am exploring an interdisciplinary solution with elements from the PC and Persuasive Computing. Our Grid Awareness MDE system intents to persuade a research group to use the MiniGrid.

The Grid Awareness MDE system is a Multi-Display Environment (MDE) that builds on concepts from calm technology to show the level of activity of the grid at a given time. The system has a desktop screen, a public touch screen and a web interface. The target organization is the molecular biology research group, which is made up of 30 researchers. The aim is to have the researchers use the MiniGrid due to the persuasive mechanisms of the MDE system.

The system will tap into the P2P network of the MiniGrid and continuously gather data about the grid usage. This data is later converted into graphics that reflect the level of activity in the grid. The MDE has to be designed and deployed in such a way that the researchers are motivated to join the MiniGrid so their participation is reflected in the MDE.

The Grid Awareness MDE system can be classified as Persuasive Technology given that it conforms to Fogg's perspectives [19]: Intentionality perspective, functional triad perspective, and level of analysis perspective.

5.4 Grid Awareness MDE Prototype

Following the ID design process, I have created Personas that will represent the users of our system. Each persona is constructed as a "rounded character" based mainly on data from the ethnographic fieldwork, with some imaginary elements defining traits at the physical and psychological levels, which are difficult to elicit during fieldwork. Using personas in this case is convenient as it allows us to trace back our design decisions to the concrete users, and create potential users that might not have been present at the moment of the fieldwork.

Thereafter, situations and scenarios where created for each persona. Once the scenarios are established, the process consists in extracting the features that the

system will have and from these the requirements. From one of the scenarios created we can extract a feature for the system like "*the user can see detailed information about the state of the grid*". And this feature can be mapped to a requirement of the kind "*The system will provide an information screen where all the tasks are shown together with the project they belong to*".

From the requirements a series of UI sketches have been built, and will be presented and enriched at a future user centric design workshop. The following images show designs for an abstract screen and a data screen:

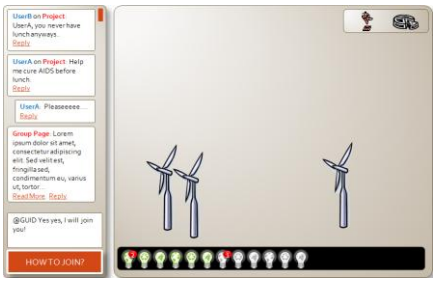

Fig. 1. Abstract screen for the Grid Awareness MDE.

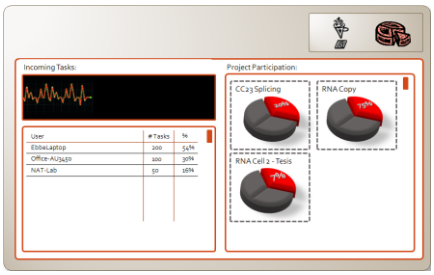

Fig. 2. Data screen for the Grid Awareness MDE.

From these elements, software artifacts can be designed and developed.

5.5 Grid Awareness MDE Evaluation

The evaluation for this project starts with a control period, in which the public and personal displays are deployed, and the researchers see the drawings both in the corridors and the web page. They will have no information as what the drawings mean or why the devices were deployed.

After a week or two an initial group of 2 researchers will start using the grid services. These two researchers have all the knowledge about the research as they took part in the Interaction Design process. As they use the MiniGrid the activity is reflected in the MDE system. Along with the new activity in the MDE a leaflet explaining the reason for changes in the screen is made available, and instructions for how to join the MiniGrid.

The response of the researchers towards the system will be tracked in terms of new MiniGrid users, information access, and usage of the CLC Bio Workbench. This data will help us resolve whether the Grid Awareness System accomplish its objective as a Persuasive Computing system.

6 Contribution to the Field

This PhD research contributes to the field of Pervasive Computing in different ways. First, it introduces the concept of Grid Awareness. Based on that, it uses calm

technology to explore the impact on users of knowing the state of distributed software. It also explores the potential of awareness information as persuasive technology for the adoption of a software product.

Second, the research explores the usage of distributed computing power in Pervasive Computing systems.

Third, it prototypes a smart laboratory bench which makes use of multiple grid enabled devices to support the work of a molecular biology researcher. For this project different pervasive computing technologies are used like context awareness and activity recognition.

Last, this research works with the paper augmented interfaces for the laboratory notebook, integrating it to the Internet by using the MiniGrid.

7 References

1. Call For Papers For “Fourth Workshop on Emerging Technologies for Next generation GRID (ETNGRID-2008)”. http://etngrid.diit.unict.it/2008/ETNGRID2008_cfp.txt.
2. I. Foster, "The anatomy of the grid: Enabling scalable virtual organizations," 2001, pp. 6-7.
3. Schraefel, G. V. Hughes, H. R. Mills, G. Smith, D. De Roure, and J. Frey, “Less is More: Lightweight Ontologies and User Interfaces for Smart Labs,” in: The UK e-Science All Hands Meeting 2004, 31st August - 3rd September, Nottingham, UK. 2004.
4. Venkataraman, N. Bardram, J. The Mini-Grid Framework: Application Programming Support for Creating and Participating in Ad-hoc, Peer-to-Peer Grids in a Volatile Execution Environment. IT University of Copenhagen.
5. H. Sharp, Y. Rogers, and J. Preece, Interaction Design: Beyond Human Computer Interaction. Wiley, March 2007.
6. Löwgren, J. (2002). How far beyond human-computer interaction is interaction design? www.boxesandarrows.com, published April 2002. Also in Digital Creativity 13(3):186–189.
7. Löwgren, J., Stolterman, E. (2004). Thoughtful interaction design: A design perspective on information technology. Cambridge, Mass.: MIT Press.
8. J. Blomberg, M. Burrell, and G. Guest, "An ethnographic approach to design," pp. 964-986, 2003.
9. J. Hughes, V. King, T. Rodden, and H. Andersen, "The role of ethnography in interactive systems design," interactions, vol. 2, no. 2, pp. 56-65, April 1995.
10. The SAGE Handbook of Qualitative Research. SAGE Publications, April 2005.
11. L. Nielsen, "From user to character: an investigation into user-descriptions in scenarios," in DIS '02: Proceedings of the conference on Designing interactive systems. New York, NY, USA: ACM Press, 2002, pp. 99-104.
12. L. Nielsen, "Engaging personas and narrative scenarios," Ph.D. dissertation, Department of Informatics, Copenhagen Business School, October 2004.

[THE OTHER REFERENCES WERE REMOVED FOR SUBMISSION]