NeuroHub: A Research Information Environment for Neuroscientists

Anne Trefethen[1], David De Roure[1], David Newman[2], David Wallom[1], Douglas Saddy[3], Garry Smith[3], Jeremy Frey[2], Maria Susana Avila-Garcia[1], Nigel Emptage[1], Philip Newland[2], Rahim Lakhoo[1],Vincent O'Connor[2],
[1]University of Oxford, [2]University of Southampton, [3]University of Reading

***Abstract*— The NeuroHub project aims to develop a research information system for neuroscientists at three different partner institutions: Oxford, Reading and Southampton. Each research group has different working practices, research methodologies and user requirements, which have lead to the development of a system that supports a wide variety of tasks in the neuroscience research life cycle. In this paper, we present how these user requirements have been translated in a research information environment that supports a community of over 70 researchers using the system for day-to-day research tasks.**

Keywords--Information environment; user requirements; neuroscience, agile development, drupal, alfresco .

I. INTRODUCTION

The NeuroHub project is developing an information environment for neuroscientists, aiming to provide an environment that allows effective management of research data through the experimental process, storing, analysing, sharing, searching and publishing; and to allow booking laboratory equipment, communicating with colleagues, and generally supporting the whole research process.

To achieve this wide-ranging vision requires building on existing components where possible, and building a system that is extensible to support easy integration of new or different tools that meet specific requirements. In order to capture those requirements there has been continuous user requirements gathering to identify which areas of the vision for NeuroHub, are of key importance to the neuroscientists and subsequently take priority within the project.

This paper describes the work done in collecting user requirements and the development methodology used; briefly introduces the neuroscience areas that define the project; outlines the development work we have carried out to date including the challenges we faced; introduces the service as it is currently used by the exemplar groups; describes the feedback we have received to and the future direction of this work.

II. NEUROSCIENCE BACKGROUND AND WORKING PRACTICES

Neuroscience requires a broad set of research processes and tools in terms of experimental approaches, subject areas and data outputs, which in turn leads to a challenge when trying to develop an information environment to support neuroscience research. The aim is to develop an infrastructure that will serve bioscience more broadly. The neuroscience groups within the NeuroHub consortium reflect the diversity of the wider field of the discipline. The research areas of the groups involved cover behavioural, psychological, clinical, systems, cellular and molecular neuroscience.

Oxford

The Oxford neuroscientists involved in NeuroHub are focussed on the study of the molecular basis of synapse formation, plasticity and the regulation of neuronal morphology in the normal and diseased brain. They are examining the mechanisms of an activity-dependent form of neural plasticity known as long-term potentiation (LTP) [1]. Techniques used include: Electrophysiology, both patch clamp and voltage clamp and Microscopy - including florescent, confocal and Total Internal-Reflection Florescence (TIRF).

The system needs to be designed to interface with any system developed to manage the data throughout the experimental lifecycle, while interfacing with their tools. The system needs to be designed for use by researchers ranging from undergraduate project students to group leaders and should allow users to easily share activities of researchers with their collaborators. It also needs to provide effective support for their research tools to annotate images and videos as they are being recorded, while making them available to other group members and collaborators. A search mechanism for images and videos, along with any associated metadata, from current and past projects is extremely desirable.

Southampton

A focus in Southampton is on the integrative analysis of brain function/dysfunction. The group utilizes different invertebrate model systems to understand particular features of the organization of neuronal networks and how different types of constituent interneurones contribute to the processing of sensory signals [2]. The insect used depends on the scientific question being addressed, but include locusts, the fruit fly, cockroaches and ants. Modelling is completed across levels of biological organisation

The NeuroHub project is sponsored by the UK Joint Information Systems Committee (JISC).

ranging from molecules, cells, tissue and systems to animal behaviour. Techniques used include: Electrophysiology, including patch clamp, voltage clamp, multi-electrode arrays and Optical Microscopy.

Like the Oxford group, Southampton requires shared access to images and videos and to the software developed for analysis. The team in Southampton have international collaborators who require access to the data and other information provided through NeuroHub. This introduces issues of authentication outside the University network and a need to introduce a level of security in the environment that might not otherwise be required.

Centre for Integrative Neuroscience and Neurodynamics (CINN) - Reading

In Reading the research is on the physiological and psychological mechanisms underpinning complex cognitive behaviours, targeting typical and atypical development and decline in individuals [3].

CINN provides dedicated facilities that house their MRI and EEG equipment, as well as interview rooms, meeting rooms and analysis laboratories. Researchers conducting experiments work with varying numbers of subjects (currently an average number for an EEG experiment appears to be 26). Researchers are responsible for hypothesis and experiment creation, arranging subject appointments, booking experiment facilities, looking after the subject while on site, preparing and executing the experiment, and recording results. Subjects are directed to the shower facilities, if required (e.g. to remove EEG brain cap gel from their hair), de-briefed, and escorted out of the centre. After an experiment, the data is post-processed, analysed and tested against a hypothesis. The results are then either published in scientific journals, or further experiments are conducted with a potentially revised hypothesis or with revised experiment stimuli. In all cases, subject data is treated confidentially and only released to other researchers (e.g. for analysis) after the data has been anonymised.

III. Requirements Gathering and User Stories

Our approach to requirements gathering has been to use a combination of embedding developers within neuroscience laboratories observing neuroscientists at work and using semi-formal interviews with the researchers and their collaborators. User needs were initially documented as notes recorded on the project wiki as well as internal reports. These were subsequently migrated into our Agile user stories page[1]. Over 100 user stories have been defined. One of the challenges for the NeuroHub project has been that the interdisciplinary teams at the three sites have different work practices, requirements and expectations. Considering this diversity of the requirements the NeuroHub team have identified the following:

- Support data collection from MRI scanners, experiment related data through the use of logbooks, input devices such as the digital pen and tablets
- Support data sharing and management by providing the functionality to store and share data with internal and external collaborators in a secure way.
- Provide secure remote data access as users require working from home or while commuting i.e. train using mobile devices as well as external collaborators.
- Support data analysis by providing mechanisms for using of scientific workflows and cloud services.
- Provide metadata management as users need to describe experimental data in a meaningful way and the mechanisms to access and search files using such metadata.
- Further requirements that were identified at the group level were the area of collaborative document preparation; and simple lab supplies ordering systems.
- There was also a requirement for community-level "social networking" with the aim to facilitate communication across the broader community of neuroscientists, sharing of ideas, data and process.

I. System design

The technology stack for the NeuroHub project uses a combination of two popular open source, and well supported technologies, namely Drupal[2] and Alfresco[3]. Figure 1, depicts NeuroHub's current architecture. Drupal provides the frontend to NeuroHub, while Alfresco provides the document store at the backend. The latter provides the facility to store and access users' files, through a WebDAV interface, in a reliable manner that scales well; whilst the former allows users to record accompanying information (metadata) about them through a Web interface. Drupal is a modular framework that allows developers to build their own bespoke content management system. It was a requirement from our users that data and information stored in NeuroHub have provenance. Therefore it was essential to utilise the versioning functionality of both systems, Drupal stores versions of the content created, while Alfresco supports file versioning.

The system provides functionality for:

- *Data Capture.* NeuroHub is trialing the use of tablet devices, such as Apple's iPad, and digital pens i.e. LiveScribe[4] digital pens with users.
- *Content, meta-data, search and management.* NeuroHub has several different content types, blogs, discussions, wiki pages, documents, log books, events etc. Metadata can be provided to facilitate discovery of appropriate files through search. To

[1] http://NeuroHub.ecs.soton.ac.uk/index.php/Agile-Development-Plans
[2] http://drupal.org/
[3] http://www.alfresco.com/
[4] http://www.livescribe.com/uk/

support the capture of this semantic file information, NeuroHub has augmented Drupal's Alfresco module. For electrophysiology data, Carmen's MINI schema will be used as a basis to define a content type for these files [4].

- *Collaboration Tools and Security.* Research groups need to be able to collaborate using external applications. NeuroHub has undertaken further development of the Google Calendar API module[5] (GCal), to allow groups to share their calendars on Google Calendar. NeuroHub also uses Drupal's Google Docs module[6] to allow groups to have collaborative access to documents shared in Google, which can then be exported as PDF or Word document from within NeuroHub.
- *Integrating with Publications Experimentation*. Management of these publications is just as important as the files and content the neuroscientist generates during experimentation. Storing these publications in a network shared Alfresco file-system publication repository, allows applications such as Mendeley[7] and Papers[8] to index and organize publications per user.

The system design presented in this section is a sustainable and extensible design that can be augmented to meet the needs of other scientific areas outside of neuroscience.

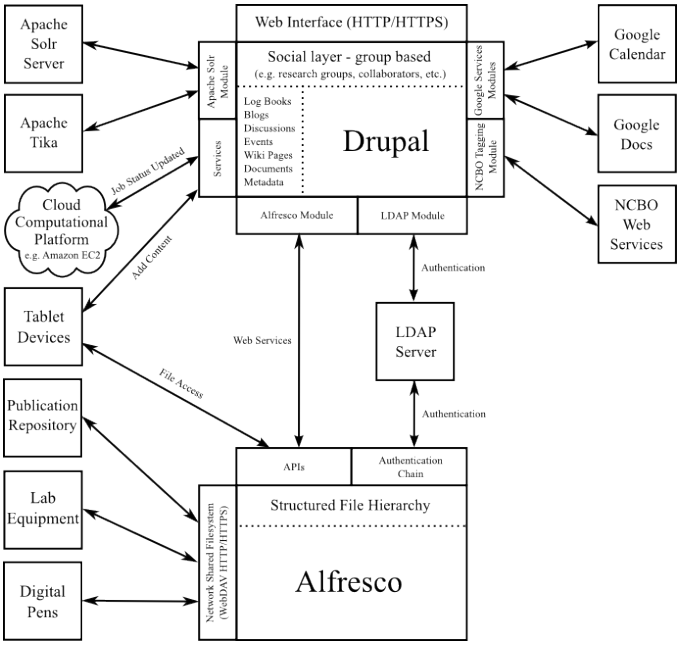

Figure 1, NeuroHub Architectural Overview

II. CONCLUSIONS

The NeuroHub environment has succeeded in building on the community efforts of Drupal developers and by strong engagement with the user community, through embedding of developers and an agile development methodology – it has succeeded at capturing the research processes of the neuroscientists and supporting them. The framework used for development has allowed a ready to install environment that at the same time is adaptable and easily grown to add further features as required – indeed to allow personalisation to suit the needs of a given laboratory. The system as developed is already in production use at CINN, Oxford and Southampton with a community of 74 researchers at the time of writing and is allowing the Southampton scientists to collaborate more effectively with groups in Brazil and monitor their researchers at home, whilst in Japan, while CINN is also using NeuroHub to communicate with collaborators in Australia.

ACKNOWLEDGMENTS

NeuroHub is an e-Research South Consortium project funded by the UK Joint Information Systems Committee (JISC) under the strand A2 for "Developing e-Infrastructure to support research disciplines". The work reported in this paper was funded in-part by the EPSRC Neurocloud project under grant EP/I016856/1.

REFERENCES

[1] Ward, B, McGuinness, L., Akermann, CJ, Fine, A. Bliss, TVP and Emptage NJ., State-Dependent Mechanisms of LTP Expression Revealed by Optical Quantal Analysis, Neuron 2006.

[2] Wilson, E, Rustighi, E, Newland, P.L. and Mace, B.R. (2011) A comparison of models of the isometric force of locust skeletal muscle in response to pulse train inputs. Biomechanics and Modeling in Mechanobiology, 11, (3-4), 519-532. (doi:10.1007/s10237-011-0330-2).

[3] Jessop M., Weeks M. and J. Austin, "CARMEN: a practical approach to metadata management", Phil. Trans of the Royal Society, no. 368, pp 4147-4159, 2010.

[4] Gibson F., Overton P.G., Smulders T.V., Schultz S.R., Eglen S.J., Ingram C.D., Panzeri S., Bream P., Sernagor E., Cunningham M., Adams C., Echtermeyer C., Simonotto J., Kaiser M., Swan D.C., Fletcher M. and P. Lord, "Minimum Information about a Neuroscience Investigation (MINI) Electrophysiology ", Nature Precedings, [online] [http://precedings.nature.com/documents/1720/version/2]

[5] http://drupal.org/project/gcal
[6] http://drupal.org/project/google_docs
[7] http://www.mendeley.com/
[8] http://www.mekentosj.com/papers/