Chapter 2.3
The PsyGrid Experience: Using Web Services in the Study of Schizophrenia

John Ainsworth
The University of Manchester, UK

Robert Harper
The University of Manchester, UK

ABSTRACT

The key aim of the PsyGrid project was the creation of an information system to ascertain and characterise a large, representative cohort of schizophrenics, beginning from their first episode of psychosis. The cohort was to be drawn from eight geographically dispersed regions of England, covering in total one-sixth of the entire population. In order to meet the current and future requirements we needed to build a secure distributed system, which not only could support remote data collection, but could also be integrated with other data sets, applications, and workflows for statistical analysis. We concluded that a service-oriented architecture was required and that the implementation technology should be Web services. In this article we present the design, deployment and operation of the PsyGrid data collection system as a case study in applying Web services to health informatics. The major problems we faced were related to the deployment of Web services into an existing network infrastructure, but overall found Web services to be the most suitable middleware technology.

INTRODUCTION

In the 1970’s Archie Cochrane (Cochrane, 1972) and colleagues alerted the medical profession to the need to weed out subjectivity and anecdote from clinical practice. At the same time there was a move to improve the safety of medicines. Since then the evidence-based care movement has grown and is now accepted by most healthcare professionals to be best practice. However, there are serious problems with the evidence on which we base healthcare: it is expensive to produce; it takes a long time to produce; it takes a long
time to influence clinical practice; it is based on clinicians’ more often than patients’ perceptions of important outcomes, which might not match; and it is crude—relating to the average participant and simple treatment definitions under ideal/trial conditions, often long ago—in other words, it gives a low-resolution picture of how your patient might respond to treatment. Most of the health informatics literature on electronic health records and “evidence into practice” is about weaving the existing evidence-base into healthcare decision-making. The role of clinical information systems in improving the evidence-base, however, has been neglected. This situation is changing: PsyGrid, which is funded under the UK Medical Research Council’s E-Science programme, is focused on providing informatics employing e-Science principles for improving clinical trials and longitudinal studies in mental health research (Ainsworth et al., 2006).

PsyGrid employs service-oriented architecture computing techniques and technologies in the implementation of a system that will remove the barriers to epidemiological research. The process of epidemiological research has three phases—the establishment and characterisation of a large, representative cohort from a geographically distributed population; the integration of the cohort data with other data sources to provide additional characterisation; and the formulation of a hypothesis and generation of the corresponding predictions. For the establishment and characterisation of a cohort, many epidemiological studies use paper based data collection systems. A computer-based data collection system which enables geographically distributed data collection would alleviate much of the labour, tedium, and error that are inherent in paper-based data collection. Such a system is required to store personal, confidential medical data, and this data must be sent across a network from remote data entry clients to a data repository server. The immediate goals in developing the system were to ensure high quality data; ensure privacy and confidentiality of sensitive patient data; enable data collection to be performed at any location; ensure that the system was scalable to cover larger populations and that it was highly available. In the long term the system needs to be enhanced and extended to provide a platform for the epidemiological study of schizophrenia and intervention research for predicting and preventing adverse outcomes.

We selected Web services as the underlying middleware because they offer firewall traversal, interoperability, modularity, and loose coupling which provides the ability to incrementally extend the capabilities of the system as it grows.

The health informatics findings from PsyGrid will be generalisable well beyond its disease focus. By treating every new patient as a participant in a longitudinal study, it will start to test a new model of care and research combined. We believe that this combination is essential to providing a timely and more flexible evidence base for future healthcare. This future could be called “high resolution healthcare”; it would encompass “personalised medicine”, self-care decision-support, efficient and opportunistic clinical trials, complex (including genomic) epidemiology, and tactical development of local services based on local environmental factors and outcomes at the population level. High-resolution care and research requires information systems to link relevant data, methods and people in a clear and timely fashion.

In the following section we examine the clinical need for PsyGrid, and in then we analyse the requirements that flow from this. Afterwards, we provide a description of the NHS Information Technology infrastructure into which PsyGrid will be deployed. The rationale for using Web services is then presented, followed by an overview of the PsyGrid system architecture and the interactions between the Web services. We report our experience of implementing Web services, and provide a detailed description of the systems functionality and implementation. Thereafter, we