Classifying Data Sharing Models for e-Health Collaborations

Prof. Richard O. Sinnott
University of Melbourne
rsinnott@unimelb.edu.au

INTRODUCTION
Seamless access to clinical and biomedical data sets is the cornerstone upon which the vision of e-Health depends. A multitude of projects and initiatives developing e-Health infrastructures providing access to a range of clinical and biomedical data sets have occurred [1-3], however by and large no clear consensus on the best way to build e-Health infrastructures has been established. Rather, different projects and initiatives have typically developed their own software solutions for their own particular needs and recycling of existing systems has been the exception as opposed to the rule. This is not surprising in many respects given the heterogeneity of many existing clinical systems and the rapid evolution taking place across the post-genomic (genomic, proteomics, metabolomics etc) space and the numerous advances in imaging and diagnostic techniques. However it is clear that the future success of the e-Health vision and its translation to personalised medicine, improved healthcare and the many other opportunities identified in the post-genomic age depends upon lessons learnt in developing e-Infrastructures, and ultimately being able to classify and compare solutions.

Ideally a common architectural framework would exist by which reference implementations could be compared – much like the OSI protocol stack. However no such overarching framework exists and it would appear that at least for the foreseeable future, e-Health infrastructures are likely to remain largely ad hoc and uncoordinated across different communities in different countries. In this context, establishing best practice and comparing different solutions is non-trivial, as they are typically developed with different scenarios and different communities involved.

AIMS
The aim of this presentation is to structure the discussion of e-Health infrastructures through fundamental architectural data sharing patterns that are at the heart of many kinds of e-Health collaboration. Thus whilst no single common architecture for e-Health infrastructures exists, it is the case that common patterns of data sharing exist to support e-Health collaborations – at least at an architectural/conceptual level, as opposed to lower level implementation patterns as found in the work of Gamma [4]. We identify such patterns and outline their advantages and disadvantages. In describing these patterns we do not focus in detail on the technologies that are used to implement them per se, but rather our focus is on the fundamental nature of the data sharing and collaboration models they support. Each pattern is illustrated with an exemplar project along with the advantages and disadvantages of the pattern itself. It is intended that this classification will help better shape the future e-Health infrastructure discussions and provide the basis for comparison of solutions as well as shed insight to others in developing their own e-Health solutions.

APPROACH
It is especially challenging to develop a single solution that meets the needs of both the clinical and the associated research community where crossing of organisational and administrative boundaries is required. However, by focusing on the architectural patterns of data sharing and collaboration, it is possible to better shape discussions on e-Health infrastructures more generally. More generally, it can be stated that architecture-driven design is not a new phenomenon. Numerous examples of how abstract architectures can be used as the basis for system development and comparison have been undertaken in a variety of application domains [5]. Similarly, work on design patterns as the basis for systems development and comparison has been widely adopted by developers and researchers in other domains [6].

The initial sets of data sharing patterns we identify and the associated e-Health projects we used to outline them in more detail are:

- Dependent, open-distributed, data sharing models;
- Independent centralised data sharing models;
- Dependent push-based (out-going) data sharing models;
- Dependent security-oriented pull based data sharing models;
- Dependent security-oriented push-based (incoming) data sharing models.

Here we classify two keys forms of data: dependent and independent. In the dependent case, we mean that the data is actual data from an existing data source, e.g. a patient record from a hospital clinical data management system or genomic data from a bio-repository. Dependent data may well be manipulated in some form, e.g. anonymised by a software system, but the key point is that it is based on pre-existing e-Health data. Independent data on the other hand, is data that does not come from a pre-existing digital resource directly or through some programmatic translation of an existing e-Health digital resource. Rather it is typically data created by clinicians and/or researchers for the purpose of a particular collaboration. This is a subtle but crucial point to e-Health data sharing infrastructures. We also note that it is quite possible to have hybrid combinations of these data sharing models.

Melbourne, Australia
5th eResearch Australasia Conference
6-10 Nov 2011
DISCUSSION

The presentation will discuss the pros and cons of all of these approaches with practical demonstrations in their utility. The projects include work in genetic causes of hypertension [7]; brain trauma research [8]; mental health and suicide [9,10], and cancer research platforms incorporating biobanks and major clinical trials [11,12]. It is noted that these patterns apply to many other research domains and indeed are currently being applied in the Australian Urban Research Infrastructure Network (AURIN – www.aurin.org.au) project.

REFERENCES

[3] Diagnostic Mammography National Database Project (eDiaMoND), www.ediamond.ox.ac.uk

ABOUT THE AUTHOR

Prof. Richard O. Sinnott is Director of eResearch at the University of Melbourne. Before this he was the Technical Director of the National e-Science Centre at the University of Glasgow; Deputy Director of the Bioinformatics Research Centre (also in Glasgow) and the Technical Director of the National Centre for e-Social Science. He has been involved in an extensive portfolio of e-Science projects in the UK, Europe and now in Australia, with specific focus on application domains where fine-grained access control is essential.