Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
A COMPUTATIONAL APPROACH TO PRIMARY HEALTHCARE INFORMATION QUALITY INDICATORS

A THESIS PRESENTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN COMPUTER SCIENCE AT MASSEY UNIVERSITY, PALMERSTON NORTH, NEW ZEALAND.

Stephen Richard Lean

2014
For my parents, Beverley and Glen. Gone much too soon, remembered everyday.
Abstract

In many countries around the world Information and Communication Technologies (ICT) are being leveraged to produce efficiency gains and cost reductions in healthcare by making health information more readily available in clinical contexts. This raises issues as to the use of health information in clinical decision making at point of care, as relying on poor quality information in this context can have serious consequences. This thesis investigates quality criteria that are used when assessing health information, with the objective of formalising those criteria for use with a prototype software system. Literature, as well as standards and currently used forms of electronic health records, were reviewed for what they offer for assessment of health information quality. A lack of criteria from these sources necessitated interviewing practicing General Practitioners (GPs) to determine criteria important to them, and how they assessed the information they want to use. Interviews were of a semi-structured type using vignettes, for clinical context. Recruitment used a Snowball methodology. Results were analysed and interpreted using Thematic Analysis and showed the GPs assessed information using criteria based on tacit knowledge, formed from community knowledge and past experience.

The Quality Criteria (QC), discovered to be integral to this process, were formalised using the developed Quality Criteria Model (QCM). A prototype system was developed to demonstrate that using a current health information standard, meta-data could be used to detect the presence of QC within health information and capture these via instantiation of the QCM. The results of successful detection of QC are then Health Information Quality Indicators (HIQI). Contributions for this thesis include the following: the set of discovered QC, thematic maps that capture the combination of criteria and the process used when applying them, the formalised model for QC (the QCM), determination that additional meta-data will be required to detect those QC categorised as being subjectively evaluated, and the demonstration that a software system can detect, and capture, QC found in health information. Implications are discussed such as that just having access to information is insufficient, and subjectively evaluated QC are problematic for implementation and use. Finally, conclusions are drawn and future work suggested such as user interface development for HIQI representation, alternative search algorithms for QC detection, and further development of the prototype toward a production system.
Acknowledgements

This thesis would not have been completed without the help of many people. I would like to take this opportunity to thank those who helped with this project.

Deep appreciation goes to my three supervisors: Prof. Hans W. Guesgen, Dr. Inga Hunter and Dr. Kudakwashe Dube. Many thanks for the guidance and expertise shown throughout. The complimentary skills that each brought to the project helped in immeasurable ways, along with the encouragement and patience. It was not only the help with the research that was appreciated, but also the support shown on a personal level.

My thanks to staff in the School of Engineering and Advanced Technology (SEAT), especially Michele Wagner, who helped in innumerable ways, and Prof. Stephen Marsland who supported the project when needed. I would also thank my fellow postgraduate students Gareth, Yuliya, Jawad and An, with who I shared an office. The help, encouragement and light relief, along with some lively discussions, provided along the way will always be remembered.

I would also like to thank Angela Baker for wise and supportive council, it really helped. Also, thanks to Janene Walter from Massey University’s Graduate Research School for assistance with the process of completing a PhD.

Thanks to Massey University for financial support in the form of the Vice-Chancellor’s Doctoral Scholarship, without which I would not have started this project. Also, to SEAT for part-time work opportunities and for the allocation of a budget for research expenses.

I am also grateful to those who participated in the interviews that formed part of this research. Without the interviews, this project would not have been possible.

Last, is thanks to family and friends who showed an interest and continued to ask how it was all going.
Contents

Abstract v

Acknowledgements vii

List of Abbreviations xix

1 Introduction 1

1.1 Background and Motivation ............................ 1

1.2 Research Aims And Goals ............................. 3

1.2.1 Research Questions ................................. 3

1.3 Research Scope ........................................ 4

1.4 Research Design ........................................ 4

1.5 Thesis Contributions ................................... 4

1.6 Thesis Plan and Organisation ........................... 5

2 Background, Motivation And Research Problem 7

2.1 Chapter Overview ....................................... 7

2.2 Health Information as a Resource ........................ 8

2.3 Health Sector Pressures ................................ 9

2.3.1 Health Sector Direction - Tension for Change ....... 11

2.4 Out-Of-Context Use of Information ........................ 12

2.5 The Problem of Providing Health Information Quality Indicators ........................ 13

2.6 Literature Review Process ................................ 13

2.6.1 Terms and Definitions ................................. 14

2.7 Literature Review ........................................ 15

2.7.1 Defining Data Quality ................................ 17

2.7.2 Why Data Quality Is Important ....................... 18
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.1</td>
<td>Rationale For “Importance In Knowing” Approach</td>
<td>79</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Quality Criteria Model - Development</td>
<td>81</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Quality Criteria - Theoretical Formalisation</td>
<td>83</td>
</tr>
<tr>
<td>4.4.4</td>
<td>Quality Criteria - Applying the Formalisation</td>
<td>85</td>
</tr>
<tr>
<td>4.4.5</td>
<td>The Quality Criteria Model - Summary</td>
<td>95</td>
</tr>
<tr>
<td>4.5</td>
<td>Chapter Summary</td>
<td>97</td>
</tr>
<tr>
<td>5</td>
<td>Prototype Development</td>
<td>99</td>
</tr>
<tr>
<td>5.1</td>
<td>Chapter Overview</td>
<td>99</td>
</tr>
<tr>
<td>5.2</td>
<td>Rationale For Prototype</td>
<td>99</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Why A Prototype?</td>
<td>100</td>
</tr>
<tr>
<td>5.3</td>
<td>Domain Environment</td>
<td>101</td>
</tr>
<tr>
<td>5.4</td>
<td>Health Information Model To Be Used - The CCR</td>
<td>101</td>
</tr>
<tr>
<td>5.5</td>
<td>Prototype - Development Environment</td>
<td>105</td>
</tr>
<tr>
<td>5.6</td>
<td>Prototype - Data Structures</td>
<td>106</td>
</tr>
<tr>
<td>5.6.1</td>
<td>Prototype - QCM Definition</td>
<td>107</td>
</tr>
<tr>
<td>5.6.2</td>
<td>Prototype - CCR Components Used For Detection Of QC</td>
<td>112</td>
</tr>
<tr>
<td>5.7</td>
<td>Prototype - Function and Operation</td>
<td>124</td>
</tr>
<tr>
<td>5.7.1</td>
<td>Prototype - Example Operation</td>
<td>124</td>
</tr>
<tr>
<td>5.8</td>
<td>QCM Application With Other Health Information Models</td>
<td>129</td>
</tr>
<tr>
<td>5.9</td>
<td>Chapter Summary</td>
<td>130</td>
</tr>
<tr>
<td>6</td>
<td>Discussion</td>
<td>131</td>
</tr>
<tr>
<td>6.1</td>
<td>Reflections On Research Questions</td>
<td>131</td>
</tr>
<tr>
<td>6.2</td>
<td>Thesis Contributions</td>
<td>138</td>
</tr>
<tr>
<td>6.3</td>
<td>Implications For Theory And Practice</td>
<td>141</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Access Alone Is Insufficient</td>
<td>141</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Form Is Important</td>
<td>142</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Subjective Criteria Are Problematic</td>
<td>143</td>
</tr>
<tr>
<td>6.3.4</td>
<td>Difficulties with Multiple Entry Criteria</td>
<td>144</td>
</tr>
<tr>
<td>6.4</td>
<td>Thesis Limitations</td>
<td>145</td>
</tr>
<tr>
<td>6.5</td>
<td>Recommendations</td>
<td>145</td>
</tr>
</tbody>
</table>
7 Conclusions And Future Work 147

7.1 Future Work ............................................. 149
  7.1.1 Generalisability of Quality Criteria .................... 150
  7.1.2 User Interface for Representation of QCM Instances to Users . 150
  7.1.3 Application Of Subjective QC ........................ 150
  7.1.4 Evaluation of Search Algorithms for Detecting Presence of QC . 151
  7.1.5 Development of Prototype Toward Production System .......... 151

7.2 General Conclusions ........................................ 152

Bibliography 153

A Ethics Approval 167
  A.1 Ethics Approval Letter .................................... 168

B Materials For Interviews 169
  B.1 Participant Information Sheet ............................ 170
  B.2 Participant Consent Form .................................. 172
  B.3 Participant Demographic Information Collection Form .......... 173
  B.4 Interview Questions List ................................. 174
  B.5 Interview Rating Exercise Form .......................... 178

C XML Schema for Quality Criteria Model: QCschema.xsd 179
  C.1 Schema for Quality Criteria Model ....................... 179
  C.2 QCM Instance Example .................................. 182
List of Figures

3.1 Candidate Quality Criteria ........................................... 34
3.2 List of 21 candidate quality criteria .................................. 43
3.3 List of 21 rated candidate quality criteria .......................... 43
3.4 Hierarchy of Emergent Themes ........................................ 47
3.5 Thematic Map for themes on information ............................ 57
3.6 Thematic Map for themes on information source .................. 58
3.7 Theoretical Framework ................................................ 64
3.8 Quality Criteria Categorised by relationship type ................ 66
3.9 Quality Criteria categorised as Objective or Subjective .......... 66

4.1 Venn Diagram showing conceptual representation of health information model as containers. ............................... 71
4.2 Health information model showing implementation as tree data structure. ........................... 72
4.3 Object-Oriented type specialisation .................................. 73
4.4 Health information entry type specialisation ......................... 73
4.5 Quality Criteria and the EHR .......................................... 81
4.6 Three dimensions for health information for QC. .................. 83
4.7 QC - Conceptual processing model .................................... 86
4.8 Quality Criteria Model ................................................ 87
4.9 QCM - Polymorphic elements dependent on QC type ............ 89
4.10 Quality Criteria processing - entries examined .................... 90
4.11 Quality Criteria processing - attestation detected ................ 91
4.12 Quality Criteria processing - C4: Attestation instanced ........ 91
4.13 Quality Criteria processing - C7: Attestor role instanced and added ... 92
4.14 Quality Criteria processing - C7: Attestor role content .......... 93
4.15 Quality Criteria processing - result of full search of entry ...... 94
4.16 Quality Criteria processing - C8: Contextual narrative content ... 95
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>New Zealand Health Sector information volumes and flows circa 2006.</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>New Zealand general practice information volumes and flows circa 2006.</td>
<td>11</td>
</tr>
<tr>
<td>2.3</td>
<td>Key Phrases/Key Word Searches</td>
<td>15</td>
</tr>
<tr>
<td>2.4</td>
<td>Data quality dimensions found in literature - Part 1</td>
<td>23</td>
</tr>
<tr>
<td>2.5</td>
<td>Data quality dimensions found in literature - Part 2</td>
<td>24</td>
</tr>
<tr>
<td>2.6</td>
<td>Example research utilising data quality frameworks.</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>Interviewed GP’s demographic data.</td>
<td>38</td>
</tr>
<tr>
<td>3.2</td>
<td>Coded Responses for candidate QC nodes by GPs ranked by total number of references</td>
<td>45</td>
</tr>
<tr>
<td>3.3</td>
<td>Coded Responses for candidate QC nodes by GPs ranked by rating exercise</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>Coded Responses for Emergent Themes</td>
<td>46</td>
</tr>
<tr>
<td>3.5</td>
<td>Across Themes Matrix Query Results - Part 1</td>
<td>53</td>
</tr>
<tr>
<td>3.6</td>
<td>Across Themes Matrix Query Results - Part 2</td>
<td>54</td>
</tr>
<tr>
<td>3.7</td>
<td>Quality Criteria abbreviated names</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Quality Criteria - groupings</td>
<td>82</td>
</tr>
<tr>
<td>5.1</td>
<td>Quality Criteria potential detection parameters.</td>
<td>116</td>
</tr>
<tr>
<td>5.2</td>
<td>FHIR extension definition fields (partial).</td>
<td>130</td>
</tr>
<tr>
<td>6.1</td>
<td>QC that can be captured by current meta-data use.</td>
<td>135</td>
</tr>
</tbody>
</table>
Listings

5.1 Type imports and namespace declarations. ............................. 107
5.2 Three types from different namespaces. ............................... 109
5.3 Simple global type declaration for eval-type. ....................... 110
5.4 QCM global type actor-type. ........................................... 110
5.5 Global complex type "qc-data". ....................................... 111
5.6 ActorRole with text value = Attestor. ............................... 118
5.7 Comment used to capture attestation. ................................. 119
5.8 Populated Specialty element for ActorType. ......................... 123
C.1 XML schema formal definition for Quality Criteria Model. ........ 179
C.2 QCM instance example from Chapter 5, Figure 5.19. ............... 182
List of Abbreviations

ASTM  American Society for Testing and Materials International
CCR   Continuity of Care Record
CDA   Clinical Document Architecture
EHR   Electronic Health Record
ePA   Electronic Prescribing and Administration
eSCRV Electronic Shared Care Record View
FHIR  Fast Healthcare Interoperability Resources
GP    General Practitioner
HISO  Health Information Standards Organisation
HIQI  Health Information Quality Indicators
HL7   Health Level 7
NZHS  New Zealand Health and Disability Sector
NZMA  New Zealand Medical Association
MOH   Ministry Of Health
OECD  Organisation for Economic Cooperation and Development
OO    Object Oriented
PMS   Practice Management System
POC   Proof Of Concept
QC    Quality Criteria
QCM   Quality Criteria Model
SCR   Summary Care Record
TDQM  Total Data Quality Management
XML   eXtensible Markup Language