Learning Analytics as a Methodology for Understanding Medical Practitioners’ Informal Learning in Online Social Networks

Xin Li
ORCID: 0000-0003-2275-7277

Submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy

September 2017

School of Computing and Information Systems
Melbourne School of Engineering
The University of Melbourne
Abstract

Online Social Networks (OSNs) are used increasingly by medical practitioners to communicate, collaborate and share ideas with each other. This activity can be thought of as a type of informal learning, thus OSNs have important implications for ongoing efforts to support Continuing Professional Development (CPD) in the medical profession. Some of these OSNs are open only to registered medical professional members. In such OSNs the nature and scope of the learning that is occurring are neither well described nor transparent. The challenge of understanding the learning occurring in these OSNs makes it difficult to determine its value as a form of CPD, and more specifically to design and facilitate this type of learning. The objective of this thesis is to examine how learning analytics may serve as a methodology for understanding the process of learning in OSNs for medical practitioners. This research is based on a multi-method design that employs a set of analytic methods, including social network analysis, topic modeling, and context analysis, to investigate different components of the learning process (i.e. learning interaction, learning content and learning context) in the case of a large online discussion forum for medical practitioners. This research found that the proposed analytic methods were useful to gain a detailed understanding of the learning process; this set of methods can reveal what is occurring in such an OSN in an efficient manner and enable the development of intervention strategies for improving learning. The understanding of the learning process in the case study forum led to proposing four intervention strategies for improving the design and operation of OSNs for medical practitioners. These intervention strategies include: organising topics of discussion, recommending subject experts, alerting special discussions, and detecting learning groups. This research elaborates the intervention strategies by describing implementation approaches and relevant application scenarios. This research contributes to greater sophistication in the understanding and improvement of OSNs for medical practitioners’ informal learning, and in turn will support more effective design, operation and evaluation of such learning environments.
Publications

The following peer-reviewed publications were published during the candidature. They were jointly written with my supervisors, Associate Professor Kathleen Gray and Professor Karin Verspoor, and advisory committee member Associate Professor Stephen Barnett, who have given their permission for incorporating parts of these publications into this thesis. I carried out the planning, data collection, and data analysis of the studies and was responsible for the main writing of these publications.


Acknowledgements

It has been a long and at times challenging journey for the past four years. This thesis would not have been finished without the support of many individuals and organisations.

My deep gratitude first goes to my supervisors, Associate Professor Kathleen Gray and Professor Karin Verspoor for providing guidance and invaluable advice. Thanks for their patience and genuine concern, helping me move from an idea to a completed study, and supporting all my travels to research conferences and training events during my candidature.

I am grateful to my advisory committee members, Professor Liz Sonenberg, Professor Pip Pattison, and Professor Fernando Martin Sanchez for their helpful advice and suggestions from various perspectives of this research project.

My appreciation also extends to Associate Professor Jelena Jovanovic, Professor John Sandars, and Associate Professor Stephen Barnett for sharing their specialised knowledge and providing assistance to my research work. Special thanks to Dr Barbara Crump for proofreading this thesis.

Many thanks to the organisations that provided research data and technical support to enable this research, and my workplace for supporting me to undertake PhD study.

Finally, I am thankful my family and friends for all their love and encouragements: for my parents, who supported me in all my pursuits. For my wife Yi, who provided me with moral and emotional support along the way. For my son Cheng, who joined us during the critical stage of this journey and gave me motivation and inspiration. There are too many friends to name individually here, but I will carry them all in my heart.

I am very grateful for having this thoughtful and rewarding journey and thank you for all your kind assistance.
Table of Contents

Abstract ................................................................................................................................. i

Publications ......................................................................................................................... ii

Acknowledgements ........................................................................................................ iii

List of Figures ..................................................................................................................... ix

List of Tables ....................................................................................................................... xi

Abbreviations ..................................................................................................................... xii

Chapter 1 Introduction ....................................................................................................... 1

1.1 Introduction .................................................................................................................. 1

1.2 Research Objectives ................................................................................................... 3

1.3 Research Approach and Methods .............................................................................. 4

1.3.1 Research Paradigm ................................................................................................ 4

1.3.2 Research Approach ............................................................................................... 5

1.3.3 Research Outline .................................................................................................. 7

1.3.4 Limitations ............................................................................................................. 9

1.4 Using a Case Study to Analyse Learning Process .................................................... 10

1.4.1 Case Selection ...................................................................................................... 11

1.4.2 Case Background .................................................................................................. 13

1.4.3 Ethical Considerations ......................................................................................... 16

1.5 Contributions ............................................................................................................. 17

1.6 Outline of the Chapters ............................................................................................. 17

Chapter 2 Literature Review ............................................................................................. 19

2.1 Introduction ................................................................................................................ 19

2.2 CPD in the Medical Profession .................................................................................. 19

2.3 CPD Through Informal Learning ............................................................................. 20
2.3.1 Informal Learning in the Workplace .............................................................. 22
2.3.2 Informal Learning in OSNs ........................................................................ 24
2.4 Theorising Learning in OSNs ......................................................................... 25
  2.4.1 Social Constructivism ............................................................................... 25
  2.4.2 Connectivism ........................................................................................... 26
  2.4.3 Networked Learning ............................................................................... 27
2.5 The Focus of Analysing Learning in OSNs ....................................................... 28
  2.5.1 Analysing Learning Outcome ................................................................... 28
  2.5.2 Analysing Learning Process ...................................................................... 29
2.6 Methods for Analysing Learning in OSNs ...................................................... 31
  2.6.1 Content Analysis .................................................................................... 31
  2.6.2 Learning Analytics .................................................................................. 33
2.7 Summary ......................................................................................................... 35

Chapter 3 Using Social Network Analysis for Analysing Learning Interaction in an
Online Social Network for Medical Practitioners .................................................. 37
  3.1 Introduction ................................................................................................... 37
  3.2 Background and Related Work .................................................................... 38
  3.3 Methods ....................................................................................................... 41
    3.3.1 Phase 1 Study ...................................................................................... 42
    3.3.2 Phase 2 Study ...................................................................................... 46
  3.4 Phase 1 Study Results and Discussion ........................................................... 49
    3.4.1 The Overall Level of Interaction ............................................................ 49
    3.4.2 The Level of Interaction for Individual Users ........................................ 50
    3.4.3 The Patterns of Overall Interaction ...................................................... 56
  3.5 Phase 2 Study Results and Discussion ........................................................... 58
3.5.1 Changes in the Level of Interaction.............................................. 58
3.5.2 Changes in the Patterns of Interaction............................................. 59
3.5.3 The Structural Properties of Interaction.......................................... 60
3.6 Conclusions and Implications.............................................................. 62

Chapter 4 Using Topic Modeling for Discovering Learning Topics in an Online Social Network for Medical Practitioners .......................................................... 65
4.1 Introduction......................................................................................... 65
4.2 Background and Related Work......................................................... 65
  4.2.1 Content Analytics.......................................................................... 65
  4.2.2 Topic Modeling ............................................................................. 66
4.3 Methods............................................................................................... 67
  4.3.1 Data.............................................................................................. 67
  4.3.2 Topic Modeling Using MALLET.................................................... 67
  4.3.3 Procedure..................................................................................... 67
4.4 Results and Discussion ........................................................................ 70
  4.4.1 Topic Model Comparison............................................................... 70
  4.4.2 Topics of Interest........................................................................... 71
  4.4.3 Topic Evaluation........................................................................... 74
4.5 Conclusions and Implications.............................................................. 74

Chapter 5 Using Context Analysis for Analysing the Context of Learning in an Online Social Network for Medical Practitioners ......................................................... 77
5.1 Introduction......................................................................................... 77
5.2 Background and Related Work......................................................... 78
5.3 Proposed Contextual Factors for Medical Practitioners’ Informal Learning in OSNs ........................................................................................................... 79
5.4 Methods............................................................................................... 82
5.4.1 Data ................................................................. 82
5.4.2 Measures ........................................................... 82
5.4.3 Procedures ......................................................... 83
5.5 Results and Discussion ............................................. 84
  5.5.1 Learner Context ................................................. 84
  5.5.2 Environmental Context ....................................... 87
5.6 Conclusions and Implications ..................................... 88

Chapter 6 Validating LA Findings about the Learning Process in an Online Social Network for Medical Practitioners ................................................................. 90
  6.1 Introduction .......................................................... 90
  6.2 Summary of Analytic Findings .................................... 91
  6.3 Methods ................................................................. 92
    6.3.1 Survey Design .................................................. 92
    6.3.2 Survey Participants ............................................ 95
    6.3.3 Procedure ....................................................... 96
  6.4 Survey Results ....................................................... 96
    6.4.1 Sample Characteristics ....................................... 97
    6.4.2 Types of OSN Being Used for Informal Learning ............. 99
    6.4.3 Learning Interaction .......................................... 100
    6.4.4 Learning Interest and Topics ................................ 102
    6.4.5 Learning Experience ......................................... 103
  6.5 Comparison of the Findings about Learning Process .......... 105
    6.5.1 Learning Interaction ......................................... 105
    6.5.2 Learning Content ............................................. 106
    6.5.3 Learning Context ............................................. 107
List of Figures

Figure 1.1. Research design........................................................................................................... 5
Figure 1.2. A snapshot of the key database tables of the forum............................................. 14
Figure 1.3. Cumulative post count by topic................................................................................. 15
Figure 1.4. Overview of chapter structure in this thesis................................................................. 18
Figure 3.1. An example of a sociogram.......................................................................................... 39
Figure 3.2. Examples of network configurations ......................................................................... 40
Figure 3.3. An example of how a network is built from a discussion forum thread 42
Figure 3.4. The original structure of posts and threads ............................................................. 45
Figure 3.5. A sample output of two-mode data matrix................................................................. 45
Figure 3.6. A sample output of one-mode data matrix................................................................. 46
Figure 3.7. Degree centrality distributions for user network ...................................................... 50
Figure 3.8. Closeness centrality distributions for the user network ........................................... 51
Figure 3.9. Betweenness centrality distributions for the user network ....................................... 52
Figure 3.10. Degree centrality distributions for the thread network .......................................... 53
Figure 3.11. Closeness centrality distributions for the thread network ....................................... 54
Figure 3.12. Betweenness centrality distributions for the thread network .................................. 55
Figure 3.13. Visualization of two-mode network....................................................................... 56
Figure 3.14. Visualization of one-mode network ....................................................................... 58
Figure 3.15. Visualization of changes in the user networks....................................................... 59
Figure 4.1. A snapshot of the probability results generated by MALLET ................................. 69
Figure 4.2. The results of F-measure............................................................................................ 74
Figure 5.1. The activity level across 24 hours............................................................................ 87
Figure 6.1. Types of OSN being used for informal learning....................................................... 100
Figure 6.2. The frequency of interaction in OSNs ............................................. 101
Figure 6.3. The learning experience in OSNs ..................................................... 104
List of Tables

Table 1.1. A summary of studies conducted in this research ........................................... 7
Table 3.1. The definitions of SNA measures ....................................................................... 43
Table 3.2. Network-level measures .................................................................................. 50
Table 3.3. Network-level measures of N1, N2, and N3 ......................................................... 59
Table 3.4. Estimation results for structural parameters ...................................................... 60
Table 4.1. The defined variables for topic models .............................................................. 68
Table 4.2. The definition of performance metrics considered for F-measure ............. 70
Table 4.3. The total number of inferred and refined topics for topic models ............ 71
Table 4.4. The refined topics for the selected topic model ................................................. 72
Table 5.1. The critical contextual factors ......................................................................... 80
Table 5.2. The measures of contextual factors for this study ........................................ 82
Table 5.3. The results for learner factors ....................................................................... 85
Table 6.1. Summarised findings about the learning process .............................................. 91
Table 6.2. Mapping between survey questions and analytic findings .................... 93
Table 6.3. The survey questions relating to environmental context factors .......... 95
Table 6.4. Characteristics of the respondents ................................................................. 97
Table 6.5. The clinical topics of discussion ................................................................. 102
Table 6.6. The non-clinical topics of discussion ............................................................ 103
Table 6.7. Comparison of findings about learning interaction .................................... 106
Table 6.8. Comparison of findings about learning content ........................................... 107
Table 6.9. Comparison of findings about learning context .......................................... 108
Table 7.1. Proposed LA methods for analysing informal learning in OSNs .......... 113
Table 7.2. Proposed intervention strategies ................................................................. 117
Abbreviations

CoI Community of Inquiry
CoP Communities of Practice
CPD Continuing Professional Development
ERGMs Exponential Random Graph Models
GP General Practitioner
LA Learning Analytics
MeSH Medical Subject Headings
MOOC Massive Open Online Course
OSN Online Social Network
PLN Personal Learning Network
SAOMs Stochastic Actor-Oriented Models
SNA Social Network Analysis
SoLAR Society for Learning Analytics Research
VCoP Virtual Communities of Practice
Chapter 1 Introduction

1.1 Introduction

Staying up to date, to be able to deliver the best evidence-based care, is crucial for medical practitioners. Medical practitioners need to be lifelong learners as medical knowledge expands and changes rapidly. Continuing Professional Development (CPD) is a mechanism for maintaining, improving and broadening knowledge and skills throughout the working life of medical practitioners (Institute of Medicine, 2010; Jackson, Manley, Martin, & Wright, 2015). In current literature the term ‘medical practitioner’ is used interchangeably with physician, general practitioner, medical doctor, or simply doctor (Rolls, Hansen, Jackson, & Elliott, 2016), and in this thesis the preferred term is medical practitioner.

The aim of CPD is to enhance and improve practice and assist in ensuring the delivery of high quality care for patients; it emphasises self-directed lifelong learning and learning from practice. However, the main approaches to CPD such as educational courses and lectures have been criticized as not being efficient or effective in meeting the needs of many medical practitioners, and thus being inadequate to improve the safety and quality of patient care (McGowan et al., 2012). These limitations, combined with the rising popularity of communicating via the Internet and social media, have led to the emergence of a more informal approach to CPD, namely, making use of Online Social Networks (OSNs) (Gorham, Carter, Nowrouzi, McLean, & Guimond, 2012; McGowan et al., 2012). OSNs are online community platforms that allow medical practitioners to connect, share interests, and interact with each other. Generic examples include but are not limited to: Discussion Forums, Facebook, LinkedIn, and Twitter (Kind & Evans, 2015).

OSNs are increasingly being used by medical practitioners to share medical knowledge, discuss practice management challenges, and network with colleagues. OSNs appear to have advantages in terms of improved access and convenience over in-person CPD activities and formally structured online CPD activities (Millar, Ho, & Carvalho, 2016). Thus, OSNs are considered important in supporting CPD (Ho & Last, 2015). However,
studies to examine the effectiveness of health CPD arising from such OSNs have shown a low impact on professional practice and patient care (Curran, Fleet, & Kirby, 2010; Scales et al., 2011). It has been found that the interaction occurring in OSNs is generally low, and they apparently fail to support the broader learning objectives of the participants (Sandars, Kokotailo, & Singh, 2012). These failures include not being able to test proposed solutions in actual practice through critical reflection in OSNs, and not being able to discuss learning topics that are relevant to the complexity of professional practice (Sandars et al., 2012).

It is evident that challenges remain in developing effective online discussions in OSNs for medical practitioners (Singh, McPherson, & Sandars, 2014). One key thing that may be making it difficult to design and facilitate this type of learning well is that there is a lack of understanding about how learning occurs in OSNs (Institute of Medicine, 2010). This creates uncertainty about how to convey what is occurring in any given OSN. Thus, analysing the process of learning in OSNs could play a major role in realising the full potential of OSNs for medical practitioners’ CPD (Sandars, Jaye, & Walsh, 2014).

Learning Analytics (LA), a recently emerging field, has been proposed as an innovative approach to understand the process of learning in OSNs for medical practitioners (Sandars et al., 2014). LA fulfils a need to derive insights about medical practitioners’ learning from large electronic data sets (so called ‘big data’) such as those generated in OSNs (Ellaway, Pusic, Galbraith, & Cameron, 2014). The Society for Learning Analytics Research (SoLAR, 2013) defines LA as “the measurement, collection, analysis and reporting of data about learners and their contexts, for the purpose of understanding and optimising learning and the environments in which it occurs”. In this thesis, LA is an umbrella term used to refer to different analytic approaches for investigating learning processes.

The aim of this thesis is to investigate the utility of LA to understand the learning process that occurs in OSNs for medical practitioners. In seeking to validate a methodology (i.e. LA) that could improve the way that health professionals work with information technology to enhance their learning, this thesis is situated squarely in the discipline of health informatics (Alexander, Kernohan, & McCullagh, 2004; Shachak, Borycki, & Reis, 2017).
1.2 Research Objectives

This research aims to investigate how LA can be used as a methodology for understanding the process of informal learning in OSNs for medical practitioners. It does this by proposing and testing a set of analytic methods (including social network analysis, topic modeling, and context analysis) to investigate medical practitioners’ informal learning in one such OSN.

The central question this research asks is:

How can learning analytics be used as a methodology for understanding the process of informal learning in online social networks for medical practitioners?

The following sub-questions frame this research and aim to understand the process of informal learning by investigating learning interaction, learning content and learning context:

- How can social network analysis be used to analyse learning interaction?
- How can topic modeling be used to analyse learning content?
- How can context analysis be used to analyse learning context?

This research is necessary because OSNs provide numerous opportunities for medical practitioners to connect professionally, to learn and share insights about their professional practices. However, this activity occurs outside of any formal educational protocols.

Despite the increasing use of OSNs by medical practitioners for their informal learning there is still insufficient understanding about the learning occurring in OSNs. The challenge of understanding this makes it difficult to determine their value as a form of CPD. Thus, a methodical approach to describing and reporting on this form of learning can help OSN providers to design and facilitate this type of learning. It can also help participating individuals and professional accrediting organisations to recognise it as CPD activity.

Use of OSNs for informal learning is likely to scale massively for medical practitioners. So, the ability to monitor these OSN participants’ activities and gain insights into their
learning is going to rely on methods that are computational and can scale to large data sets. In this new way of supporting learning, LA, as one of these computational methods, is becoming very important; but so far, its systematic use in this context is untested.

1.3 Research Approach and Methods

1.3.1 Research Paradigm

All research should make explicit its research paradigm, since it is this underlying set of assumptions that determines how the research is considered valid and which research methods are appropriate (Orlikowski & Baroudi, 1991). According to Creswell (2013), there are four classifications of research paradigm: postpositivism, constructivism, transformative, and pragmatism.

Postpositivist research aims to test theory or describe an experience in order to increase predictive understanding of phenomena; it is most commonly aligned with quantitative methods of data collection and analysis. Constructivist research aims to inductively develop a theory in order to understand phenomena through accessing the meanings that participants assign to them; it may include quantitative and qualitative methods of data collection and analysis, or mixed-methods. Transformative research attempts to critically evaluate and transform the social reality under investigation by following an action agenda; it may include quantitative and qualitative methods of data collection and analysis, or mixed-methods. Pragmatic research focuses on the ‘what’ and ‘how’ of the research problem using all approaches available to understand the problem; it is most aligned with multiple methods and different forms of data collection and analysis.

A pragmatic stance is considered most appropriate for this research, because this research seeks to explore what is required in terms of technological innovation to turn an Internet phenomenon (i.e. OSNs) into a sustainable system that supports medical practitioners’ CPD. Achieving this aim requires employing the methods that are most appropriate to analyse the variety of electronic datasets associated with the learning process in OSNs.
1.3.2 Research Approach

To use LA to investigate the informal learning of medical practitioners in OSNs, a case study strategy was employed. There are two main reasons to employ the case study method to explore the strength and weakness of LA: 1) applying LA in a particular case enables understanding of how it acts on specific datasets related to the learning process in an OSN; 2) an in-depth study of a particular case demonstrates the ability of LA to generate insights into the learning process of medical practitioners in a real-world OSN setting. Figure 1.1 presents the research design. It depicts how the case study was conducted using a LA approach to analyse the learning process occurring in an OSN for medical practitioners. The analytic methods proposed in this thesis were applied to an online discussion forum for medical practitioners, which was the case selected to study in this research.

Figure 1.1. Research design
To explore the selected case in this research, a multi-method research design was used. Multi-method research may be broadly defined as the practice of employing two or more different methods of research within the same study or research program rather than confining the research to the use of a single method (Brewer & Hunter, 2006). The reason for using multi-method research design is because the interactions occurring in a social learning network are highly complex and usually require a multi-dimensional approach to analyse their effects (Joksimovic, Gasevic, & Hatala, 2014). It is necessary to combine different types of analytic methods so that LA can provide a detailed understanding of learning processes (Kovanović, Joksimović, Gašević, Hatala, & Siemens, 2015).

Many researchers employ a multi-method approach to study online learning communities. Most of them combine Social Network Analysis (SNA) with content analysis to examine the quantity and quality of interactions (De Laat, Lally, Lipponen, & Simons, 2007; Haya, Daems, Malzahn, Castellanos, & Hoppe, 2015; Lipponen, Rahikainen, Lallimo, & Hakkarainen, 2003). Others combine SNA with qualitative questionnaires, for example to investigate social awareness of collaborative learning (Lambropoulos, Faulkner, & Culwin, 2012). One multi-dimensional research framework suggested for analysing networked learning combines SNA with content analysis and context analysis to produce a complete picture of the learning process in a social learning network (De Laat & Schreurs, 2013). This framework has been successfully applied in a study of teachers’ learning in the workplace and has been suggested as a useful approach for social network-focused LA (Buckingham Shum & Ferguson, 2012).

In this thesis, we built on the De Laat and Schreurs (2013) multi-method research framework but proposed an alternative approach that considered the raw data about learning process in an OSN, and provided analytical strategies for analysing the data about different components of the learning process (i.e. as shown in Figure 1.1, social network analysis to analyse C1-Learning Interaction; topic modeling to analyse C2-Learning Content; and context analysis to analyse C3-Learning Context). The analytic studies were conducted sequentially, and for each study conclusions and inferences were developed that represented its method and dataset. Measures were taken to avoid a
situation where the findings from one study biased the next, including: continuous oversight by two research supervisors with differing methodological expertise, and progressive presentation of the findings to the case study site owner for critical feedback.

Once all analytic studies were completed, validation of the findings was sought through a cross-sectional survey with a similar group of medical practitioners who used OSNs for informal learning. Such validation is necessary to generalise the analytic findings for a case study and understand the strength and weakness of the applied analytic methods.

Finally, intervention strategies to improve the informal learning of medical practitioners in OSNs were derived through the use of this framework in this thesis. The next section provides an overview of the studies conducted in this research and outlines the methods employed in each study.

1.3.3 Research Outline

The objectives, data collected, methods and expected outcomes for each study are presented in Table 1.1.

Table 1.1. A summary of studies conducted in this research

<table>
<thead>
<tr>
<th>Study</th>
<th>Objective</th>
<th>Data collected</th>
<th>Method(s)</th>
<th>Expected outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Understanding overall learning interaction</td>
<td>All users (n = 621) and threads (n = 723) in the online discussion forum</td>
<td>SNA network-level measures</td>
<td>The level of overall interaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SNA individual-level measures</td>
<td>The level of interaction for individuals</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SNA visual analysis</td>
<td>The patterns of overall interaction</td>
</tr>
<tr>
<td>2</td>
<td>Understanding changes in learning interaction</td>
<td>Active users (n = 48) in the online discussion forum and their threads from the year 2012 to 2014 (n = 106, 101, 51)</td>
<td>SNA network-level measures</td>
<td>Changes in the level of interaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SNA visual analysis</td>
<td>Changes in the patterns of interaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Stochastic Actor-Oriented Models</td>
<td>The structural properties of interaction</td>
</tr>
<tr>
<td>Study</td>
<td>Objective</td>
<td>Data collected</td>
<td>Method(s)</td>
<td>Expected outcome</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
<td>----------------</td>
<td>-----------</td>
<td>------------------</td>
</tr>
<tr>
<td>3</td>
<td>Understanding learning content</td>
<td>Posts of active users in the online discussion forum (n = 1604)</td>
<td>Topic modeling</td>
<td>Topics of interest</td>
</tr>
<tr>
<td>4</td>
<td>Understanding learning context</td>
<td>Membership data and usage log files maintained in the online discussion forum for active users (n = 48)</td>
<td>Quantitative context analysis</td>
<td>Learner characteristics and learning environment</td>
</tr>
<tr>
<td>5</td>
<td>Validating the findings about the learning process</td>
<td>Valid survey responses from medical practitioners (n = 149)</td>
<td>Survey</td>
<td>Validated findings about learning interaction, content and context</td>
</tr>
</tbody>
</table>

In Study 1 SNA is used to investigate the network of interactions in the forum. These interactions may provide insight into learning behaviours. Together, learning behaviours effectively signify a learning process, and can be important indicators that learning may have occurred. The aim is to understand the level of overall interaction in the forum through the analysis of all users’ connections derived from their participation threads. A thread (also called a topic) is contained in the forum and can contain any number of posts. SNA network-level measures (include density, centralization, diameter and average path length) are used to understand the overall activity status of the network. SNA individual-level measures (include degree centrality, betweenness centrality, and closeness centrality) are performed to identify the activity levels of the individual users. SNA visual analysis is performed to enrich the findings of interaction.

In Study 2 SNA is used to understand the changes in patterns of interaction over time among the active users in the forum by applying SNA network-level measures (include density, centralization, diameter and average path length). SNA visual analysis is performed to enrich the findings of interaction. Stochastic actor-oriented models (the statistical models for longitudinal networks) (Snijders, 1996) are used to evaluate associated structural network effects and individual characteristics that could explain these changes as features of the learning process.
In Study 3 topic modeling is used to examine the learning content (i.e. posts) produced by the active users in the forum. Topic modeling is a statistical method that analyses the words of the original texts to discover the themes that run through them; it is used for discovering the topics of interest for the active participants. Since sharing knowledge and profession-related information is generally considered as a form of informal learning (Corrigan & Curtis, 2017; Hager & Halliday, 2007), the topics of interest emerge from OSN users are the reflection of what knowledge is shared and hence, provides indirect evidence of learning.

In Study 4 quantitative context analysis is used to understand the learning context that influences the understanding and interpretation of learning process, to the extent that context data could be gathered from the forum and associated files for the active users. The learning context is analysed using descriptive statistics about several context elements that are considered useful, based on the literature for understanding the learning context in OSNs for medical practitioners. These context elements include learner factors (demographics, work experience, educational background, and learning interest), and environmental factors (time, location, activity, relations, recognition and application).

In Study 5 an online survey was conducted to validate the summarised findings from the three LA approaches. Participation in the survey was open to any medical practitioners who used OSNs for informal learning in a manner similar to the case study, not limited to users of the specific OSN used in the case study. Thus cross-validation was done, involving a group not analysed in the initial studies.

1.3.4 Limitations

Applying the proposed LA approaches had three major limitations as research to understand medical practitioners’ informal learning in OSNs.

First, it was not possible to understand the learning process of passive users (i.e. those who may learn by reading the online posts but do not participate in any discussion), since the forum under study was not able to track the activities of passive users. This was because the forum under study had not implemented technology that would have enabled it to track the activities of passive users.
Second, the findings of the learning process were identified based on an online discussion forum for medical practitioners, so they may not generalise to the other types of health professionals. However, in our review of OSNs for health professionals (in Appendix A), we identified that medical practitioners were the major users of current OSNs for health professionals, so a study of medical practitioners was warranted as a way to provide a meaningful starting point for evidence about this aspect of health professional learning.

Third, very few context data about individual participants were collected in the forum, so the analysis and understanding of context of learning in this specific OSN were limited. It is acknowledged that collection (or integration) of additional context information could enhance the results of the present study, or could lead to a different interpretation of the learning process in the forum.

The next section will give background to the case study strategy, describe the process involved in selecting the online forum for this research, present preliminary analysis of key details about the forum, and discuss the ethical issues involved in the research.

1.4 Using a Case Study to Analyse Learning Process

Case studies are used extensively in a wide variety of disciplines, particularly in the social sciences. A case study is “an empirical inquiry that investigates a contemporary phenomenon within its real-life context” (Yin, 2013, p. 16). It allows the researcher to obtain an in-depth, multi-faceted understanding of an issue, event or phenomenon of interest. One of the criticisms of the case study method is that the case under study may not be representative of a wider social setting and therefore it is argued that the results of the research cannot be used to make generalisations. However, through validation, results may be generalised (Yin, 2013).

There are three main types of case study: intrinsic, instrumental and collective (Stake, 2008). An intrinsic case study is typically undertaken to learn about a unique phenomenon. In contrast, the instrumental case study uses a particular case to provide insights into a wider issue or phenomenon. The case is of secondary interest to what may be ‘typical’; it plays a supportive role and facilitates the understanding of something more broadly applicable. The collective case study involves studying
multiple cases simultaneously or sequentially to generate a still broader appreciation of a particular issue. The research in this thesis may be considered as an instrumental case study, as the research objective is not only to understand the learning process for a particular case (i.e. the chosen online discussion forum), but also to gain insights into the learning process of medical practitioners in OSNs in general based on the chosen OSN.

1.4.1 Case Selection

Several types of OSN are changing the way that medical practitioners are learning and collaborating professionally including Discussion Forum, Facebook, LinkedIn, and Twitter (Panahi, Watson, & Partridge, 2012; Rolls et al., 2016). Increasingly popular are private online communities that require the verification of medical practitioners’ credentials before membership and offer targeted professional content and services. These private OSNs are designed for the explicit use of medical practitioners and are platforms in which medical practitioners bring together their specialised knowledge and individual experiences to share within the community online, and thus allow for informal learning.

The main interest of this thesis is to explore a methodology for understanding informal learning in the context of the growing phenomenon of private OSNs for medical practitioners. Therefore, three criteria were defined for selecting an OSN case study for this research: 1) It had to be an online community platform that allows medical practitioners to connect, share interests, and interact with each other for informal learning, where informal learning is defined as non-formal learning, i.e. unstructured, unofficial, and unscheduled learning. 2) It had to be a private online community that requires verification of professionals’ credentials before membership. 3) It had to be an online community with at least 1,000 members.

An Internet search was conducted in July 2013 to identify potential OSNs based on the defined case selection criteria. Appendix A presents an indicative list of the OSNs retrieved from this search, which included both Australian and internationally recognised OSNs. In addition to being private groups (that is, not open to participation by the general public) some of these OSNs operate as private profit-making enterprises.
The details of their business operations have commercial in-confidence aspects. Therefore, for confidentiality, the thesis may not disclose whether the OSN that was finally selected for this research is included in or excluded from this list.

A review of these OSNs found that medical practitioners were the main users of OSNs. Most OSNs (70%) had free membership, whereas others required paid membership to get access, or else access was linked with belonging to a professional organisation (e.g. Royal Australian College of General Practitioners (RACGP), Australian Medical Association (AMA)). Although these OSNs differed in terms of their target users, look-and-feel and features, they all promoted knowledge sharing, collaboration, and development of connections. For example, Sermo\(^1\) provided medical practitioners with the ability to crowd-source answers to common and not-so-common clinical questions; Doximity\(^2\) was focused on developing professional connections; QuantiaMD\(^3\) encouraged social learning and collaboration through interaction and learning from experts and peers. Other types of OSNs included the private groups which operated as discussion boards within public platforms such as Facebook or LinkedIn, used mostly by those with similar clinical interests who wish to share information and find answers (MacWalter, McKay, & Bowie, 2016).

The selection process involved confirming that the researcher would be permitted the required access to an OSN in order to conduct the research. This proved to be challenging. Very few of these OSNs provide a contact method (e.g. email, phone, physical address, or online contact form) that allows the public to make inquiries. The online registration form requires details to be supplied about a medical practitioner’s identifying board registration or government license to practice. This made it difficult for the non-medical-practitioner researcher, as an outsider, to contact OSN site owners to discuss the access required for this research. For those few OSNs providing contact details, the researcher tried to contact them but had little success. Most said that they had a policy of not sharing any system data with outside parties regardless of use. After numerous attempts, one OSN’s owners agreed to provide the access to their OSN for

---

1 www.sermo.com
2 www.doximity.com
3 www.quantiamd.com
this research. This OSN is an online discussion forum restricted to medical practitioners, established in 2009 and run by an online health CPD provider.

There were three reasons to choose this forum as the case to study for this research: 1) The site owners were amenable to research in a way that almost no other OSNs were, and provided not only data but also support to access it – the researcher was granted access to their support staff and the back-end data which were essential for analysis (feasibility). 2) It is a large online community with more than 10,000 medical practitioners, which offered the opportunity to learn through studying a ‘typical’ case (authenticity). 3) It allowed study of the changes occurring in the learning process over time because it was established in 2009 and has data available across several years, to cater for longitudinal analysis (temporality). The next section provides more detailed background about this online discussion forum.

1.4.2 Case Background

The forum was set up with the aim of fostering a community of medical practitioners who move towards using technology as a tool for creating better patient outcomes. It allowed the users to discuss industry issues, share best practices and promote conversation within the health community. All medical practitioners seeking to join were verified by entering their registration-to-practice details, which were then checked against the registration database before users gain entry. The forum was moderated and facilitated by a panel of medical experts.

From the initial discussions with the site owners, it was evident that there was an interest in gaining a better understanding of learning activities occurring in the forum to improve the learning experience of community members and keep them better engaged over time. The owners made many suggestions of potential areas for research, which included: understanding the users’ participation behaviours, why people leave the community, who are the active contributors and how to identify the key people early on and provide the support they need, as well as knowing the users’ expectations of the community and their interest in providing relevant learning materials.

Prior to applying any analytic methods, a preliminary analysis was conducted to further understand the data collected in the forum. The analysis was enabled by the technical
support and cooperation of the operational staff of the forum. Summaries of the analysis results are presented in the following paragraphs.

Figure 1.2 presents a snapshot of the key database tables in relation to the forum. By querying the tables se_users and se_profile_info, it was found that there was a total of 10,056 registered users in the forum, mainly with backgrounds of General Practice (n=7632), but also including Nursing (n=775), Cardiology (n=125), and General Medicine (n=122). Of these registered users, 5,839 users never logged into the system and 955 users logged into the system only once or twice, which means only 32.4% (n=3,262) of the users used the forum to any extent useful for research. This included both active and passive users.

Figure 1.2. A snapshot of the key database tables of the forum
By further querying the tables `se_forum_posts` and `se_forum_categories`, it was found that the forum had low participation overall. There were not many active users. In the study period 2009 – 2014, there were 40 sub-forums, with a total of 723 threads and 7,089 posts. 621 users posted at least once. Of these 621, most users made less than 100 posts; five made more than 200 posts; however, three of them were known to have formal roles as moderators. Eighty-one percent (n=2641) of users were passive users, as calculated by counting the number of the active forum participants.

When the post count for each sub-forum was analysed, these post events were found not only to be varied over time but also to be different across different sub-forums as shown in Figure 1.3. This preliminary inspection of these sub-forums showed that the topics for discussion tended to be highly professionally focused, whether on the business challenges of running a successful practice or on getting help with a difficult diagnosis. ‘Politics/IT/Administration’, ‘Doctors’ Life’, and ‘Cardiovascular/Vascular’ were the most popular sub-forums. Many topics were discussed in 2011 but stopped after 2012. Discussion on specialised topics (e.g. Paediatrics, Neurology) was not very intense but had periodic spikes, in particular, in the years 2012 and 2014.

![Figure 1.3. Cumulative post count by topic](image)

Figure 1.3. Cumulative post count by topic
1.4.3 Ethical Considerations

Meeting human research ethics standards has been considered one of the key challenges in doing LA research since there is no agreed method for researchers to obtain informed and ongoing consent for the use of data (Nunn, Avella, Kanai, & Kebritchi, 2016). As indicated earlier, confirming access to an OSN for the purpose of a case study was a major issue encountered in the early stage of this research because of these ethical concerns.

Since the chosen online discussion forum was closed to external observers, gaining access to the data from the forum involved formal approval from the site owners, a commercial enterprise. No explicit consent was obtained from the users of the forum, because the forum is privately owned. As stated in their site Terms of Service, “all content appearing on or included in the site and produced, published, displayed, transmitted or created by [company], including site layout, design, images, programs, text and other information (collectively, the "Content") is the property of [company].”

As part of the Human Research Ethics approval for this research, a written plain language statement was provided to the site owners to gain their written consent to use their data for the proposed research. As requested by the host company, for protecting their members’ confidentiality, the student researcher also signed a confidentiality agreement with the host company. The confidentiality agreement acknowledged that the student researcher could use the online discussion forum as a research site and gave permission for the student researcher to: 1) share the data and information from the online discussion forum with other nominated researchers for assisting with research advice and support and collaborating on publications; and 2) publish the research findings of the online discussion forum within academic publications (including the PhD thesis, journal articles, book chapters, conference papers, the media, working papers) so long as the identity of the online discussion forum and its participants were anonymised.

---

4 This study has been approved by the Human Research Ethics Committee of the University of Melbourne (HREC Reference #1340990).
1.5 Contributions

This research makes major contributions in three areas.

1. This research proposes and validates a LA framework for investigating the informal learning of medical practitioners in OSNs. It also proposes and demonstrates the general capability of a set of analytic methods to investigate different components of the learning process in OSNs, on a wider scale.

2. This research provides an understanding of the learning process in the case of one OSN for medical practitioners, which contributes to understanding how informal learning through OSNs functions effectively as a form of CPD. The results are used to infer intervention strategies that can contribute to improve the quality of this form of learning.

3. This research helps to build an evidence base about learning processes that can be supported in OSNs. This will allow medical practitioners, OSN hosts and CPD accreditation agencies to evaluate other OSNs as a form of CPD.

1.6 Outline of the Chapters

Figure 1.4 indicates the thesis structure in relation to the aims of the research.
Chapter 2 presents a detailed literature review on the conceptual foundations of this thesis. In Chapters 3, 4 and 5 the three analytic studies conducted to understand the learning process in the case study setting are presented. Each chapter discusses relevant work and presents the details of the analytic approaches including associated research methods and procedures, as well as the results and implications. Chapter 6 presents the synthesis of findings from the three analytic studies, and the validation of these findings about the learning process. Chapter 7 discusses the key contributions of this thesis, including: the LA methodology for understanding informal learning of medical practitioners in OSNs that has been proposed in this thesis; the intervention strategies to improve the informal learning of medical practitioners in OSNs that were derived through the use of this framework in this thesis; the opportunities for future work to refine this approach to OSNs as a form of CPD in the medical profession; and reflections on how LAs may assist better use of OSNs in future to support medical practitioners’ informal learning.
Chapter 2 Literature Review

2.1 Introduction

This chapter reviews the relevant literature for this research. Section 2.2 provides the background to CPD for the medical profession. Section 2.3 outlines the informal approach to CPD occurring in workplaces and OSNs. Section 2.4 discusses the theories that are relevant for learning in OSNs. Section 2.5 describes how learning in OSNs for medical practitioners has been analysed previously. Section 2.6 discusses the methods developed for analysing the process of informal learning including traditional approach of content analysis and the novel approach of Learning Analytics (LA). Finally, this literature review and the research opportunities that arise from the current literature are summarised.

2.2 CPD in the Medical Profession

As introduced in Chapter 1, CPD is a mechanism for broadening knowledge and skills throughout the life of medical practitioners (Institute of Medicine, 2010). CPD is a lifelong commitment, which encompasses all forms of learning and professional development relevant to the practice of individual medical practitioners. The ultimate aim of CPD in the medical profession is to promote up-to-date and high-quality patient care by ensuring that medical practitioners have access to the necessary learning opportunities to maintain and improve their ability to practice (Légaré et al., 2015). CPD also enables an individual to verify that their professional practice is similar to that of their professional peers (Schostak et al., 2010).

Participation in CPD has been introduced as a mandatory revalidation scheme in most developed countries. Revalidation is a process by which doctors are required to demonstrate to a regulatory body or professional organisation that they maintain competence and excellence in their practice, and to meet the requirement for registration, also known as licensing (Archer & de Bere, 2013). For example, in Australia, the medical practitioners who are engaged in any form of medical practice are required to participate in CPD and complete a minimum of 50 hours of CPD per year (Medical Board of Australia, 2016).
Nevertheless, concerns have been raised over the last decade regarding the effectiveness of CPD (Sklar, 2016). In a major review of CPD effectiveness for health professionals by the Institute of Medicine (2010, p. 2), it was noted that “there are major flaws in the way continuing education is conducted, financed, regulated, and evaluated.” The focus of current CPD is on meeting regulatory requirements and enforcing minimal, narrowly defined competencies. Therefore, the CPD through the mandatory revalidation schemes does not encourage learning behaviours focused on addressing the particular learning needs of medical practitioners.

Current CPD of medical practitioners is delivered primarily using formal approaches. However, didactic methods (e.g. lectures, conferences and workshops) alone have little direct impact on changing professional practices as most CPD activities delivered through these methods focus on dissemination of knowledge rather than the application of knowledge (Légaré et al., 2015; Sachdeva, 2016). The formal approaches do not offer opportunities for collegial engagement which are necessary for confirmation of practice and of medical practitioners’ own competence (Cantillon, 2016).

Further, because medical practice changes rapidly, there are acknowledged gaps between what research evidence suggests for optimal care and the care provided by medical practitioners. The formal CPD approach does not provide them with readily accessible resources to address many of the case related questions that arise in their practice. Thus, there is a need to identify informal approaches for doing CPD that promote effective and timely knowledge dissemination (Millar et al., 2016), and focus on meeting individual learning needs of medical practitioners in order to help them improve. The next section introduces informal approaches to CPD, and describes how they are currently applied by medical practitioners.

### 2.3 CPD Through Informal Learning

To improve the delivery and effectiveness of CPD, one approach to CPD that is being increasingly developed is self-directed learning (Lindsay, Wooltorton, Hendry, Williams, & Wells, 2016). Self-directed learning has been suggested as a promising methodology for lifelong learning in medical education (Murad & Varkey, 2008). Further, the Liaison Committee on Medical Education (LCME) in the United States
endorsed accreditation standards in 2004 that promote flexibility and innovation in learning and provide medical students with skills necessary for self-directed learning (Simon & Aschonbrener, 2005).

To date, the most common definition of self-directed learning comes from Knowles (1975, p. 18), who described it as “a process in which individuals take the initiative, with or without the help of others, in diagnosing their learning needs, formulating learning goals, identifying human and material resources, choosing and implementing appropriate learning strategies, and evaluating learning outcomes”. In the Personal Responsibility Orientation (PRO) to self-directed learning model (Brockett & Hiemstra, 1991, p. 26), the researchers described self-directed learning as a composite of “instructional method processes (self-directed learning) and personality characteristics of the individual learner (learner self-direction)”, where “the individual assumes primary responsibility for a learning experience”.

Self-directed learning relies primarily on informal learning methods. Many researchers have attempted to define informal learning but there is still no agreed definition of the concept (Fahlman, 2013). Hager and Halliday (2007, p. 236) say that the key features of informal learning are that it is “indeterminate, opportunistic, involves internal and external goods, and is an ongoing process”. The aspects of indeterminacy are that such learning is significantly contextual, arises from practice, and can be viewed as a continually evolving process; and because informal learning situations are typically indeterminate, much valuable informal learning is opportunistic and contingent both at the individual and communal levels. Informal learning is enhanced in situations that emphasise both internal (e.g. values, behaviours) and external (e.g. skills base, demonstrated competency. Informal learning is considered to be a process in which “learning is about ongoing becoming rather than about attaining a particular state as a preparation for something else” (Hager & Halliday, 2007, p. 238).

Eraut (2004, p. 250) defines informal learning as “learning that comes closer to the informal end than the formal end of a continuum. Characteristics of the informal end of the continuum of formality include implicit, unintended, opportunistic and unstructured learning and the absence of a teacher”. The author considers that this is the kind of learning in which the individual often has little awareness of having learned something,
since the learning may not be immediately translated into their practice but is likely to be stored and applied when the appropriate opportunity arises in the future.

Corrigan and Curtis (2017, p. 23) consider that “informal learning is largely a process of social interaction and collaboration to establish a common understanding and share knowledge and experience”. The aspect of social interaction and collaboration for informal learning has been considered to be particularly important for medical practitioners, as it offers opportunities for interactions that can facilitate behaviour changes and so lead to practice improvements (Ibrahim, 2015; Parboosingh, Reed, Caldwell Palmer, & Bernstein, 2011). This type of informal learning occurs often with colleagues in the workplace, and more recently in OSN settings (McGowan et al., 2012; Moorhead et al., 2013).

In the upcoming sub-sections, we will start by reviewing the informal learning that occurs in medical practitioners’ workplaces, then the literature about their informal learning in OSNs.

2.3.1 Informal Learning in the Workplace

How medical practitioners learn in the workplace has been researched extensively (Cantillon, 2016; Van de Wiel, Van den Bossche, Janssen, & Jossberger, 2011). Evidence suggests that most of the learning that takes place at work in healthcare is informal (Wihak & Hall, 2011). Informal learning at work may occur when there are identified learning needs that arise through daily practice. The learner employs a deliberate process to access a wide variety of learning resources to meet their learning needs (Kelly & Hager, 2015).

Eraut (2004) found that there were four main types of work activity that regularly gave rise to learning: 1) participation in group activities; 2) working alongside others; 3) tackling challenging tasks; and 4) working with clients. Also, medical practitioners learned most from the patient cases they encountered and the discussions with colleagues about patients (Van de Wiel et al., 2011).

The study of workplace learning for medical practitioners often has been conducted using Social Network Analysis (SNA) as a method. The social network data in these
studies were mostly collected through traditional research methods, such as surveys (Chambers, Wilson, Thompson, & Harden, 2012). For example, Creswick and Westbrook (2007) used SNA to examine how a network of staff in a renal ward of a metropolitan Australian teaching hospital sought medication advice. They found that most communication occurred within professional groups. Medication advice was sought from several key individuals in the ward both within and across professional groups. It is interesting that through SNA, they found that there was a relatively low level of advice-seeking about medication-related decisions and tasks. This suggests that SNA may provide different insights than other methods, and may help test some assumptions about informal learning based on evidence collected through conventional methods.

Similarly, Keating, Ayanian, Cleary, and Marsden (2007) evaluated the network of influential discussions among primary care physicians in a US hospital-based academic practice. SNA was used to compare communication patterns and to examine factors predictive of a physician’s location in the network. Physicians obtained information from colleagues with greater expertise and experience as well as from colleagues who were accessible based on location and schedule. Physicians were found to be more likely to have discussions with physicians of the same gender.

Wagter, van de Bunt, Honing, Eckenhausen, and Scherbier (2012) employed SNA to examine informal interprofessional learning among senior physicians, residents, and nurses within an intensive and medium care unit in a large hospital. They confirmed that stability and physical proximity of team members not only provide continuity of patient care, but may also offer opportunities to develop professional working relationships. More importantly, their study supports the assertion that SNA is a useful research tool to study informal learning relations in the workplace.

Understanding the structure, characteristics and functions of networks is vital (Cunningham et al., 2012). The use of SNA to understand communication and collaboration of medical practitioners in their workplaces is increasing rapidly and has proven to be a useful approach, but current studies tended to use single point-in-time, small-group designs that limit our understanding of the potential of this approach. When such network data are automatically collected online, longitudinal study can be
conducted on a large scale through more sophisticated analysis techniques. In the next section, we will review the informal learning occurring in OSNs for medical practitioners.

2.3.2 Informal Learning in OSNs

Over the past decade, OSNs have received increasing attention as a platform for sharing medical knowledge and approaches to practice among medical practitioners; they have been considered as having important implications for ongoing efforts to support CPD (Ho & Last, 2015). An OSN refers to an online community, which can be formed through a variety of social networking technologies (e.g. Discussion Forums, Twitter, LinkedIn, YouTube, and Facebook), that allows people to connect, share interests, and interact with each other in different ways (Singh et al., 2014). Research findings suggest that OSNs could complement (or even replace other modes of) CPD for doctors as an informal learning channel and that doctors will adopt specific online communities for the purpose of exchanging medical information with each other (McGowan et al., 2012).

OSNs are said to contribute to medical practitioners’ informal learning since they offer ease of access and the opportunity to interact with peers who provide support for needed changes in practice (Lindsay et al., 2016). This helps reduce professional isolation from peers, long recognised as an important risk factor for reduced medical competence (Cantillon, 2016; Norton, Dunn, & Soberman, 1997). According to St George (2006), professional isolation usually describes doctors who, for one reason or another, are working without regular contact with other doctors. It has been demonstrated that engagement with peers in OSNs reduces feelings of professional isolation among doctors and leads to demonstrable improvement in practice (Barnett et al., 2014).

Further, OSNs allow medical practitioners to share tacit knowledge with peers which is essential to ensure that explicit knowledge can be translated into their daily professional practice (Panahi, Watson, & Partridge, 2013). The professional practice of medical practitioners is characterised by dealing with complex situations that require the integration of explicit and tacit knowledge. The explicit knowledge includes generalised information that is often of little practical value for dealing effectively with the unique situations that medical practitioners face in their daily practice, whereas the tacit
knowledge is based on clinical skills that are built through experience and reflection and shared through social interaction with others. Such tacit knowledge has a significant impact on the quality and delivery of patient care (Panahi et al., 2012).

2.4 Theorising Learning in OSNs

OSNs are a social phenomenon - they connect learners, they create networks of learning. Explanations of learning and professional development in OSNs are based on a wide range of theories. Three of these theories – social constructivism, connectivism, and networked learning – have particular relevance for the present research. This section outlines these theories and explains how they are useful to understand learning in OSNs as it is explored in this thesis.

2.4.1 Social Constructivism

Social constructivism is classified within the constructivist paradigm (Joey, 2017). Constructivism (Vygotsky, 1978) views learning as the individualised construction of meanings by the learner herself or himself. According to constructivism, learners themselves actively and often proactively create knowledge, based on interaction with the world and with other learners. Social constructivism is considered as a development of constructivism by highlighting the importance of social group working. Compared to constructivism, social constructivism (Vygotsky, 1987) is said to favor collective over individual meaning, and to consider how meaning is constructed through communication and interactions with others, with knowledge being created as the result of consensus (Brown, Collins, & Duguid, 1989).

Social constructivism has been helpful to design and evaluate the use of technological developments such as Web 2.0 (Ravenscroft, 2009), or the social web (García-Peñalvo, Colomo-Palacios, & Lytras, 2012), to support forming social learning groups where learning mainly takes place through exchanging ideas and information, determining solutions, and building innovations online (McMahon, 1997). For example, Moodle, a widely used online Learning Management System, is based on this theory. Moodle’s design emphasises the social interactive aspect of learning where learners participate in composing and creating learning materials. The participatory aspect means learners
acquire new knowledge from interaction based on their background knowledge and through interaction within like-minded peers (Forment, 2006).

Another useful concept arising from this theory of socially constructed learning is Communities of Practice (CoP). CoP are defined as “groups of people who share a concern, a set of problems, or a passion about a topic, and who deepen their knowledge and expertise in this area by interacting on an ongoing basis” (Wenger, Richard, & William, 2002, p. 4). The CoP model consists of four domains: learning as community, learning as meaning, learning as identity and learning as practice (Lave & Wenger, 1998). Building on the Lave and Wenger (1998) notion of community of practice, the evolution of OSNs has been considered as a way to provide Virtual Communities of Practice (VCoP) (de Souza, Farinelli, Jamil, de Vasconcelos, & Dias, 2014), which have been defined as a “network of individuals who share a domain of interest about which they communicate online” (Gannon-Leary & Fontainha, 2007).

2.4.2 Connectivism

It is acknowledged that constructivism theories have emerged within the context of traditional methods of learning and teaching, and therefore, they do not fully take account of the influences of technological innovation on learning engagement. To respond this need, connectivism (Siemens, 2005) was presented as “a learning theory for the digital age”, which views learning as making and continuously updating connections with others and with knowledge sources. According to connectivism, learning resides in networks and these networks are formed from the social interactions between learners. In this context, learning is not purely content driven, but begins and is maintained through the connections that learners make. In navigating knowledge, learners need the ability to assess valid options and incorrect information. The following summarises the major principles of connectivism (Siemens, 2005):

- **Learning and knowledge rest in diversity of opinion.**
- **Learning is a process of connecting specialized nodes or information sources.**
- **Learning may reside in non-human appliances.**
- **Capacity to know is more critical than what is currently known.**
- **Nurturing and maintaining connections is needed to facilitate continual learning.**
• *Ability to see connections between fields, ideas, and concepts is a core skill.*

• *Accurate, up-to-date knowledge (currency) is the aim of all connectivist learning activities.*

• *Decision-making is a learning process in itself.* (p. 5)

Connectivism has been considered as a theoretical framework for understanding learning occurring in technology-enabled environments (Goldie, 2016). There is consensus that connectivism is a particularly relevant learning theory informing the use of social media in medical education (Flynn, Jalali, & Moreau, 2015). Although some scholars believe that connectivism is not a new theory of learning but rather a pedagogical approach (Verhagen, 2006), it is well accepted that it at least connects existing theories and builds on them by incorporating social technology to advance education (Kop & Hill, 2008). Accordingly, Massive Open Online Courses (MOOCs), adopted by many leading campus-based universities, are an example of a connected and networked learning approach. The way that learning occurs in a MOOC is determined by participants according to their abilities and educational backgrounds, in their own personal learning environments. MOOC participants can benefit from Rich Site Summary (RSS) technologies which accumulate different learning materials created through online activities such as blogs, discussion forums and wikis. Each learner can combine activities and material based on their own preferences and most importantly to suit their own learning objectives.

### 2.4.3 Networked Learning

Further, the theory of connectivism is considered as one of the networked learning theories that are useful to understand CPD in OSNs. Networked learning was defined by Goodyear (2004, p. 2) as the “learning in which information and communications technology is used to promote connections between learners, and between a learning network and its learning resources”. The study of networked learning aims to understand the learning process by investigating how people develop and maintain a web of social relations for their learning. It focuses on the diversity of social relationships (rather than the development of long-lasting relationships), as well as the value this creates for learning.
Both theories of connectivism and of networked learning are closely related to the methodologies of social network theory. This highlights the need for understanding the structure of a network since this helps reveal important evidence of the information flow and knowledge sharing practices. SNA is proposed as the method to study the structure of a network in both theories. It offers a way to proceed from these theories to undertake studies of learning interaction in OSNs (Schrire, 2006).

2.5 The Focus of Analysing Learning in OSNs

Since OSNs are increasingly used by medical practitioners, it is necessary to understand how their learning occurs in OSNs (Singh et al., 2014). The main way of analysing learning in conventional CPD has an emphasis on learning outcome, aimed at understanding the effectiveness of learning. However, the focus of analysing learning occurred in OSNs is to investigate learning process, with the aim of understanding how and why the learner participated (Sandars et al., 2014). Reviewing current literature identifies that the analysis of medical practitioners’ learning in OSNs is focused on both learning outcome and learning process. In the following sections, these two areas of analysis are discussed.

2.5.1 Analysing Learning Outcome

There is a need to demonstrate the effectiveness of CPD, especially its impact on patient care and health outcomes (Grant, 1999; Ratanawongsa et al., 2008). The Kirkpatrick model of evaluation of training interventions (Kirkpatrick & Kirkpatrick, 2005) is a widely used approach to measuring effectiveness of CPD in OSNs. The model suggests the evaluation of development levels (i.e. level 1 – reaction, level 2 – learning, level 3 – behaviour, level 4 – results) and assumes a linear relationship between educational input and impact of learning. Level 1 – reaction measures the individual participants’ reaction to the program; level 2 – learning measures what participants have learned from involvement in the program; level 3 – behaviour refers to the extent to which the participants apply what they learned; level 4 – results measures the extent to which the training program has made an organisation-level difference in relation to the need it was designed to meet.
Current literature lacks evidence about the efficacy of OSNs to support the CPD that results in changes to practice or patient outcomes. Most evaluations of learning effectiveness in OSNs have been at level 1 and 2. For example, Curran et al. (2010) used pre- and post-knowledge assessment surveys to evaluate the outcome of facilitated asynchronous discussion among medical practitioners but only examined satisfaction, knowledge and confidence to explore how and why their practice changed. Barnett et al. (2014) conducted surveys and interviews with medical practitioners to understand the usefulness of VCoP. Based on self-reported data, they suggested that training, technology, and valuable content may help build a successful VCoP.

Increasingly there have been attempts to measure the level 3 effectiveness of learning from a behavioural perspective. However, few studies have used valid methods and demonstrated the effectiveness or impact of the learning in OSNs. Earlier studies have shown evidence of low impact on professional practice and patient care (Bloom, 2005; Curran & Fleet, 2005; Wutoh, Boren, & Balas, 2004). More recently, Walsh, Sandars, Kapoor, and Siddiqui (2010) evaluated the impact of an e-learning module on a clinical guideline. Based on a survey, they reported that most users believed that the module had helped them put the guideline into practice. However, the reported behaviour changes were based on self-reporting, and also the impact over the long-term is unknown. Further, Scales et al. (2011) reported improved care as a result of interactive learning, however, they used a randomised controlled trial to measure the impact without questioning the applicability of a medical research methodology to an educational intervention.

### 2.5.2 Analysing Learning Process

A different point of view and less established way of regarding the analysis of learning in OSNs for medical practitioners is to focus on the essential process of learning, which is the focus of this thesis. It has been suggested that most relevant learning for the development of professional knowledge and expertise is informal, thus, it cannot be readily analysed for the evaluation of outcomes (Kelly & Hager, 2015). Further, learning outcomes are difficult to assess as medical practitioners may acquire new knowledge and skills through OSNs but still be unable to apply them in daily practice because of constraints within which they work. Therefore, investigation of the learning
process has been suggested, since an improved understanding of the activities in OSNs has the potential to assist CPD providers to better serve medical practitioners (Margolis & Parboosingh, 2015). Sandars et al. (2014) suggest that new methods to evaluate learning in OSNs need to recognise how medical practitioners learn (i.e. understanding of learning process) in this context because improving the process of learning may empower medical practitioners to change practice. They suggest applying LA to analyse learning interaction, learning content, and learning context when investigating the learning process in OSNs for medical practitioners (Sandars et al., 2014; Sandars & Walsh, 2014).

Studies that investigate the process of learning in OSNs for medical practitioners have emerged but are limited in the literature. Casebeer, Bennett, Kristofco, Carillo, and Centor (2002) used surveys to identify the online learning behaviour of medical practitioners. They found that medical practitioners demonstrated self-directed learning behaviours, as they learned online quite strategically to address patient problems as and when these came up. They considered that OSNs were used as “Intensional Networks”, which were characterised by emergency, being used for a task at hand, and a sense of history, being based on known relationships and a shared experience of working on similar tasks. Curran and Sibte Raza Abidi (2007) evaluated the impact of an online discussion forum created for 18 months to support information exchange for nurses and physicians. Through the content analysis of discussion topics, they found that the online medium stimulated an opportunity for these emergency practitioners from multiple sites to engage in dialogue around topics that were relevant to their practice learning needs.

Other studies employ the data collected in online systems to investigate the patterns of interaction. For example, Ikioda, Kendall, Brooks, De Liddo, and Buckingham Shum (2013) used SNA to visualise the connections established between UK health visitors in a VCoP, finding that a VCoP is likely to have a mixture of lurkers, observers, passive and active contributors. They considered that the interaction of a health VCoP may be influenced by network size, geographical dispersion and topic relevance. Stewart and Abidi (2013) investigated the sharing of experiential knowledge among medical practitioners within a paediatric pain discussion forum. They believed that the use of SNA helped identify the content experts and active subgroups in the network.
Although scholars have studied the process of informal learning in OSNs, few have systematically investigated the learning process of medical practitioners in an OSN by considering their interactions either with content, or in context. This may be due to the difficulty of gaining access to the required data (Sie & De Laat, 2014), or limitations of current analysis methods (Sandars et al., 2014). The next section reviews the methods developed for analysing the process of learning in OSNs.

2.6 Methods for Analysing Learning in OSNs

2.6.1 Content Analysis

Earlier work on the analysis of online informal learning was dominated by analysis of asynchronous discussion forums, with a focus on identifying effective learning and knowledge building processes using content analysis (Garrison, Anderson, & Archer, 2001; Henri, 1992). Learning in social settings is considered as a set of processes by which the learner constructs meaning, and new ideas based on their previous experiences. This suggests that it is possible to search for evidence of the learners’ online interactions such as conversations, in order to identify cognitive and social processes in which learners engage to give meaning to their new ideas.

Content analysis is a well-established research technique widely used in online education, and it makes use of specifically designed coding schemes to analyse the latent variables of text artefacts with respect to the defined research objectives. However, this requires considerable resources and effort for manual data coding of the learning content in which learners engage (Strijbos, Martens, Prins, & Jochems, 2006). Many scholars (Garrison, Anderson, & Archer, 2000; Gunawardena, Lowe, & Anderson, 1997; Henri, 1992) have proposed coding schemes that may be used to analyse the contents of conversations and thus to characterise the processes of knowledge construction in online learning environments. The following paragraphs review these main coding schemes proposed for analysing learning content.

Henri (1992) proposed five dimensions to be used for evaluating computer-mediated conferences: participative, social, interactive, cognitive and metacognitive. Gunawardena et al. (1997) pointed out that Henri’s model was not specific enough for evaluating the process of knowledge construction, so they presented an analysis
framework that consists of five phases of knowledge construction to study the process of knowledge construction. Both models identify varying numbers of stages or phases with different theoretical perspectives, but they describe knowledge construction in online learning environments as a logically sequenced developmental process.

Based on the work of Gunawardena et al. (1997), Garrison et al. (2000) developed the Community of Inquiry (CoI) model that explains different dimensions of social learning occurring in an online community of education professionals. The model proposed three conceptual elements, also known as presences (including cognitive, social and teaching presence), which together provide a comprehensive understanding of community-based learning. Cognitive presence refers to “the extent to which the participants in any particular configuration of a community of inquiry are able to construct meaning through sustained communication”. Social presence is defined as “the ability of participants to project their personal characteristics into the community, thereby presenting themselves to the other participants as ‘real people’”. Teaching presence is viewed as “a means to an end to support and enhance social and cognitive presence for realising educational outcomes”.

The CoI model has been widely researched in formal learning but received many critiques. It was argued that the value of social presence to online learning experience was questionable due to its observed effect on the cognitive presence (Akyol, Garrison, & Ozden, 2009; Annand, 2011). Shea and Bidjerano (2009) reported that low social presence but high teaching presence still leads to high cognitive presence and vice versa. Annand (2011) pointed out that the present CoI research work was mostly done in the context of online higher education, thus the understanding of the application and the effect of social and cognitive presence in the context of social learning is limited.

Despite critiques, content analysis and the proposed coding schemes have been applied to investigate the learning content in OSNs. For example, De Laat (2002) applied the Gunawardena et al. (1997) analysis framework to assess the quality of the dialogue in terms of social knowledge construction in an online community for the police. The author identified that the knowledge construction occurs mainly in the information sharing phase. Schrire (2006) applied the coding scheme of CoI model to investigate the collaborative knowledge-building process in a discussion forum used in a higher
education context. The author found a high proportion of higher-order thinking based on all three models tested. Both authors acknowledged the limitation of content analysis and the coding schemes in terms of the large effort required for manual data coding. However, classification techniques for automating different content analysis coding schemes have been investigated subsequently (Kovanović et al., 2016).

In the next section, we will introduce different approaches to LA for analysing learning in OSNs.

2.6.2 Learning Analytics

LA has been proposed as an innovative approach to understand the process of learning in OSNs for medical practitioners (Sandars et al., 2014). LA is an overarching analytical approach for aggregating potentially very large volumes of online learning data in order to analyse learners’ performance and interactions within online environments. The Society for Learning Analytics Research (SoLAR, 2013) defines LA as “the measurement, collection, analysis and reporting of data about learners and their contexts, for the purpose of understanding and optimising learning and the environments in which it occurs”.

LA is an interdisciplinary field. As such, LA selects and uses analysis techniques from a variety of disciplines as appropriate to achieve the goal of improving learning. According to Dawson, Gašević, Siemens, and Joksimovic (2014), LA uses various approaches including visual data analysis techniques, SNA, statistics, and educational data mining to analyse the data.

LA was initially used in formal educational settings. Its main application includes tracking and predicting learners’ performance, as well as identifying potential problematic issues and students at risk (Nunn et al., 2016). For example, the Signals project at Purdue University helped the institution to predict students who were at risk of performing badly based on data collected from the course management system (Arnold & Pistilli, 2012). There is also some research on using student interaction data to establish indicators of more complex concepts such as creativity (Dawson, 2010), engagement (Olmos & Corrin, 2012; Richards, 2011), and self-regulated learning (Biswas, Jeong, Kinnebrew, Sulcer, & Roscoe, 2010).
Viewing medical professional CPD from a social learning perspective highlights types of LA that can be employed to understand learner activity in OSNs. Emerging analytic approaches for investigating social learning include: social network analytics, content analytics, discourse analytics, and context analytics (Ferguson & Shum, 2012).

Social network analytics involves using SNA techniques to investigate the network processes and properties of ties, relations, roles and network formations, and to understand how people develop connections to support their learning within networks (Haythornthwaite & De Laat, 2010). For example, SNA centrality measures can be used to study the patterns of interactions occurring in an online community among teachers (Schreurs & De Laat, 2012). SNA visualisation techniques can be also used by teaching staff to identify patterns of learner interaction in a discussion forum (Dawson, Bakharia, & Heathcote, 2010).

Content analytics is a broad heading for the variety of automated methods that can be used to examine online learning content. It focuses on building analytical models that rely on computational tools to process content, including learner-produced content such as online discussion messages (Kovanović et al., 2015). Some popular techniques include: topic modeling for identifying themes and topics in learning content (Tobarra, Robles-Gómez, Ros, Hernández, & Caminero, 2012); classification techniques for automating different content analysis coding schemes (Kovanović et al., 2016); and collaborative filtering techniques for discovering and recommending educational resources (Walker, Recker, Lawless, & Wiley, 2004).

Discourse analytics is the collective term for various approaches to identify where and how learning happens, while the quality of contribution to discourse is viewed as a key indicator of meaningful learning. It includes processing of open response questions in educational contexts, and a large proportion of research in the area focuses on assessment of writing. For example, natural language processing techniques is used to examine linguistic and rhetorical properties of written text, so as to enable automatic assessment of the quality of an argument in an essay (Shermis, 2014; Simsek, Shum, De Liddo, Ferguson, & Sándor, 2014). Discourse analytics also includes the analysis of a series of communication events in OSNs. For example, it has been applied to analyse asynchronous discussions between students and tutors, showing how groups of learners
create and maintain community and coherence through the use of discursive devices (Lapadat, 2007).

Context analytics involves automatically collecting and using the available context data such as learner profile information, timestamps and location to support recommendation (Beale & Lonsdale, 2004). A useful approach is tracking of learner context via mobile technologies, such as geospatial location apps (Schäper & Thalmann, 2014). Another approach is using technology standards. For example, MedBiquitous focuses on developing a common technology framework that will support reforms in healthcare education and competence assessment, proposing a ‘Professional Profile’ data standard in XML format that provides a common format for exchanging clinician contact, education, training, certification, and membership information (Smothers, Clarke, & Van Dyck, 2006). This could potentially help the exchange of learner context information, if the learner is given the ability to update their profile and import and manage their profile.

LA for workplace and professional learning is now on the research agenda of several ongoing European Union projects (Ley, Klamma, Lindstaedt, & Wild, 2016). Suggestions have been made to combine different types of analytics to further enrich the understanding of learning processes (Buckingham Shum & Ferguson, 2012). One approach suggested is to combine SNA with content analysis and context analysis to gain a richer picture of networked learning, investigating not only who is talking to whom, but what they are talking about and why they are talking in this way (De Laat & Schreurs, 2013).

2.7 Summary

Conventional CPD has shortcomings in the way that it can meet individual learning needs of medical practitioners and improve their professional practice. Specifically, these shortcomings include: 1) no ability to access CPD over large distances; 2) limited opportunity to scale up to provide CPD for large numbers of people; and 3) limited ability to provide CPD at the time when it is most impactful. Based on social learning theories, OSNs may contribute to improving these aspects of CPD, but their exploitation
for this purpose depends on developing methods that support understanding how learning occurs in such environments.

In order to understand and identify ways of improving learning in OSNs for medical practitioners, previous researchers have taken two directions: one is to measure the outcome of learning, the other is to study the process of learning. Since most of the learning in these OSNs is informal, the electronic data that is generated by the learners cannot be used readily for the direct evaluation of outcomes. Therefore, the study of the learning process is more feasible and meaningful as a way to gain insights about learning in these OSNs.

Formal learning in OSNs is extensively studied, but OSNs for just-in-time informal learning of medical practitioners are new and rarely studied. This is partly due to the limited availability and accessibility of data, and partly due to the limitations of current analysis methods.

The emerging field of LA is recognised as an innovative approach to working with such data, but no previous studies have systematically investigated the learning process in OSNs for medical practitioners through the use of LA.

The three chapters that follow present original research in an OSN for medical practitioners in the real world. They explore uses of three LA methods to analyse electronic data from this OSN and they demonstrate how these methods can provide three views of the learning process: interactions among learners; interaction with content; and the context of interaction.
Chapter 3 Using Social Network Analysis for Analysing Learning Interaction in an Online Social Network for Medical Practitioners

The relevant content of the following publications has been integrated into this chapter:


3.1 Introduction

This research reported in this thesis aims to understand how Learning Analytics (LA) can be used as a methodology for understanding the process of informal learning in OSNs for medical practitioners. As outlined in Chapter 1, an online discussion forum for medical practitioners was analysed based on a multi-method research design, in order to get a detailed understanding of the learning process through LA methods.

The first step in such analysis was to describe the interaction within this learning environment. This in turn would help understand learning behaviours, which effectively signify a learning process, and can be important indicators that learning may have occurred.

Social Network Analysis (SNA) has already proven to be an effective technique to analyse interaction in online learning environments (Ferguson & Shum, 2012; Sie, Ullmann, et al., 2012), as outlined in Chapter 2. Using SNA, the nodes (learners, or learning resources) and ties (relationships) in a network can be visualised and analysed using quantitative measures and graphical representations to examine the flow of
interactions. Network modeling techniques can be used to further understand changes in patterns of interaction over time and to test associated structural network effects that could explain these changes as features of the learning process.

This chapter introduces SNA in more depth and demonstrates how SNA, as a method of LA, can be used to interpret the patterns of interactions among medical practitioners in the online discussion forum.

### 3.2 Background and Related Work

SNA first appeared in 1930s, and has been applied in many research disciplines including physics, life sciences, computer science, statistics and social sciences to study different types of relations and problems (Prell, 2012). Increasingly, it has been used to study learning environments and collaborative learning problems (Haythornthwaite & De Laat, 2010). Basically, SNA is used to study relationships established among different individuals within social networks. The individuals are represented as nodes of the network, while their relationships are represented as links. Networks are created when nodes are connected by links. Learning may be considered to be occurring when learners make links to other learners (Carolan, 2013).

A social network can be classified as a one-mode network or a two-mode network (Borgatti, Everett, & Johnson, 2013). In most network analysis, nodes represent a class of people and links represent some sort of social construct (e.g. learning relations) that connects them. This is referred to as a one-mode network, as the connections are between a single class of nodes. Two-mode networks, in contrast, represent two different classes of nodes, and links exist only between classes. There are two main types of node sets – one is people and the second is the defined events – and the link goes from one node to another indicating a person has attended an event. This type of relation is considered transitive which means it can be converted into a one-mode network. This simplifies the analysis process of the social network and allows applying the normal SNA measures (Sie, Ullmann, et al., 2012).

In networked learning research, SNA is used in four ways: analysis, visualization, simulation and intervention (Sie, Ullmann, et al., 2012). We describe each of these approaches in the following paragraphs.
Analysis involves the mathematical analysis of interactions, relationships and positions in the network, and the network itself. The mathematical quantitative analysis of a social network is accomplished on three parts of the network (Borgatti & Halgin, 2012). These levels are the whole network, the subgroup, and individual level. Choosing the level of analysis depends on the focus of study and the types of relationships that need to be understood. There are measures for each level that facilitate understanding the overall structures and relationships of social networks. Individual-level measures include centrality measures, such as degree centrality, betweenness centrality, and closeness centrality. They can point out the different roles of the nodes like leaders, isolates or bridges. Subgroup measures include measures that detect cliques, cohesive subgroups, and components. Network-level measures capture the overall centralization and density of the network.

Visualization focuses on generating images of the networks of learners (nodes) to show which learners are interconnected through their learning relationships (links). It is considered an essential supplement to the mathematical analysis of the network. The main objectives for visualization include discovering: 1) centralized actors with higher influences over others; 2) isolated individuals with minimal number of connections; 3) brokers who play an important role in keeping the network connected; and 4) subgroups and clusters that are formed based on shared interest (Freeman, 2004). Visualization of networks is done by sociograms. Sociograms are visual representation of individuals and their relationships to others in a group (Prell, 2012). Figure 3.1 presents an example of a sociogram. As shown, individuals are represented as points, and the relationships linking the individuals together are represented as lines.

Figure 3.1. An example of a sociogram
Simulation is about using network models (e.g. stochastic models) to model network behaviours and explain the structure of networks. Many statistical models have been developed for the analysis of social network data. The best known of these statistical models are Exponential Random Graph Models (ERGMs) (Robins & Lusher, 2012) and Stochastic Actor-Oriented Models (SAOMs) (Snijders, 1996). ERGMs is the preferred model to use for making statistical inferences about cross-sectional network data, whereas SAOMs was developed for analysing longitudinal network data and is used to analyse the changes in network structure over time.

One of the key reasons of using these statistical models for network data over other forms of network analysis is to determine whether lower-level network configurations are commonly observed in a given network. In other words, the statistical models use network configurations to explain the structure of a network. A network configuration refers to a small set of nodes with a subset of ties amongst the nodes (Robins, Snijders, Wang, Handcock, & Pattison, 2007). Examples (in Figure 3.2) include an edge, where two nodes are connected by one mutual tie; a 2-star, which consists of three nodes where one node is connected by a tie to each of the other two; and a triangle, where three nodes are connected through three mutual ties.

![Network Configurations](image)

**Figure 3.2.** Examples of network configurations

Intervention may involve setting up activities through analytic tools, data, and reports. For example, interventions can be undertaken by using SNA to promote network awareness and to uncover hidden connections (Fetter, Rajagopal, Berlanga, & Sloep, 2011), and to strength certain types of relationship between learners (Sie, Drachsler, Bitter-Rijpkema, & Sloep, 2012).

The application of SNA to online learning is still at an early stage (Cela, Sicilia, & Sánchez, 2015; Sie, Ullmann, et al., 2012). Visualization and analysis are the major approaches in current SNA research for online learning (Cela et al., 2015). The majority
of studies neither employ simulation techniques to explain network behaviours, nor use SNA to intervene in real-world settings (Sie & De Laat, 2014). Further, there have been very few studies conducted in the context of professional learning in any specific field.

Basic SNA measures such as centrality have been used to study the patterns of interactions occurring among teachers in an online community (Schreurs & De Laat, 2012), and to understand the flow of experiential knowledge sharing among medical practitioners within a paediatric pain discussion forum (Stewart & Abidi, 2013). These studies conducted cross-sectional analysis, which reflected only a temporary state of the network. However, learning is an on-going process, so longitudinal network analysis has been suggested to overcome this challenge by studying the interaction changes over time and the driving factors behind such changes (Sie & De Laat, 2014).

Only a few studies have employed simulation techniques to explain changes of network structure and learning behaviours (Sie & De Laat, 2014). Uddin, Thompson, Schwendimann, and Piraveenan (2014) employed ERGMs to explain the formation of the network in different stages of a university course by studying the dynamics of longitudinal email communication among students. They found that the students used increased numbers of email communications as their study load increased. Wild and Sigurdarson (2011) developed a simulation model to simulate the impact of facilitation on a blog-based learning network in a university; they found that improved facilitation could have a positive impact on the density and connectedness of the network.

### 3.3 Methods

SNA is employed to identify the patterns of interaction occurring in the case study discussion forum. The analysis of interaction is done in two phases: Phase 1 is a cross-sectional study that aims to understand the pattern of overall interaction; Phase 2 is a longitudinal study that aims to understand the changes in patterns of interaction. The following sections describe the approach and methods employed in these two analysis studies.

To perform SNA, network(s) must be constructed using the data collected from the discussion forum. The general approach to the network construction using forum data is applied in both phases of the analysis studies. The online discussion forum can be
considered as a two-mode network, which represents how actors are tied to events (i.e. an actor-by-event network). As mentioned in the above section, a common method of analysing a two-mode network is to transform the data into two one-mode networks (Scott & Carrington, 2011). One is an actor network, in this case, created from the forum users. A tie is created between them if they communicated on the same thread. Another is an event network, created from discussion threads. A tie is created between two threads if the same user communicated on both. Figure 3.3 presents an example of a two-mode network that evolves from the participation of two Threads (T) by five Users (U), and a one-mode actor network that is transformed from these five users who participated on the same thread.

![Two-mode network](image1.png)

**Figure 3.3.** An example of how a network is built from a discussion forum thread

### 3.3.1 Phase 1 Study

#### 3.3.1.1 Data

Data for this study were selected for all users (n=621) who participated in discussion in the forum. This includes all the threads (n=723) that were involved between the period 2009 and 2014.

#### 3.3.1.2 Measures

Table 3.1 depicts the SNA mathematical measures employed in this study. This involves using four network-level structural measures: density, centralization, diameter and average path length, and three individual-level centrality measures: degree, betweenness, and closeness centrality (Wasserman & Faust, 1994). The network-level measures reveal the interaction level and connectivity in the entire network. The individual-level measures provide information about the activity levels of the individual
users, along with the overall activity status of the network; they help understand how interactions take place by summarising the individual users and thread-level characteristics. Lastly, visual analysis was done by representing the relationships between users/threads through graphs, to enrich the findings of mathematical analysis.

Table 3.1. The definitions of SNA measures

<table>
<thead>
<tr>
<th>SNA measure</th>
<th>Mathematical definition</th>
<th>Descriptive definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Network-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Density</td>
<td>[ d = \frac{L}{n(n-1)/2} ] (where ( L ) refers to actual number of lines present in the network, and ( n ) to the number of nodes present in the network)</td>
<td>The number of present relationships as a ratio of the possible number of relationships in a network; represents the overall connection between users/threads.</td>
</tr>
<tr>
<td>Centralization</td>
<td>[ C = \frac{\sum(Max(C_{Di}) - C_{Di})}{n^2 - 3n + 2} ] (where ( Max(C_{Di}) ) is the network’s maximum centrality score, ( C_{Di} ) indicates individual actors’ centrality scores, and ( n ) is the network’s size)</td>
<td>The extent to which the connectedness is focused around a user/thread.</td>
</tr>
<tr>
<td>Diameter</td>
<td>[ d = \text{Max}_{ij} l(i, j) ] (where ( l(i, j) ) is the length of the shortest path between node i and j)</td>
<td>The longest path between any pair of users/threads in a network.</td>
</tr>
<tr>
<td>Average path length</td>
<td>[ a = \frac{\sum l(i, j)}{\frac{n(n-1)}{2}} ] (where ( l(i, j) ) is the length of the shortest path between node i and j)</td>
<td>The average path between any pair of users/threads in a network.</td>
</tr>
<tr>
<td>Individual-level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree centrality</td>
<td>[ d(l) = \sum_{j} a_{ij} ] (where ( a_{ij} ) is the total number of ties from node i to node j)</td>
<td>The number of direct relationships a user/thread has with others in a network, which provides an indication of their popularity and influence.</td>
</tr>
</tbody>
</table>
**SNA measure** | **Mathematical definition** | **Descriptive definition**
---|---|---
Belowenness centrality | \[ b(i) = \sum_{ij} \frac{g_{ikj}}{g_{ij}} \] (where \( g_{ij} \) is the number of shortest paths from node \( i \) to node \( j \), and \( g_{ikj} \) is the number of shortest paths from \( i \) to \( j \) that pass through \( k \)) | The number of times a user/thread sits on the shortest path linking two other users/threads together in a network; it helps identify important users/threads.

Closeness centrality | \[ c(i) = \sum_j d_{ij} \] (where \( d_{ij} \) is the number of links in the shortest path from actor \( i \) to actor \( j \)) | How quickly a user/thread can reach all other users/threads within the entire network; it provides an indication of the speed of information distribution.

### 3.3.1.3 Procedure

To perform SNA, a two-mode network was constructed first based on the raw data extracted from the forum database. Then, the two-mode network data file (a csv file) was loaded into RStudio for analysis.

To illustrate how the raw data were extracted from the forum database and transformed to a two-mode network data file, the original structure of sample posts and threads in the forum is presented in Figure 3.4. A thread is contained in the forum and can contain any number of posts. The column named ‘parent_id’ represents the unique identity of a thread. The column named ‘id’ represents the unique identity of a post. If the value of ‘parent_id’ is NULL, it indicates the record is the first post of a thread, and the value of ‘parent_id’ equals to the value of ‘id’. For example, the ‘parent_id’ in the record 1 is NULL, indicating that this record is the first post of the thread (id=4473); the records 2 to 6 are identified as associated posts to the thread (id=4473) as the value of their parent_id equals to ‘4473’.

---

5 RStudio is a free and open-source integrated development environment (IDE) for R, a programming language for statistical computing and graphics. https://www.rstudio.com
Figure 3.4. The original structure of posts and threads

Basically, there are two steps involved in constructing the two-mode network for the Phase 1 study: 1) extracting raw data from the database table `se_forum_posts` (including the unique ids of users and the threads attended) into an Excel Spreadsheet Editor through Structured Query Language (SQL); 2) converting the data to matrix format using Excel Spreadsheet Editor. Figure 3.5 provides a sample output of the two-mode data matrix based on the sample posts provided above. The two-mode network data are generated and saved in csv format and are ready to be imported into RStudio for performing SNA.

Figure 3.5. A sample output of two-mode data matrix

The STATNET library (version 2014.2.0) in R was loaded into RStudio to perform mathematical analysis. Firstly, the two-mode network was transformed into two one-mode networks (one user network and one thread network). A sample output of one-mode data matrix based on the above example is shown in Figure 3.6. Then, the
individual-level and network-level measures for both user and thread networks were calculated.

Figure 3.6. A sample output of one-mode data matrix

The IGRAPH library (version 0.7.1) in R was used for visual analysis. The two-mode network was visualised based on the layout option `fruchterman.reingold` since it directs most connected nodes into the centre of the graph. The user network and thread network were visualised separately. Given the large size of the network (users=621, threads=723), both networks were thinned by displaying only those ties that satisfy a cut-off point (specifically, only those ties that have a tie weight greater than the mean tie weight plus one standard deviation) were kept.

3.3.2 Phase 2 Study

3.3.2.1 Data

This phase of the study focused on the interactions of the forum participants who were more engaged in discussion than others (i.e. the active users). Based on the preliminary analysis results we reported in Section 1.4.2, the period 2012 to 2014 has the most recent and complete years available at the time of data collection, so an active user in
this study was defined as a user who contributed to any discussion in three consecutive years (2012, 2013, 2014) in the forum.

Discussion forum data were selected and grouped into three datasets of one-year duration. Within these datasets, data were extracted on the activities of the 48 active users.

3.3.2.2 Measures

The same SNA network-level measures as defined in the Phase 1 study (i.e. density, centralization, diameter and average path length) were applied to reveal the interaction level and connectivity among participants. The patterns of interactions were also visualised to enrich the findings about the network measures.

To identify the structural properties of interactions and test the significance of their effects in the network, SAOMs (Snijders, 1996) was employed. This approach considers the totality of all possible network configurations of a given set of actors (the medical practitioners, in this case) as the state space of a stochastic process, and models the observed network dynamics by specifying parametric models for the transition probabilities between these states.

In this study, SAOMs was used to examine which micro structures might play a statistically significant role in the process of learning. Specifically, the network effect known as transitivity was considered. Transitivity refers to the situation where actor \( i \) is tied to actor \( j \); actor \( j \) is tied to actor \( k \); and actor \( i \) is likewise tied to actor \( k \) (Wasserman & Faust, 1994). It is a key concept that helps explain the structure of a network, which is relevant to the network configuration of a triangle. Establishing transitivity helps form exclusive learning groups over time so that learners can build learning relationships, construct knowledge and learn from each other (Carolan, 2013). Therefore, it may be important when designing a social learning network for medical practitioners.

In terms of attribute effects, homophily was tested. Homophily refers to the extent to which individuals interacted with others who were similar to themselves (Prell, 2012). This similarity could mean either those from the same organisational settings (e.g.
hospitals, professional associations), or those with similar personal attributes (e.g. age, sex, educational background). In this study, the homophily effect was used to determine whether the participants’ characteristics (included gender and geographic location details that were available at the time of data collection) might affect changes in the patterns and structure of interaction.

### 3.3.2.3 Procedure

The same procedure as explained in Phase 1 study was followed to construct two-mode networks but used different SQL to query the 48 active users and the threads where they were active between 2012 and 2014 from the table `se_forum_posts` for longitudinal analysis. Three two-mode networks were formed by extracting the data (user_id, thread_id) separately for each year, so three two-mode network data files (two-mode network_2012.csv, two-mode network_2013.csv, two-mode network_2014.csv) were generated and imported into RStudio for performing SNA. Also, the gender and geographical location (country and state) attributes for each user were extracted from the table `se_profile_info` for network modeling.

The STATNET library (version 2014.2.0) in R was loaded into RStudio to perform mathematical analysis. All three two-mode networks were transformed into one-mode networks, and so three one-mode user networks (one for each year) were generated. Then network-level measures for the three user networks were calculated to observe the changes that occurred during these years.

The IGRAPH library (version 0.7.1) in R was used for visual analysis. The three one-mode user networks were visualised and compared. The graph visualization is based on the layout option `fruchterman.reingold` since it directs most connected nodes into the centre of the graph. To reveal the patterns of interaction, all networks were thinned by displaying only those ties that passed a minimum threshold (specifically, only those ties that had a weight greater than the mean weight plus one standard deviation were kept).

Statistical modeling was performed using the SIENA library (version 1.1-232) in R called RSiena. To model the network, four key steps were followed below, as guided by the current RSiena users’ manual (Ripley, Snijders, & Preciado, 2011). Appendix B presents the major R code used for this analysis.
1. Defining the data by loading all three one-mode user interaction network csv files as dependent variables, and user characteristics (gender and geographic location attributes) as explanatory (independent) variables.

2. Combining dependent and independent variables to define dataset and obtain the basic effects object.

3. Defining these effects into the model and performing estimation of the parameters by fitting the specified model to the dataset until reaching convergence (via ‘convergence t-ratio’)

4. Testing the significance of valid effects using t-value, which is defined by dividing the parameter estimate by its standard error.

3.4 Phase 1 Study Results and Discussion

3.4.1 The Overall Level of Interaction

SNA network-level measures are used to identify the overall level of interaction in the network. Table 3.2 presents the results of network-level measures for both user and thread networks. Regarding the user network, the density score of 0.04 indicates a low level of participation and connection among the users in the network. The centralization score of 0.59 indicates that the user network is centralized, which implies that the interaction is centralized in a small set of users and a considerable number of users are not engaged and interact very little. This finding is consistent with other research (Beck, Fitzgerald, & Pauksztat, 2003; Nonnecke, Andrews, & Preece, 2006) that found evidence of a small set of users producing the bulk of the discussion within online communities. A diameter of 5.00 and average path length of 2.17 indicates that the users are not very close to one another, which confirms the low density of the network, meaning that the users may not easily reach each other or share knowledge.

The thread network has a much higher density of 0.40 compared to the user network, which may imply that the same users initiated many threads, and that those threads were connected through them. The centralization score of 0.46 indicates that most threads were equally active and only a few attracted more attention than others. The diameter of 4.00 and average path length of 1.66 indicate that the threads sit close to each other in
general, however some are distant from each other. This implies that some threads were initiated and commented on by different sets of users and so there is a chance that small group learning was occurring. After inspecting these threads to cross-check, it was found webinar case discussions had been conducted in a few specific sub-forums.

**Table 3.2.** Network-level measures

<table>
<thead>
<tr>
<th>Network level measure</th>
<th>User (N=621)</th>
<th>Thread (N=723)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0.04</td>
<td>0.40</td>
</tr>
<tr>
<td>Centralization</td>
<td>0.59</td>
<td>0.46</td>
</tr>
<tr>
<td>Diameter</td>
<td>5.00</td>
<td>4.00</td>
</tr>
<tr>
<td>Average path length</td>
<td>2.17</td>
<td>1.66</td>
</tr>
</tbody>
</table>

**3.4.2 The Level of Interaction for Individual Users**

SNA individual-level measures are used to identify the level of interaction for the individual users in the network. Figure 3.7 presents the degree of centrality distributions for the users. As shown, degree distribution is quite skewed, ranging from 0.001 to 0.42 with a median of 0.003. It is a highly centralized network, in which a minority of users who have a degree of 0.20 or above are producing the bulk of the discussion.

![Figure 3.7. Degree centrality distributions for user network](image-url)
Cross-checking their identities in the forum database showed that these users were mainly moderators and core members. Moderators were people who had been given explicit responsibility by the forum operator to initiate and facilitate discussions, while ‘core member’ is the researcher’s own label for the key users who initiated and participated most in discussions.

Figure 3.8 presents the closeness centrality distributions for the users. As shown, closeness centrality ranges from 0.26 to 0.59 with a median of 0.37. Most of the users are not very close to each other, confirming that the network is loose. However, most users are similarly positioned, which implies that the users have potential access to learn from each other.

![Closeness centrality distributions for the user network](image)

**Figure 3.8.** Closeness centrality distributions for the user network
Figure 3.9 presents the betweenness centrality distributions for the users. As shown, betweenness centrality has a highly-skewed distribution. There are many users (N=290) who have 0.00 betweenness. Only a few users (N=14) have betweenness centrality of 0.02 or above. This indicates that these users are important participants who acted as leaders, knowledge brokers or experts; so identifying them might help connect disconnected learners and provide new learning opportunities. Median betweenness centrality is 4.83886e-06 which is very low, showing that most of the users are part of a cluster; the cluster forms the core of the network and the users are not well connected each other. This low connectivity further emphasises the relatively low levels of participation in the forum.

![Figure 3.9. Betweenness centrality distributions for the user network](image-url)
Figure 3.10 depicts the degree centrality distributions for the threads. Degree distribution ranges from 0.002 to 0.05 with a median of 0.006, indicating there is no single thread that everyone commented on. The thread network is not centralized compared to the user network. There are only a few threads (N=25) with relatively higher degrees of 0.024 or above, as most threads had fewer than 13 comments.

**Figure 3.10.** Degree centrality distributions for the thread network
Figure 3.11 depicts the closeness centrality distributions for the threads. As shown, closeness centrality ranges from 0.43 to 0.59 with a median of 0.45, indicating that most of the threads sit close to each other, and are similarly positioned. This suggests that users who participated in one learning topic did not find it difficult to participate in other learning topics.

**Figure 3.11.** Closeness centrality distributions for the thread network
Figure 3.12 depicts the betweenness centrality distributions for the threads. As shown, betweenness centrality ranges from 0.00 to 0.02 with a median of 0.001. Though many threads (N=71) have 0.00 betweenness, this number is much lower than in the user network. Approximately 25% of the threads have betweenness centrality of 0.003 or above, showing that the threads are independent, and there is no one thread required to initiate other threads, or to keep other threads going.

Figure 3.12. Betweenness centrality distributions for the thread network
3.4.3 The Patterns of Overall Interaction

Figure 3.13 presents the visualization of the two-mode network, which has nodes of user and thread. As shown, there are only a small number of users sitting in the centre of threads who were really engaged and actively participating in threads: about seven active users controlled most threads where comments were posted.

Figure 3.13. Visualization of two-mode network
Figure 3.14 presents the visualization of two one-mode networks (i.e. the user and thread networks). As shown, the user network (node size represents the number of threads where the user contributed) is quite centralized. The conversation mostly occurs among a few active users and the rest are not engaged much. The thread network is generally decentralized. Some threads received more attention than others, but were quite independent of each other. These findings are confirmed by SNA mathematical analysis results.
3.5 Phase 2 Study Results and Discussion

3.5.1 Changes in the Level of Interaction

SNA network-level measures are used to calculate each of the three user networks. Table 3.3 presents the results of the network-level measures for these networks. The overall low density scores (<0.5) indicate a low level of participation and connection among these participants. The decline in the density score indicates a decreasing level of participation over time. The low centralization score of N1 and N3 (<0.5) indicates the network was not centralized continuously over the three years of networked learning that we examined. The centralization of N2 (0.68) demonstrates that the interaction in the middle year was centralized, while most were not engaged and interacted infrequently. It is unknown from this result why centralization was happening only in the middle year. The overall high diameter result (>2) shows that in general participants
were not very close to each other in terms of interaction steps, but the low average path length confirms that most of them were as close to each other as one or two interaction steps. This therefore indicates that there were only a few participants who were distant from the core group and thus might not easily share knowledge.

Table 3.3. Network-level measures of N1, N2, and N3

<table>
<thead>
<tr>
<th>Network measure</th>
<th>N1</th>
<th>N2</th>
<th>N3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>0.31</td>
<td>0.27</td>
<td>0.17</td>
</tr>
<tr>
<td>Centralization</td>
<td>0.43</td>
<td>0.68</td>
<td>0.49</td>
</tr>
<tr>
<td>Diameter</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Average path length</td>
<td>1.77</td>
<td>1.78</td>
<td>2.13</td>
</tr>
</tbody>
</table>

3.5.2 Changes in the Patterns of Interaction

Figure 3.15 presents the visualization of the three user networks. In N1 (2012), approximately 40% of participants (N=18) sat in the centre of the network; they were engaged and actively participating. Analysis of OSN member identification data showed that this included two moderators (orange dot) who had a formal facilitation role in the discussion forum. Unsurprisingly these two contributed the most to the discussion threads.

Figure 3.15. Visualization of changes in the user networks

In N2 (2013), less participation overall was seen as compared to N1. There was one member (yellow dot) informally leading the network – this core member interacted directly and frequently with a number of other users, but little conversation occurred
directly between those other users. The moderators moved more towards the side and one moderator became a “broker” who helped to bridge the discussion among users as indicated by one of the orange dots.

In N3 (2014), more participants moved to the edge of the network. More conversation started happening among the participants in the centre, led by the same core member who led N2.

By examining the user identity numbers in the three networks, it was found that most active participants (other than the moderators and the core member) did not remain in the centre over time. In other words, they generally did not stay engaged for a long period of time.

Furthermore, a visualization of interactions provides some insights on why the network was centralized only in N2 – because when the core member took over the effective lead role from the moderators, the discussion occurred between the core member and other participants. As shown, the participants only started interacting with each other again, independently of the core member, in the following year. This reflects a tendency for the network to become temporarily centralized if there is a change in the lead role.

### 3.5.3 The Structural Properties of Interaction

Network modeling (via SAOMs) was done to identify the structural properties of interaction in the network to explain the changes in patterns of interaction. The identified network effects on structural parameters and their results are summarised in Table 3.4. The significance of each effect can be tested using a t-value, which is defined by dividing the parameter estimate by its standard error. It measures how many standard errors the estimate is away from zero. Generally, any t-value greater than +2 or less than -2 is acceptable.

#### Table 3.4. Estimation results for structural parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>T-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate period 1</td>
<td>27.16</td>
<td>5.93</td>
<td>4.58</td>
</tr>
<tr>
<td>Rate period 2</td>
<td>19.03</td>
<td>2.58</td>
<td>7.38</td>
</tr>
<tr>
<td>Parameter</td>
<td>Estimate</td>
<td>Standard error</td>
<td>T-value</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>----------</td>
<td>----------------</td>
<td>---------</td>
</tr>
<tr>
<td>Degree</td>
<td>-2.67</td>
<td>0.11</td>
<td>-23.85</td>
</tr>
<tr>
<td>Transitive triads</td>
<td>0.25</td>
<td>0.14</td>
<td>1.84</td>
</tr>
<tr>
<td>Distance-2</td>
<td>0.13</td>
<td>0.03</td>
<td>4.43</td>
</tr>
<tr>
<td>Same gender</td>
<td>0.03</td>
<td>0.06</td>
<td>0.54</td>
</tr>
<tr>
<td>Same geographic location</td>
<td>0.11</td>
<td>0.07</td>
<td>1.48</td>
</tr>
</tbody>
</table>

The rate parameter indicates a frequency at which network changes are estimated to occur. The changes of network structure are affected when a participant starts or stops interacting with other participants. Since we have two consecutive observations within the three-year period, Rate period 1 and Rate period 2 indicate the first and second observation, respectively. As shown, the considerable greater value of Rate period 1 (27.16) in relation to the value of Rate period 2 (19.03) suggests the interaction dynamics slowed down from the first to the second observation year, indicating a stabilised interaction in the network. The Degree parameter estimates the change in the number of connections for a participant between two periods, so its negative effect in this case confirms the stabilising tendency within the network.

Both Transitive triads and Distance-2 parameters are indicators of transitivity. Transitive triads estimate the number of transitive patterns (i.e. if participant A interacts with participant B and participant B interacts with participant C, then participant A interacts with participant C). The Distance-2 parameter estimates the number of shortest path lengths equal to 2, and it expresses transitivity inversely.

As shown, the estimation results of the Transitive triads effect is not statistically significant so we do not consider it further. However, the estimation of Distance-2 parameter indicates a positive (0.13) and statistically significant value, so we conclude that network did not reveal any transitivity effect. This therefore suggests that there were numerous null connections and little groups formed in the network, indicating a relatively sparse network in which information (or learning resources) would have difficulty flowing from one part of the network to another. This may suggest that there was little knowledge shared among groups – the medical practitioners did not form
collaborative groups. Instead, they interacted with one or more selected individuals to obtain information in the network. This relates to the earlier work in understanding how medical professionals behave online (Casebeer et al., 2002; Singh, McPherson, & Sandars, 2012), which found medical practitioners were highly strategic online in seeking information to solve problems and build careers, and used OSNs as “Intensional Networks”, which were characterised by emergency, being used for a task at hand, and a sense of history, being based on known relationships and a shared experience of working on similar tasks. This finding may explain why, the interaction in the network is low and appears to fade quickly.

Both the parameters Same gender and Same geographical location are homophily effects considered in this network, to examine whether they indicate any preference for medical practitioners of the same gender or geographic location to interact. Although the concept of homophily associates certain network structures with the similar actor attributes within a network (Snijders, 1996), the estimation results for both parameters of homophily in this study indicate a positive but not statistically significant homophily effect, which means there is no pattern of interaction common to either characteristic.

3.6 Conclusions and Implications

This study employed SNA techniques to analyse the interactions occurring among the medical practitioners in an online discussion forum. The analysis included using SNA individual-level measures (degree, betweenness, and closeness centrality), network-level measures (density, centralization, diameter and average path length), visual analysis, and stochastic actor-oriented models.

We started by analysing the pattern of overall interaction using SNA individual-level and network-level measures and visual analysis. We found that the interaction in the network was low in general and the network was loosely connected, which may imply that knowledge was not easy to share and collaboration for learning was difficult. The network is highly centralized, indicating a small set of active users (includes moderators) producing the bulk of discussion. In turn, this implies not many users were engaged in learning in the network.
To further understand how the users interacted, we shifted the attention of our analysis to the active users. We conducted longitudinal analysis by focusing on the changes in patterns of interaction among the active users. We employed SNA network-level measures and visual analysis to observe the changes in the level of interaction and connectivity; and we used network models to identify the structural properties of interaction that may help explain the changes in interaction behaviours.

The longitudinal analysis shows that the level of interaction among those active users decreased but that their interactions stabilised over time. They might easily access and learn from each other as only a few were distant from the core group. The two moderators (also counted as active users) stayed in the centre of the network (i.e. initiated and contributed to most discussions) over time. By testing structural network effects, we found that these active users were reluctant to share knowledge and collaborate in groups. Possibly they were joining the forum simply for the purpose of seeking specific information, and then leaving. This provides clues as to why the level of interaction in the network is low and appears to fade quickly. Based on available user attributes, we found that there was no pattern of interaction common to their gender or geographic locations, which indicates there was no preference for medical practitioners of the same gender or geographic location to interact.

The level of interaction may be an indicator of the level of learning that occurred in this specific OSN. Only a small percentage of users appear to have engaged substantially in ways of learning that are theorized in social constructivism, connectivism and networked learning. The use of SNA allows us to understand the role of these few users including as leaders and as knowledge brokers. Since they are highly influential and control what knowledge and information are shared, engaging them is essential to furthering the interaction or learning process in the network. Importantly, their information seeking behaviours prompts us to reconsider the purpose of designing OSNs for medical practitioners’ informal learning. Perhaps, instead of encouraging them to form and sustain collaborative learning groups, we should support medical practitioners to assess connections rapidly and target interactions precisely, to help them to exploit learning opportunities whose meaning, and value are more readily recognised and rewarded in their workplaces.
One limitation of this study is its inability to track passive users, due to the limitations of the case study data source. Understanding how passive users may interact with online content may contribute the design and operation of OSNs which in turn might improve their learning experience. Another limitation is the small scale of the longitudinal study (48 participants). Even so, capturing and analysing the interaction of 48 participants through a long period of time allows us to identify the changes in patterns of interaction and test associated structural network effects to enhance our understanding of their interaction.

In this study, we identified that there was some small group learning that occurred in the form of case-based webinar discussion. Future study may be worthwhile to investigate small group learning activities by employing SNA subgroup measures (e.g. component analysis) to understand how learning groups formed and how their group members collaborated. Also, further study may be worthwhile to investigate how other homophily effects (e.g. clinical area of interest or domain of expertise) might affect changes in the patterns and structure of interaction.

Although the present study does not give a complete picture of the learning process, using SNA allows us to identify the level and pace of interaction and make inferences about how and why learning occurred. In the next chapter, the content of discussion from these active users will be investigated, which will provide indirect evidence of their learning, and which will have implications for determining their learning and development needs.
Chapter 4 Using Topic Modeling for Discovering Learning Topics in an Online Social Network for Medical Practitioners

The relevant content of the following publication has been integrated into this chapter:


4.1 Introduction

From the study reported in Chapter 3, we have gained an understanding of informal learning behaviours in an online discussion forum for medical practitioners, by using SNA to identify their patterns of interaction. However, there is still much to be explored in terms of the textual dialogues among them, particularly regarding how those dialogues support the process of learning in this forum. The aim of the study in Chapter 4 is to demonstrate how topic modeling can be used to discover the topics of interest to the medical practitioners in this forum. Identifying these topics is expected to provide indirect evidence of their learning, and to have implications for determining their learning and development needs. This in turn can inform the design of learning content that is relevant to their professional practice and motivates them to participate in this community.

4.2 Background and Related Work

4.2.1 Content Analytics

As more and more textual data is generated online and human annotation becomes impossible, the field of content analytics come to the fore. It is considered as one type of Learning Analytics (LA) (Ferguson & Shum, 2012). Content analytics focuses on building analytical models that rely on computational tools to process content, including learner-produced content such as online discussion messages. Content analytics is
defined as “automated methods for examining, evaluating, indexing, filtering, recommending, and visualizing different forms of digital learning content, regardless of its producer (e.g. instructor, student) with the goal of understanding learning activities and improving educational practice and research” (Kovanović et al., 2015).

Applications of content analytics are diverse. As noted in Chapter 2, examples are: topic modeling to identify themes and topics in learning content (Tobarra et al., 2012); classification techniques to automate different content analysis coding schemes (Kovanović et al., 2016); and collaborative filtering techniques to discover and recommend educational resources (Walker et al., 2004).

### 4.2.2 Topic Modeling

Topic modeling is one popular method in content analytics (Kovanović et al., 2015). It is a statistical method that analyses the words of the original texts to discover the themes that run through them, how those themes are connected to each other, and how they change over time (Blei, 2012). There are many techniques that are used to obtain topic models, but the simplest and most widely used is Latent Dirichlet Allocation (LDA). LDA is a generative probabilistic model (McCallum, 2002). The basic idea of LDA is that documents are presented as a random mixture of topics, where each topic is a probability distribution over a given vocabulary of words (Blei, Ng, & Jordan, 2003).

Many researchers have employed topic modeling to explore the themes in dialogues in online learning environments. However, to the best of our knowledge, its application in discovering topics discussed within an online community of medical practitioners is novel. Tobarra et al. (2012) used it to discover topics of interest in the forum of a Learning Management System for improving the structure and contents of education courses. Portier et al. (2013) used it together with sentiment analysis to identify improvements that enhance social support in an online cancer patient community. More recently, Ezen-Can, Boyer, Kellogg, and Booth (2015) used it to understand the topics of discussion in the forum of an open online course for educators.
4.3 Methods

4.3.1 Data

This study focused on identifying the topics of interest from the active users in the case study forum and understanding how the learning content (i.e. online discussion) supported their informal learning. The data for the present study comprised all the posts made by the 48 active users during the three-year period from 2012 to 2014. Based on the data selection criteria, we found 154 discussion threads, each receiving between one and 58 replies. In total 1604 forum posts (105,063 words) were extracted from the dataset.

4.3.2 Topic Modeling Using MALLET

To identify the topics of interest using the extracted data, topic models were generated using the MALLET library (version 1.0) in R. The MALLET (Machine Learning for Language Toolkit) implements the LDA algorithm and automates the process of topic discovery from a large volume of text. In this study, each forum post is considered as a document that may contain a mixture of topics. The MALLET program was used to generate clusters of words (i.e. topics) that frequently occur together within forum posts.

4.3.3 Procedure

Data preparation: The raw text data were pulled from the se_forum_posts database table into Excel Spreadsheet Editor using SQL queries and saved in csv format, then the data file (a csv file) was loaded into RStudio\(^6\) for analysis.

The TM library (version 0.6) in R was loaded into RStudio to pre-process the text. The text was cleaned by removing anything other than English letters or spaces. To improve the coherence of generated topics, stop words were removed from the full text based on the standard list of stop words of MALLET\(^7\). The popular words (e.g. lol, cheers, pretty, nice, yrs) and any specific words associated with country/state/city and personal names that appeared in this dataset were also removed. All words were also stemmed to

\(^6\)RStudio is a free and open-source integrated development environment (IDE) for R, a programming language for statistical computing and graphics. https://www.rstudio.com

\(^7\)For instructions to download the standard list of stop words of MALLET, please go to http://mallet.cs.umass.edu/import-stoplist.php.
retrieve their stems so that various forms of a word would be counted together when counting word frequency. These pre-processing steps help obtain a cleaned full text and reduced the number of words in the dataset from 105,063 to 54,873.

**Topic model generation:** To generate topic models using MALLET, four key steps were followed below, as guided by the topic modeling tutorials on MALLET website (http://mallet.cs.umass.edu) and R Reference manual for MALLET\(^8\). Appendix C presents the major R code used for this analysis.

1. Defining and creating the topic model using cleaned full text.

2. Identifying the optimal number of topics for the topic model by defining two variables (number of topics, number of sampling iterations). Different numbers of topics were specified to generate four models (Models 1 – 4). The initial number of topics was set to 15 by inspecting all the 154 thread titles and noting from inspection that there are approximately 14 broad topics in the dataset.

3. Training the model by specifying the number of iterations. The dataset of this size usually has the default sampling iteration set to 400. Since increasing the number of iterations may improve topic coherence (McCallum, 2002), the iteration was increased to 800 to generate two additional models (Models 5 – 6). Table 4.1 depicts the variables defined for these different topic models.

4. Generating topic model results including the probability of topics in documents, the probability of words in topics, and top keywords for each topic.

**Table 4.1.** The defined variables for topic models

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Topics (T)</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>30</td>
<td>20</td>
<td>25</td>
</tr>
<tr>
<td>Number of Iterations (I)</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>400</td>
<td>800</td>
<td>800</td>
</tr>
</tbody>
</table>

**Topic inference:** Based on an inductive analysis approach (Thomas, 2006), topics were manually inferred by the researcher using clusters of words produced by topic models.

\(^8\) https://cran.r-project.org/web/packages/mallet/index.html
Since each topic is a probability distribution over words, the top ten words were chosen to inspect for each cluster. This is based on the assumption that more words per cluster might make it more difficult to infer a meaningful topic for each cluster.

**Topic refinement:** The inferred topic labels were refined by selecting and reviewing the contents of the top five posts with consideration for each word cluster (topic). The top five posts for each topic were identified by inspecting the probability of each topic appearing in each post, which was obtained by employing the function `mallet.doc.topics` in MALLET.

Figure 4.1 presents a snapshot of the probability results generated by MALLET for a topic model in excel. The column value (V1 – V10) is the unique identity assigned to each post, the row value (1 – 10) is the unique identify assigned to each topic, the numbers (e.g. 0.000314) are the probability of each topic appearing in each post.

![Table]

<table>
<thead>
<tr>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
<th>V5</th>
<th>V6</th>
<th>V7</th>
<th>V8</th>
<th>V9</th>
<th>V10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.000314</td>
<td>0.006216</td>
<td>0.009768</td>
<td>0.003943</td>
<td>4.28E-05</td>
<td>0.00182</td>
<td>0.000184</td>
<td>0.01225</td>
<td>2.42E-05</td>
</tr>
<tr>
<td>2</td>
<td>0.004277</td>
<td>0.000117</td>
<td>2.10E-05</td>
<td>5.63E-05</td>
<td>3.11E-05</td>
<td>0.005411</td>
<td>0.000134</td>
<td>4.17E-05</td>
<td>0.007213</td>
</tr>
<tr>
<td>3</td>
<td>0.000310</td>
<td>0.000163</td>
<td>2.92E-05</td>
<td>7.82E-05</td>
<td>4.32E-05</td>
<td>1.62E-05</td>
<td>0.000186</td>
<td>5.80E-05</td>
<td>2.44E-05</td>
</tr>
<tr>
<td>4</td>
<td>0.000318</td>
<td>0.000164</td>
<td>2.94E-05</td>
<td>7.87E-05</td>
<td>4.35E-05</td>
<td>1.63E-05</td>
<td>0.000187</td>
<td>5.84E-05</td>
<td>2.46E-05</td>
</tr>
<tr>
<td>5</td>
<td>0.000323</td>
<td>0.000167</td>
<td>2.98E-05</td>
<td>8.08E-05</td>
<td>3.10E-05</td>
<td>1.65E-05</td>
<td>0.000191</td>
<td>5.93E-05</td>
<td>2.50E-05</td>
</tr>
<tr>
<td>6</td>
<td>0.000325</td>
<td>0.000168</td>
<td>3.00E-05</td>
<td>8.02E-05</td>
<td>3.08E-05</td>
<td>1.66E-05</td>
<td>0.000191</td>
<td>5.96E-05</td>
<td>2.51E-05</td>
</tr>
<tr>
<td>7</td>
<td>0.000287</td>
<td>0.000148</td>
<td>2.64E-05</td>
<td>7.09E-05</td>
<td>3.91E-05</td>
<td>1.00364</td>
<td>0.000168</td>
<td>5.26E-05</td>
<td>2.21E-05</td>
</tr>
<tr>
<td>8</td>
<td>0.000312</td>
<td>0.000161</td>
<td>2.88E-05</td>
<td>7.71E-05</td>
<td>4.26E-05</td>
<td>1.60E-05</td>
<td>0.000183</td>
<td>5.72E-05</td>
<td>2.41E-05</td>
</tr>
<tr>
<td>9</td>
<td>0.000289</td>
<td>0.000149</td>
<td>2.67E-05</td>
<td>7.14E-05</td>
<td>3.94E-05</td>
<td>1.48E-05</td>
<td>0.00017</td>
<td>5.30E-05</td>
<td>2.23E-05</td>
</tr>
<tr>
<td>10</td>
<td>0.000358</td>
<td>0.000185</td>
<td>3.13E-05</td>
<td>8.86E-05</td>
<td>8.01E-05</td>
<td>1.84E-05</td>
<td>0.000211</td>
<td>6.58E-05</td>
<td>2.77E-05</td>
</tr>
</tbody>
</table>

**Figure 4.1.** A snapshot of the probability results generated by MALLET

To identify the top five posts for each topic, the matrix was first transposed using Excel Spreadsheet Editor. That is, transposing topic IDs to columns and post IDs to rows. Then, the top posts for each topic can be identified by sorting the probability of each topic from Largest to Smallest. The process of refinement helped identify further duplicates and improve the accuracy of the inferred topic labels.

**Topic evaluation:** The performance of topic models is typically evaluated using quantitative intrinsic methods such as computing the probability of held-out documents. However, it has been shown that this measure is not always a good predictor of human judgment (Chang, Gerrish, Wang, Boyd-Graber, & Blei, 2009). In this study, the topics
were evaluated based on human judgment using F-measure (Van Rijsbergen, 1979), which is often used in the field of information retrieval. There are four performance metrics considered (i.e. accuracy, precision, recall, and F-score, as defined in Table 4.2).

**Table 4.2.** The definition of performance metrics considered for F-measure

<table>
<thead>
<tr>
<th>Performance metric</th>
<th>Description</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>The percentage of the posts identified are expected to belong to a refined topic.</td>
<td>TP + TN / (TP + FP + FN + TN)</td>
</tr>
<tr>
<td>Precision</td>
<td>The percentage of posts correctly identified as belonging to a refined topic.</td>
<td>TP / (TP + FP)</td>
</tr>
<tr>
<td>Recall</td>
<td>The percentage of posts identified as belonging to any topic.</td>
<td>TP / (TP + FN)</td>
</tr>
<tr>
<td>F-score</td>
<td>The harmonic mean of precision and recall, which can be interpreted as a weighted average of the Precision and Recall.</td>
<td>2 × (Precision × Recall) / (Precision + Recall)</td>
</tr>
</tbody>
</table>

Forty forum posts were randomly selected from the dataset to validate the refined topics using F-measure. It is considered that a post is True Positive (TP) when any part of the post content matches a refined topic; a post is False Positive (FP) when the post content does not match a refined topic; a post is False Negative (FN) when the post content suggests a discernable topic (may be a refined topic or any new topic that has not been identified); a post is True Negative (TN) when the post content does not suggest any discernable topic.

**4.4 Results and Discussion**

**4.4.1 Topic Model Comparison**

After inferring the topics of generated topic models, the number of refined topics was compared from each topic model. It became evident that a topic model with \( T = 20 \) would be more appropriate than \( T = 15 \), or \( T = 25 \), or \( T = 30 \). As shown in Table 4.3, Model 1 (\( T = 15 \)) generated 9 refined topics which indicate that setting too few topics could result in not covering all topics. Model 3 (\( T = 25 \)) generated 11 refined topics which indicate that setting too many topics could result in duplications (five pairs of word clusters represent the same topic). Model 4 (\( T = 30 \)) generated only 9 refined
topics which indicate that setting too many topics could even result in uninterpretable topics.

For this dataset, a topic model with $I = 400$ would be more appropriate than $I = 800$. The number of refined topics generated from Model 5 and Model 6 suggests that increasing the number of iterations did not result in better topic models, as the composition and quality of the resulting topics only increased to a certain point and then levelled off. From the results, we concluded that Model 2 ($T = 20, I = 400$) seems to be the topic model that best describes the topics of interest discussed by medical practitioners in this forum.

Table 4.3. The total number of inferred and refined topics for topic models

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Model 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of inferred topics</td>
<td>11</td>
<td>14</td>
<td>15</td>
<td>13</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>Number of refined topics</td>
<td>9</td>
<td>13</td>
<td>11</td>
<td>9</td>
<td>12</td>
<td>9</td>
</tr>
</tbody>
</table>

4.4.2 Topics of Interest

Table 4.4 shows the inferred, refined topics and their associated word clusters for the selected topic model (i.e. Model 2). Of the 20 word-clusters generated from the model, 14 unique topics were inferred; two word-clusters were interpreted as referring to the same topic, and four word-clusters were indicated as “Not Applicable” (N.A.) as no meaningful topic could be inferred. After refining the inferred topics by reviewing the selected posts, 13 unique topics were found. They are identified as a mixture of clinical and non-clinical topics. The Clinical Topics (CT) include ‘palliative care’, ‘rheumatology’, ‘evidence-based medicine’, ‘statins use’, ‘vitamin use’, ‘vaccines’, ‘women’s health check’, and ‘fibromyalgia’; the Non-Clinical Topics (NCT) include ‘patient fees’, ‘training’, ‘pharmaceutical industry’, ‘policy’, and ‘workload’. 
Table 4.4. The refined topics for the selected topic model

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Topic weight</th>
<th>Topic words</th>
<th>Inferred topic labels</th>
<th>Refined topic labels</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.19594</td>
<td>pay medicare bulk work charge fee private service money government</td>
<td>Bulk billing</td>
<td>Patient fees (NCT)</td>
</tr>
<tr>
<td>2</td>
<td>0.18415</td>
<td>prescription script pharmacy addict drug pharmacist pbs authority day write</td>
<td>Pharmaceutical industry</td>
<td>Pharmaceutical industry (NCT)</td>
</tr>
<tr>
<td>3</td>
<td>0.0322</td>
<td>food car house fridge eat poor store hot change balance</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td>4</td>
<td>0.1898</td>
<td>patient pay medicine stress society reduce current government rate finance</td>
<td>Health cost</td>
<td>Patient fees (NCT)</td>
</tr>
<tr>
<td>5</td>
<td>0.08499</td>
<td>care nurse hour visit service palliative home medical provide rural</td>
<td>Palliative care</td>
<td>Palliative care (CT)</td>
</tr>
<tr>
<td>6</td>
<td>0.04848</td>
<td>medical profession racgp public doctor ahpra nurse health wrote report</td>
<td>Training</td>
<td>Training (NCT)</td>
</tr>
<tr>
<td>7</td>
<td>0.05038</td>
<td>point vaccine medicine understand body view generate form base suggestion</td>
<td>Vaccines</td>
<td>Vaccines (CT)</td>
</tr>
<tr>
<td>8</td>
<td>0.03467</td>
<td>refer comment expert issue interest person lack call present programme</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
<tr>
<td>9</td>
<td>0.07868</td>
<td>restrict country year work moratorium hospital system area dws law</td>
<td>Policy</td>
<td>Policy (NCT)</td>
</tr>
<tr>
<td>10</td>
<td>0.10449</td>
<td>trial statin evidence effect prevent group side benefit interest study</td>
<td>Statins use</td>
<td>Statins use (CT)</td>
</tr>
<tr>
<td>11</td>
<td>0.07429</td>
<td>effect level side dose disease vitamin high symptom difference drug</td>
<td>Vitamin use</td>
<td>Vitamin use (CT)</td>
</tr>
<tr>
<td>12</td>
<td>0.09575</td>
<td>pain inject joint muscle guidance bursa knee stretch tear elbow</td>
<td>Elbow bursa treatment</td>
<td>Rheumatology (CT)</td>
</tr>
<tr>
<td>13</td>
<td>0.07989</td>
<td>evidence point medicine base comment treatment understand body view ahpra</td>
<td>Evidence-based medicine</td>
<td>Evidence-based medicine (CT)</td>
</tr>
<tr>
<td>14</td>
<td>0.07777</td>
<td>examination check breast pap women history year present cancer diagnose</td>
<td>Women’s health checks</td>
<td>Women’s health checks (CT)</td>
</tr>
<tr>
<td>15</td>
<td>0.1173</td>
<td>human organ end therapy cell central</td>
<td>N.A.</td>
<td>N.A.</td>
</tr>
</tbody>
</table>
There are more clinical than non-clinical topics identified from the dataset. However, the weights of the topics imply that non-clinical topics were more frequently discussed than clinical ones.

With regards to the clinical topics, palliative care, rheumatology, and evidence-based medicine appeared to generate some in-depth discussion among the participants. By inspecting a number of specific posts on the topics relating to women’s health checks, fibromyalgia, the use of statins, vaccines, and vitamin, we noted that the participants were interested in benchmarking their practices. This is understandable as clinical practice can be conducted differently in different places; OSNs have been found to enable medical practitioners to share different ways of performing the same practice and benchmark the most effective one (Millar et al., 2016).

The non-clinical topics identified from this dataset are mostly controversial (including policy, workload and patient fees). This finding is consistent with previous studies that have demonstrated that medical practitioners are particularly interested in discussing controversial topics in an OSN (Panahi, Watson, & Partridge, 2014). In addition, the participants were keen to keep themselves up-to-date on the latest information and news in the field; this is reflected in the topics relating to policies, training, and the pharmaceutical industry.
4.4.3 Topic Evaluation

The 13 identified topics were evaluated using F-measure against 40 randomly selected posts from the dataset. Figure 4.2 presents the results of TP, FP, FN and TN, as well as the four performance metrics considered for F-measure.

![Figure 4.2. The results of F-measure](image)

The results of accuracy, precision, recall and F-score were calculated based on the formulas presented in Table 4.2. The Accuracy of 0.53 indicates that the topic model is likely to capture 53% of the topics in any randomly selected posts. The F-score of 0.66 informs us that the topic model correctly captures 66% of the overall topics in this random selection of 40 posts.

4.5 Conclusions and Implications

This study employed topic modeling to identify the topics of interest that emerged from the medical practitioners active in an online discussion forum. The results suggest that topic modeling could be used to identify learning topics efficiently from the large amount of textual dialogue generated in an OSN. We found the active users in this forum engaged in sharing knowledge and profession-related information, which can be considered as forms of informal learning. We found they were interested in discussing both clinical and non-clinical topics. Clinical topics were related to their practices, sharing practical and experiential knowledge and providing benchmarks, which included ‘palliative care’, ‘rheumatology’, ‘evidence-based medicine’, ‘statins use’, ‘vitamin use’, ‘vaccines’, ‘women’s health check’, and ‘fibromyalgia’. The non-clinical topics were related to controversial topics, latest news and information in the field, which included ‘pharmaceutical industry’, ‘patient fees’, ‘policy’, ‘workload’, and ‘training’.
Identifying topics using topic modeling could provide OSN designers and operators with guidance on facilitating online discussion that is most relevant to the learning needs of medical practitioners. In this OSN, it has been found that non-clinical topics were more frequently discussed than clinical ones by these medical practitioners. Without knowing the context of learning, we could not advocate having non-clinical topics as the focus of their online discussion, but it is important to consider how to help medical practitioners to deal with the challenge of keeping themselves up-to-date on non-clinical and other work-related information. In addition, it might be worth considering proposing common clinical topics relating to their clinical practices that allow them to share practical and experiential knowledge and meet their needs for benchmarking.

This study has several limitations. First, the evaluation of the topic model suggests that the topic model captured 66% of the overall topics in a random selection of 40 posts in the forum. However, as we have found no previous work on topics discussed by an OSN for medical practitioners to compare our results with, it is inconclusive whether the topics we identified are typical or atypical of those discussed by medical practitioners. A second limitation is regarding the aptness of labels that were assigned to topics, based on inferences by a non-biomedical researcher using an inductive analysis approach. In a future study, topics could be improved by mapping to a structured Medical Subject Heading (MeSH) lexicon that provides formal representation of the knowledge contained within the network. The last limitation is that passive users were not tracked in our study due to the limitations of the case study data source. This means the topics identified only apply to the active participants of this OSN.

In future work, additional metadata may be fitted into the topic model, for example, including the identity of the authors of discussion forum posts. This would enable investigation of authors’ similarity based on their discussion of topics. This could help identify the groups of medical practitioners who may have similar learning needs; there may be value over time in gathering them together in the OSN for small group learning activities.

In the next chapter, the context of learning for these active users will be investigated by analysing context data collected in the case study forum. This will complement findings
about content and interaction, to complete the series of three studies designed to show how using LA systematically can give a detailed understanding of the learning process in an OSN for medical practitioners.
Chapter 5 Using Context Analysis for Analysing the Context of Learning in an Online Social Network for Medical Practitioners

The relevant content of the following publication has been integrated into this chapter:


5.1 Introduction

Previous chapters have used analytic methods to investigate learning behaviours by identifying the patterns of the interaction among medical practitioners (Chapter 3), and the content of textual dialogues among medical practitioners to understand how those dialogues support the process of learning (Chapter 4). However, working within theories about how learning occurs in OSNs requires understanding the learning context to get a more detailed understanding of the learning process in the case study forum.

The aim of the study in this chapter is to demonstrate how context analysis can be used to understand the learning context of the active users studied in the forum. This study first introduces the contextual factors that are critical for analysing the context of learning in OSNs for medical practitioners, then it uses these factors to analyse context data collected from the forum. The intention of this study is to identify the learning context information that may help facilitate meaningful learning experiences in OSNs for medical practitioners, as well as to enhance the interpretation of the findings about learning interaction and learning content from previous studies.
5.2 Background and Related Work

All learning occurs within a context. The nature of the context and how this context relates to the concepts being learned has been widely shown to have an effect on learning outcomes (Balsam, 2014; Tessmer & Richey, 1997). The nature of learning occurring in OSNs is self-directed and this requires medical practitioners to have self-regulated learning skills (Sandars & Walsh, 2014), but the ability to self-regulate learning is shaped by both personal-psychological and contextual factors (Zimmerman & Schunk, 2012). Further, it is essential that these processes be learned “in context” or in relation to the specific tasks of interest, in order for self-directed learning to be relevant to medical education (Cleary, Durning, Gruppen, Hemmer, & Artino, 2013).

The role of context is important in medical professional learning, as highlighted by learning theories that are used to guide learning development interventions for medical practitioners (Mann & Sargeant, 2013). These theories include: adult learning theory (Knowles, 1980), which identifies experience, personal goals and practical needs as motivating individual learning; complexity theory (Griffiths, 2002), which highlights both environmental and personal factors that influence learning; self-directed learning theory (Candy, 1991), which emphasises the role of motivational beliefs, and the importance of contextual and environmental factors.

People learn more effectively if the context of learning makes it meaningful to them. In essence, learning context inevitably affects learning and performance (Jarvis, 2011). Dey (2001) gave the most widely accepted definition for context, which is “any information that can be used to characterize the situation of an entity” (a learner, in this case). Hager and Halliday (2007, p. 159) view context as “the surroundings in which learning occurs and the possible influences that these surroundings have on what is learnt”. Zimmermann, Lorenz, and Oppermann (2007) consider that elements for the description of context information fall into five categories: individuality, activity, location, time, and relations, whereas others suggest four categories: location, identity, time, and environment (Gross & Specht, 2001).

Various analysis frameworks have been proposed for understanding the learner’s context in online learning settings. Notably, Tessmer and Richey (1997) pointed out that
contextual analysis is important but has been missed within instructional design models for learning management systems. They consider context to be composed of levels (prior, during and after learning) as well as critical factors related to the learner, immediate and organisational environments. Each of these levels contains a variety of contextual factors. Leung (2003) suggested a generic conceptual model, which integrates four contextual issues in design and implementation of computer-based learning. These issues are topic selection, authenticity, complexity, and multiple perspectives.

The analysis of context in online learning is limited, although it has been widely researched in mobile computing (Musumba & Nyongesa, 2013). As part of a study of the role of teachers within a network learning community, De Laat et al. (2007) collected context data separately from the online setting, via interviews with learners, to understand the patterns of online interaction. Bicans (2015) analysed the context information (including time, activity, resources and user) collected in an intelligent tutoring system, and demonstrated how the context information may be used to improve learning experiences. Hood, Littlejohn, and Milligan (2015) investigated how learners’ context may influence their ability to self-regulate in a Massive Open Online Course (MOOC); however, that study considered only the job role context of learners.

5.3 Proposed Contextual Factors for Medical Practitioners’ Informal Learning in OSNs

A standard model for describing the learning context would aid consistent collection and analysis of data about medical practitioners’ informal learning in OSNs. As described in the previous section of this chapter, a literature review of context models proposed for online learning shows inconsistent approaches. Here, we identify the contextual factors that we consider to be critical for analysing the informal learning context in OSNs for medical practitioners, based on our synthesis of the context models in the literature. As shown in Table 5.1, these factors are of two main types: learner factors and environmental factors.
Table 5.1. The critical contextual factors

<table>
<thead>
<tr>
<th>Contextual factor</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner factors</td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>Age, gender, practice location, etc.</td>
</tr>
<tr>
<td>Work experience</td>
<td>Job role, years of experience, other professional activity, etc.</td>
</tr>
<tr>
<td>Education</td>
<td>Qualifications, professional memberships, etc.</td>
</tr>
<tr>
<td>Clinical areas of interest</td>
<td>General Practice, mental health, etc.</td>
</tr>
<tr>
<td>Environmental factors</td>
<td></td>
</tr>
<tr>
<td>Time and Location</td>
<td>When and where the learner commonly engages in learning activities</td>
</tr>
<tr>
<td>Activity</td>
<td>Other learning activities that the learner is currently involved in or will be in future, apart from in an OSN, e.g. exam preparation.</td>
</tr>
<tr>
<td>Relations</td>
<td>The social relations relevant to learning that the learner has with other people outside of an OSN, e.g. mentoring in their workplace.</td>
</tr>
<tr>
<td>Recognition</td>
<td>Forms of recognising the learner’s effort in an OSN, e.g. through CPD points</td>
</tr>
<tr>
<td>Application</td>
<td>Perceived and actual use for what is learned, e.g. a complex patient case</td>
</tr>
</tbody>
</table>

Learner factors are considered since learners are often regarded as a context in themselves (Tessmer & Richey, 1997; Zimmermann et al., 2007). So that personalised learning can be provided, we have determined that the basic learner factors include selected demographic characteristics, work and educational background, and clinical areas of interest.

Environmental factors are considered since learning occurs in one or more environments aside from in the OSN itself. These factors include the context relating to time, location, activity, and relations (Zimmermann et al., 2007), as well as recognition and application (Tessmer & Richey, 1997). We elaborate on these below.

Time as a context identifies the time of day when a learner usually is online in the OSN. Location as a context provides information on the physical learning location of the learner. Both of these help us to understand the learning situation of the learner so that
appropriate learning content and activity based on preferred time and location can be provided in the OSN.

Activity as a context covers the external learning activities that the learner is currently involved in or will be in future, apart from the interaction in the OSN under study. Knowing these activities of learners help us to understand what they want to achieve so that more responsive learning support and targeted discussion can be provided in the OSN.

Relations as context captures the social relations a learner has with other people outside of the OSN. The social relations are considered because prior interpersonal relationships and interactivity among medical practitioners may impact on the learning process (Margolis & Parboosingh, 2015). For example, a learner may interact with someone from their workplace in the OSN; this may have unforeseen consequences if the learner intended that the topic was to be kept confidential from co-workers. More positively, the relations of learners make possible the discovery or establishment of common ground so that formation of relevant learning groups can be facilitated in an OSN.

Both recognition and application are context information that is particularly relevant to self-directed learning. Recognition means the ways of recognising learners’ learning efforts in an OSN and this may motivate learners and contribute to their long-term engagement in an OSN (Sandars & Langlois, 2005). It has been suggested that recognition can be given through provision of CPD points which may lead to increased acceptance for OSNs (Millar et al., 2016). Several countries have created national credit systems for CPD, although there is a wide variation in the accreditation systems as there is in the regulatory standards governing their operation (Silva et al., 2012; Sriharan, Murray, Pardell, & Silva, 2009).

Application covers the learner’s perception of the utility of a learning activity, as well as the actual application of learned information or skills in practice. Context about perceived and actual utility of the learning can support the design of learning opportunities that are useful and relevant to a learner’s routine practice and long term development (Tessmer & Richey, 1997). For example, information can be collected on
the use-value of learned information or skills and the frequency with which they are applied in practice.

5.4 Methods

5.4.1 Data

Context analysis of the case study OSN was conducted based on the proposed contextual factors in this study. The relevant context data were collected from forum database tables including the tables se_user, se_profile_info, and se_forum_post. For this study, the context data of the 48 active users in the forum were collected. By conducting context analysis on the active users, this study identifies common aspects of the learning context that may be shared among the active group in this community.

5.4.2 Measures

The learning context of the active users in this forum was analysed based on the contextual factors we have proposed for medical practitioners’ informal learning in OSNs. Table 5.2 presents the contextual factors considered in this dataset with associated measures and data sources.

Table 5.2. The measures of contextual factors for this study

<table>
<thead>
<tr>
<th>Contextual factor</th>
<th>Measure</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>Age</td>
<td>date_of_birth</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>gender</td>
</tr>
<tr>
<td></td>
<td>Practice Location</td>
<td>location</td>
</tr>
<tr>
<td></td>
<td>Jurisdictional location</td>
<td>location</td>
</tr>
<tr>
<td>Work experience</td>
<td>Job role</td>
<td>about_me</td>
</tr>
<tr>
<td></td>
<td>Years of experience</td>
<td>about_me</td>
</tr>
<tr>
<td></td>
<td>Practice status</td>
<td>about_me</td>
</tr>
<tr>
<td></td>
<td>Other professional activity</td>
<td>about_me</td>
</tr>
<tr>
<td>Educational background</td>
<td>Place of graduation</td>
<td>university</td>
</tr>
</tbody>
</table>
### Contextual Factor Measures Data Source

<table>
<thead>
<tr>
<th>Contextual factor</th>
<th>Measure</th>
<th>Data source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional memberships</td>
<td>college_membership</td>
<td></td>
</tr>
<tr>
<td>Clinical areas of interest</td>
<td>Self-reported interests/speciality</td>
<td>about_me</td>
</tr>
<tr>
<td>Environment factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>Time of participation</td>
<td>Forum activity log</td>
</tr>
<tr>
<td>Location</td>
<td>Location of participation</td>
<td>Forum activity log</td>
</tr>
</tbody>
</table>

#### 5.4.3 Procedures

All context data relevant to learner factors were retrieved from the `se_users` and `se_profile_info` tables of the forum database via SQL. After performing data cleansing, ‘demographics’ and ‘education background’ were analysed first by using raw data (date_of_birth, gender, location, university, college_membership), then the context information for other contextual factors was obtained by processing the raw data (i.e. analysing, validating and adding metadata to the raw data).

All the free text in data attribute `about_me` were manually inspected to obtain context information for ‘work experience’ and ‘clinical areas of interest’. This includes 48 short paragraphs of free text (one supplied in the forum by each user). Below provides an example of the free text in `about_me`, to demonstrate how the information in `about_me` can be analysed to identify the context information about ‘work experience’ and ‘clinical areas of interest’.

“**General Practitioner** in a busy rural practice undertaking complex older person primary care. Dr [person name] is an experienced General Practitioner (over 10 years of work experience) with an interest in **chronic disease management**. He also has an interest in **skin examination** and has performed minor skin cancer removal surgery for many years. He has worked around the world in many and varying positions including on the ski fields and on board a cruise liner. He is an **examiner** for the Royal Australian College of General Practitioners and the Australian Medical Council. He is **teaching** at the Medical School of the [name of university] and still works at the Emergency Department of [hospital name].”

In this example, we can know that the user’s job role was General Practitioner and had over 10 years of work experience; his practice status was active; his other professional
activities included teaching (lecturer) and examiner; his clinical areas of interest included dermatology and chronic disease management.

To get the context data relevant to environment factors, the interaction history of each user was analysed to obtain the preferred times and locations that learners go online. The value of created (date and time) of all the posts created by the 48 active users was retrieved from the database table forum_post into Excel Spreadsheet Editor using SQL. Then the raw data in Excel were loaded into RStudio for analysis.

The GGLOT2 library (version 2.1.0) in R was loaded into RStudio for visualising the daily distribution for time intervals. All records were re-coded, first based on defined time periods (including morning, afternoon, evening, and night), then the distribution of participation time was visualised using bar charts.

Actual data about what location a user was in when accessing the forum proved to be unavailable from any source in the case study OSN. For the purpose of simulating the potential of analytic methods to process such context data, the location of users’ participation was inferred by the researcher, based on the time-of-day pattern of actual participation.

5.5 Results and Discussion

5.5.1 Learner Context

Demographics. As shown in Table 5.3, the majority of the participants were males aged 55+ years, which shows that experienced medical practitioners actively participated in the forum. This implies that it is important to consider the learning needs of older medical practitioners when designing the learning in an OSN, and support their learning in later stages of their career so that they can provide their valued contribution to the community as long as possible (Joyce, Wang, & McDonald, 2015).

Using the Australian Standard Geographical Classification (Australian Bureau of Statistics, 2011), we found that the majority of participants were practicing in a major city in the most populous Australian States (i.e. Queensland, New South Wales, and Victoria). These findings overturn assumptions that this type of learning is most relevant to professionals in regional and remote areas (Brown, Ryan, & Harris, 2014).
Table 5.3. The results for learner factors

<table>
<thead>
<tr>
<th>Learner factors</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td><strong>Results</strong></td>
</tr>
<tr>
<td>Age group</td>
<td>35-44 years (17%), 45-54 years (15%), 55-64 years (46%), 65-74 years (13%), 75+ years (4%)</td>
</tr>
<tr>
<td>Gender</td>
<td>Male (70%), Female (30%)</td>
</tr>
<tr>
<td>Practice location</td>
<td>Major cities (67%), Inner regional (22%), Outer regional (7%), Remote (0%), Very remote (4%)</td>
</tr>
<tr>
<td>Jurisdictional location</td>
<td>QLD (35%), NSW (30%), VIC (15%), Others (20%)</td>
</tr>
<tr>
<td><strong>Work experience</strong></td>
<td><strong>Results</strong></td>
</tr>
<tr>
<td>Job role</td>
<td>GP (96%), Physician (4%)</td>
</tr>
<tr>
<td>Year of experience</td>
<td>&lt;10 years (11%), 10-19 years (15%), 20-29 years (9%), 30+ years (61%)</td>
</tr>
<tr>
<td>Practice status</td>
<td>Active (57%), Retired (22%), Semi-retired (16%)</td>
</tr>
<tr>
<td>Other activity</td>
<td>22% undertaking extra professional activities, included working as a medical officer at hospitals, teaching or lecturing at universities, providing online consultation, working in community services.</td>
</tr>
<tr>
<td><strong>Educational background</strong></td>
<td><strong>Results</strong></td>
</tr>
<tr>
<td>Place of graduation</td>
<td>Australia (72%), outside Australia (22%)</td>
</tr>
<tr>
<td>Professional qualification</td>
<td>FRACGP (61%), RACGP (22%), None recorded (17%)</td>
</tr>
<tr>
<td>Clinical areas of interest</td>
<td>General practice, dermatology, women’s health, chronic disease management, general medicine, diabetes, and sports medicine</td>
</tr>
</tbody>
</table>

*Work experience.* 96% of the participants were General Practitioners (GPs). Of those, 56% were principal GPs and/or practice owners, implying that those with supervisory and/or management responsibility were more active in the case study forum. The distribution of years of experience is consistent with the age factor, which confirms that more experienced medical practitioners with 30+ years of work experience were more active in the case study forum for learning. Examining their practice status showed that 38% of participants were retired or semi-retired, which reveals that even in retirement
GPs are willing to engage, and keen on using OSNs to maintain connection with peers and overcome isolation (Barnett, Jones, Bennett, Iverson, & Robinson, 2016).

We identified 22% participants as having ‘portfolio careers’ (Eyre, Mitchell, Milford, Vaswani, & Moylan, 2014), that is, multiple types of professional activities. For example, they were in private practice and at the same time also working as a medical officer at one or more hospitals, teaching or lecturing at a university, providing online consultations, or involved in community services. This shows that even very busy professionals were prepared to commit to learning in the case study forum.

**Educational background.** The majority of participants (61%) held the Fellowship of the Royal Australian College of General Practitioners (RACGP), and the majority of participants (72%) graduated in Australia, suggesting a high level of homogeneity in this group. However, 22% graduated outside Australia, and this suggests that the learning design of OSNs should consider cultural and linguistic differences in the backgrounds of medical practitioners.

**Interests.** All participants expressed a major interest in general practice. More than half of the participants (54%) were interested in or had developed knowledge of sub-specialties. The clinical areas of most interest among the participants were dermatology, women’s health, chronic disease management, general medicine, diabetes, and sports medicine. This finding is somewhat consistent with a recent national GP survey (NPS MedicineWise, 2014) that identified the top three clinical areas of interest to GPs as chronic pain management, cardiovascular health and diabetes. This implies that GPs are keen to develop sub-specialist areas. Thus, supporting their learning by facilitating more discussion on clinical topics in such an OSN seems advisable.

In comparing their clinical areas of interest from context analytics here, with the topics of interest which were identified from topic modeling in Chapter 4 (i.e. ‘palliative care’, ‘rheumatology’, ‘evidence-based medicine’, ‘statins use’, ‘vitamin use’, ‘vaccines’, ‘women’s health check’, and ‘fibromyalgia’), we found that the majority of these topics were discussed in this forum. This suggests that their clinical areas of interest align within the topics they state that they wish to discuss or find valuable online. This also demonstrates the potential of combining analytic methods to assess how well such
alignment is occurring in an OSN, and to guide interventions to address major discrepancies.

5.5.2 Environmental Context

Although environmental context is critical to the design and management of learning in OSNs (Cleary et al., 2013), it was possible to collect only limited environmental data (i.e. time of access) from the dataset available. None of context data relating to learner activity, relations, recognition, and application were collected by the operator of this forum.

Time. By analysing the interaction history of the participants, we identified the time of day the participants go online, revealing their self-directed learning schedules. As shown in Figure 5.1, evening was found to be more common than morning or afternoon; some participants were online after midnight, or very early in the morning. This is consistent with GPs’ very busy schedules during normal office hours (Joyce et al., 2015).

![Figure 5.1. The activity level across 24 hours](image)

Location. Since evening was the most common learning time for the participants, and most are GPs, we can infer that a home office is likely to be their physical learning
locations. However, for those medical practitioners who don’t work during normal office hours (e.g. those who do shift work or work in hospitals or elsewhere as part of their portfolio career), their physical locations for learning online may not be in their home office but at a workplace. Future implementation of the forum may consider tracking the IP addresses of computers logged in to the OSN to obtain better understanding of the participants’ physical learning environments, so that appropriate learning content and activity can be suggested.

5.6 Conclusions and Implications
Informal learning in OSNs is thoroughly influenced by and located in the context of the learners. This study proposed contextual factors that are critical for understanding learning context in OSNs for medical practitioners, and demonstrated how these contextual factors could be analysed, in the case of a small number of active users in an online discussion forum. This analysis produced non-stereotypical findings and demonstrated its usefulness as a method for gaining better insight into medical practitioners’ informal learning processes in OSNs.

To understand the context of learning in OSNs for medical practitioners, we suggest considering two types of contextual factors – learner factors and environmental factors. Learner factors include key demographics, work and educational background, and clinical areas of interest. Environmental factors include time, location, activity, relations, recognition and application.

Based on these contextual factors, we identified the typical characteristics of the active users who were engaged in learning in this forum. We found that those users were 55+ year-old men practicing in a major city of a populous State. They were highly experienced GPs with supervisory and management responsibility. Some worked part-time and had a portfolio career. They were graduates of the country where they practiced, and they held advanced qualifications. Their clinical areas of interest included general practice, dermatology, women’s health, chronic disease management, general medicine, diabetes, and sports medicine.

Based on limited environmental context data that is available in this forum, we found these active users were participating online in the evening more commonly than at other
times of day, and were likely to access the forum from their home office. More environmental context data should be systematically collected in this forum in the future including the data relating to learner activity, relations, recognition, and application. Therefore, future work should investigate approaches and methods to capture environmental context data. An useful approach is tracking of learner context via mobile technologies, such as geospatial location apps (Schäper & Thalmann, 2014).

In terms of the acquisition of learning context information, not all could be collected automatically through the technical system; some information relied on manual analysis of input by participants (e.g. personal profile details). This could change to take greater advantage of LA in future, if an OSN enabled participants to import relevant information about themselves from other online professional databases (e.g. LinkedIn or Twitter).

This study provided an understanding of the learning context of the active users in the forum, and thus contributes to developing more personalised and just-in-time learning for them. While the sample of our context data is too small to infer the general characteristics of the medical practitioners in OSNs, it shows a possible way for OSN operators to increase learner engagement by collecting and analysing the data about learning context. For instance, OSN operators could use this information to help the participants to make connections with others with similar practice profiles and invite them to relevant discussions to meet their practice needs.

So far, we have gained understandings of learning interaction, learning content and learning context in this forum, however, the validity of these finding about aspects of learning process, as well as the generalisability of these findings to medical practitioners in other OSNs, are unknown. In the next chapter, results of the survey study that was conducted to address these concerns are reported.
Chapter 6 Validating LA Findings about the Learning Process in an Online Social Network for Medical Practitioners

The relevant content of the following publication has been integrated into this chapter:


6.1 Introduction

Few previous studies have systematically investigated the learning process of medical practitioners in OSNs. To gain an understanding of the process of informal learning occurring in an online discussion forum, we carried out a series of studies to investigate the learning interaction, learning content and learning context using different Learning Analytics (LA) methods. We reported the findings in previous chapters: Chapter 3 described how Social Network Analysis (SNA) was used to analyse learning interaction; Chapter 4 described how topic modeling was used to analyse learning content; and Chapter 5 described how context analysis was used to understand learning context. In each of those chapters we developed a set of conclusions and inferences, based on the dataset and analytic method of each study.

This chapter aims to summarise these analytic findings, validate them via an online survey, and explore whether these findings may generalise to other medical practitioners who use other OSNs for informal learning. Section 6.2 presents a summary of the findings from previous analytic studies. Section 6.3 introduces the survey approach and describes how the survey questions were developed based on each study’s findings. Section 6.4 presents the survey results. Section 6.5 discusses and compares the survey results with the analytic findings. Section 6.6 draws conclusions about this study and discusses the implications of results.
6.2 Summary of Analytic Findings

Table 6.1 provides the summarised findings about the learning process from the LA studies conducted as part of this research. The detailed explanation for each of the individual findings was provided in the corresponding chapter of this thesis. According to the research design framework presented in Chapter 1 (in Figure 1.1), a detailed understanding of the learning process in the discussion forum was obtained by analysing the learning interaction, learning content and learning context. The findings of learning interaction helped understand how the learning might occur when the users interacted with each other. The findings of learning content provided the indirect evidence of learning responding to their interaction behaviours. The findings of learning context provided the background of their informal learning, which helped understand if their discussion content responded their learning needs and further interpreted the patterns of their learning interactions.

<table>
<thead>
<tr>
<th>Analytic finding</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Interaction</td>
<td></td>
</tr>
<tr>
<td>A large number of passive users.</td>
<td>Chapter 1</td>
</tr>
<tr>
<td>A small number of active users.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>An overall low level of interaction.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>A decreasing level of interaction over time.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>Discussions were dominated by moderators.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>Active users may easily interact with others.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>Small group learning occurred in the form of case discussion.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>Tendency to seek information.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>No pattern of interaction common to gender or location.</td>
<td>Chapter 3</td>
</tr>
<tr>
<td>Learning Content</td>
<td></td>
</tr>
<tr>
<td>Discussions of clinical topics are related to practices, sharing practical and experiential knowledge and providing benchmarks.</td>
<td>Chapter 4</td>
</tr>
<tr>
<td><strong>Analytic finding</strong></td>
<td><strong>Source</strong></td>
</tr>
<tr>
<td>----------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Discussions of non-clinical topics are related to the controversial topics, latest news and information in the field.</td>
<td>Chapter 4</td>
</tr>
<tr>
<td>Non-clinical topics were discussed more frequently than clinical topics.</td>
<td>Chapter 4</td>
</tr>
<tr>
<td><strong>Learning Context</strong></td>
<td></td>
</tr>
<tr>
<td>Males aged 55+ years-old, practicing in a major city of populated state in Australia.</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>Participating at home in the evening.</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>Highly experienced GPs with management responsibility. Some worked part-time and had portfolio career.</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>Graduated in Australia with a specialist qualification.</td>
<td>Chapter 5</td>
</tr>
<tr>
<td>Clinical areas of interest: general practice, dermatology, women’s health, chronic disease management, general medicine, diabetes, and sports medicine.</td>
<td>Chapter 5</td>
</tr>
</tbody>
</table>

### 6.3 Methods

#### 6.3.1 Survey Design

An online questionnaire was used to validate the findings about the learning process from prior analytic studies. It was based on the summarised findings in the above table and consisted of items measuring learning interaction, learning content and learning context. The survey used a combination of categorical, true/false, open-ended questions, and five-point Likert scale style response items. It had 30 items and took approximately five minutes to complete. Table 6.2 presents the survey items and the analytic findings with which they were associated. A copy of the survey is presented in Appendix D.

The survey consisted of two parts. The first part was a personal information section which included items relating to demographics, work experience, and educational background. These questions were aimed at validating the analytic findings about learning context. The second part consists of items validating the findings about the learning interaction and learning content. The survey asked the respondents to answer these questions based on an OSN they used most often. In this survey study, an OSN was defined as an online community platform that allows connecting, sharing interests,
and interacting with each other. Generic examples include but are not limited to: Discussion Forums, Facebook, LinkedIn, and Twitter.

**Table 6.2.** Mapping between survey questions and analytic findings

<table>
<thead>
<tr>
<th>Analytic finding</th>
<th>Corresponding survey question(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning Interaction</strong></td>
<td></td>
</tr>
<tr>
<td>A large number of passive users.</td>
<td>How often do you read information in this online social network? + How often do you ask questions in this online social network? + How often do you answer questions in this online social network? + How often do you contribute new information or ideas to this online social network? (Categorical options: Daily / A few times a week / Once a week / Once or twice per month / Never)</td>
</tr>
<tr>
<td>A small number of active users.</td>
<td></td>
</tr>
<tr>
<td>An overall low level of interaction.</td>
<td></td>
</tr>
<tr>
<td>A decreasing level of interaction over time.</td>
<td>Have your interactions in this online social network decreased over time? (Categorical options: Yes / No / Neither Yes nor No)</td>
</tr>
<tr>
<td>Discussions were dominated by moderators.</td>
<td>Are your discussions in this online social network facilitated? (Categorical options: All / Some / None / Don’t know)</td>
</tr>
<tr>
<td>Active users may easily interact with others.</td>
<td>The online social network allows me to easily share knowledge with others. (5-point scale ranging from ‘Strongly disagree’ to ‘Strongly agree’)</td>
</tr>
<tr>
<td>Small group learning occurred in the form of case discussion.</td>
<td>What is your main purpose in using this online social network? (Categorical options: Seek information / Build connections / Benchmark practice / Discuss specific cases / Other)</td>
</tr>
<tr>
<td>Tendency to seek information.</td>
<td></td>
</tr>
<tr>
<td><strong>Learning Content</strong></td>
<td></td>
</tr>
<tr>
<td>Discussion of clinical topics are related to practices, sharing practical and experiential knowledge and providing benchmarks</td>
<td>In your interactions in this online social network, please list up to three clinical topics that have been most important to your learning.</td>
</tr>
<tr>
<td>Discussion of non-clinical topics are related to the controversial topics, latest news and information in the field.</td>
<td>In your interactions in this online social network, please list up to three non-clinical topics that have been most important to your learning.</td>
</tr>
<tr>
<td>Analytic finding</td>
<td>Corresponding survey question(s)</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Non-clinical topics were discussed more frequently than clinical topics.</td>
<td>My interactions in this online social network are more often about Non-Clinical topics than Clinical topics. (Categorical options: Yes / No / Neither Yes nor No)</td>
</tr>
</tbody>
</table>

Learning Context

| Males aged 55+ years-old, practicing in a major city of populated state in Australia. | What is your gender? + What is your age? + Where is your practice located? + In what type of area is your practice located? (Categorical options: Major cities / Inner regional / Outer regional / Remote / Very remote / Don't know) |
| Participating at home in the evening.                                           | When do you usually use this online social network? (Categorical options: Morning / Afternoon / Evening / Night / No fixed time) + Where do you usually use this online social network? (Categorical options: Home / Workplace / Internet café / When I am mobile or moving from place to place / Other / No fixed location) |
| Highly experienced GPs with management responsibility. Some worked part-time and had portfolio career. | What is your current status? (Categorical options: working full-time / working part-time / Retired / Doctor in training / Doing unpaid work or voluntary activities / Other) + What is your main job title? + How many years have you worked in this role? + Do you have any other job? If yes, please provide details. |
| Graduated in Australia with a specialist qualification.                         | Was your first medical qualification completed in Australia? + What is your professional qualification? (Categorical options: FRACGP / ACRRM / FRACP / Other / None) |
| Clinical areas of interest: general practice, dermatology, women’s health, chronic disease management, general medicine, diabetes, and sports medicine. | What are your clinical areas of interest? |

The second part of the survey included additional questions relating to learning context (in Table 6.3). These questions were aimed to collect environmental context data that was not available in the discussion forum and included ‘Application’, ‘Relation’, and ‘Recognition’. It was expected that knowing this additional context information would
help to interpret these respondents’ learning experiences, and so enhance understanding of the learning process in OSNs.

Table 6.3. The survey questions relating to environmental context factors

<table>
<thead>
<tr>
<th>Environmental context factor</th>
<th>Corresponding survey question(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>The online social network has provided valuable information for my learning. + The online social network has contributed to my ability to address challenges in professional practice. + I can easily participate any discussion topic in the online social network. (5-point scale ranging from ‘Strongly disagree’ to ‘Strongly agree’)</td>
</tr>
<tr>
<td>Relation</td>
<td>How many different people have you interacted with in this online social network since you started using it? + Of those people you have interacted with, how many did you interact outside of this online social network? (e.g. in your workplace)</td>
</tr>
<tr>
<td>Recognition</td>
<td>It would be valuable to me to have my participation in this online social network accredited towards my Continuing Professional Development (CPD) points. (5-point scale ranging from ‘Strongly disagree’ to ‘Strongly agree’)</td>
</tr>
</tbody>
</table>

6.3.2 Survey Participants

Participation in the survey was open to any medical practitioners who used OSNs for informal learning in a manner similar to the case study, not limited to users of the specific OSN used in the case study. Thus cross-validation was done, involving a group not analysed in the initial studies. In the event, participants were recruited opportunistically, via the e-mailing distribution list of an Australian-based online health CPD provider that expressed interest in the research. The online health CPD provider is a health professional education company that provides education and technology solutions to medical practitioners, organisations and educators. The e-mail list includes 3,233 registered medical practitioners who have joined this community to take part in online courses and webinars for CPD.

It should be mentioned that the initial intent of the thesis was to survey the 48 active users from the forum that served as the basis for the case study. However, cross-validation was determined to be a stronger method. Given that the case study data were
generated between the years 2011 and 2014, it was very likely that those users might not recall what they did that far in the past. In addition, they may have used other OSNs and may not have been able to distinguish their experiences in the case study forum from that in other OSNs they accessed.

### 6.3.3 Procedure

Before distributing the survey, the survey was piloted with three medical practitioners who were experienced in using OSNs. Their feedback was incorporated into the last version of the survey. Following Human Research Ethics approval\(^9\), the survey took place in late 2016 and early 2017. Data were collected using REDCap software (https://www.project-redcap.org). An invitation to the study with a link to an online questionnaire was emailed by the online health CPD provider to its participants. The survey was open for completion for a total of four weeks. No incentive for completion was offered, and no reminder email was sent. The questionnaires were anonymous and completing a questionnaire was on a voluntary basis. Consent by respondents was assumed if the surveys were returned.

To analyse the survey data, the raw data were exported from REDCap via its automated export procedure to Excel, and were analysed using standard descriptive statistics to show trends in the data. Positive (i.e. ‘agree’ and ‘strongly agree’) and negative (i.e. ‘disagree’ and ‘strongly disagree’) responses were combined for the purpose of analysis.

Once the survey data were analysed, the findings were summarised and transformed thematically using a narrative approach (Fetters, Curry, & Creswell, 2013). Each of the survey-based findings was then compared with the corresponding analytic findings from the prior studies conducted during this research.

### 6.4 Survey Results

Of the 3,233 medical practitioners who received the invitation email, 254 logged on to the survey and 191 completed it, yielding a response rate of 6%. The database was subsequently cleaned of 29 records missing critical answers (i.e. those not answering

---

\(^9\) This study has been approved by the Human Research Ethics Committee of the University of Melbourne (HREC Reference #1648003).
questions relating to interaction, topics, or learning experience). In addition, 13 records were removed from respondents who said that they accessed only formal online learning (e.g. doing formal CPD activity in a Learning Management System). The final sample consisted of 149 valid responses. The following sections describe the sample characteristics, the types of OSN being used most often for informal learning, and the interactions, topics, and experiences of the respondents in these OSNs.

### 6.4.1 Sample Characteristics

The characteristics of the 149 respondents are summarised in Table 6.4. As shown, gender was evenly split. Most respondents (63%) were aged between 35 and 54 years. 86% of them were drawn from the most populated Australian States (i.e. Queensland, New South Wales, and Victoria), and identified as working in a non-rural setting. In terms of their educational background, 62% of them graduated in Australia, and 66% were Fellows of the Royal Australian College of General Practitioners (RACGP). GPs accounted for 95% of respondents and the remainder were specialists. Most (68%) had 10 to 19 years’ work experience in General Practice. Sixty-four percent worked full-time, 31% worked part-time, and the remainder were in training. We also identified 24% of the respondents as having portfolio careers, that is, multiple professional roles. For example, they were also doing remote locums, working as a medical officer at one or more hospitals, teaching or lecturing at a university, or working in a local hospital Emergency Department.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>79</td>
<td>52.8%</td>
</tr>
<tr>
<td>Male</td>
<td>70</td>
<td>47.2%</td>
</tr>
<tr>
<td>Age group (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25-34</td>
<td>15</td>
<td>10.2%</td>
</tr>
<tr>
<td>35-44</td>
<td>47</td>
<td>31.3%</td>
</tr>
<tr>
<td>45-54</td>
<td>47</td>
<td>31.3%</td>
</tr>
<tr>
<td>Characteristics</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>----------------------------</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>55-64</td>
<td>25</td>
<td>17.0%</td>
</tr>
<tr>
<td>65-74</td>
<td>14</td>
<td>9.4%</td>
</tr>
<tr>
<td>75+</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td>State</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australian Capital Territory</td>
<td>2</td>
<td>1.6%</td>
</tr>
<tr>
<td>New South Wales</td>
<td>45</td>
<td>30.5%</td>
</tr>
<tr>
<td>Queensland</td>
<td>48</td>
<td>32.0%</td>
</tr>
<tr>
<td>South Australia</td>
<td>6</td>
<td>3.9%</td>
</tr>
<tr>
<td>Tasmania</td>
<td>6</td>
<td>3.9%</td>
</tr>
<tr>
<td>Victoria</td>
<td>36</td>
<td>24.2%</td>
</tr>
<tr>
<td>Western Australia</td>
<td>5</td>
<td>3.1%</td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td>Location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major cities</td>
<td>62</td>
<td>41.3%</td>
</tr>
<tr>
<td>Inner regional</td>
<td>31</td>
<td>20.6%</td>
</tr>
<tr>
<td>Outer regional</td>
<td>50</td>
<td>33.3%</td>
</tr>
<tr>
<td>Remote</td>
<td>6</td>
<td>4.0%</td>
</tr>
<tr>
<td>Very remote</td>
<td>0</td>
<td>0.0%</td>
</tr>
<tr>
<td>Don't know</td>
<td>1</td>
<td>0.8%</td>
</tr>
<tr>
<td>Place of graduation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>93</td>
<td>62.1%</td>
</tr>
<tr>
<td>Outside Australia</td>
<td>56</td>
<td>37.9%</td>
</tr>
<tr>
<td>Professional qualification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRACGP</td>
<td>98</td>
<td>66.1%</td>
</tr>
<tr>
<td>FACRRM</td>
<td>3</td>
<td>2.3%</td>
</tr>
<tr>
<td>FRACP</td>
<td>3</td>
<td>2.1%</td>
</tr>
</tbody>
</table>
6.4.2 Types of OSN Being Used for Informal Learning

Figure 6.1 presents the types of OSN being used most often by the respondents for informal learning. As shown, the OSN that was most often named was GPDU, a private Facebook group that operates as an online community for all GPs working in Australia and New Zealand, followed by LinkedIn, ShareGP, and Twitter. Other OSNs included Figure1, Medcast, UpToDate, and AMA forum.
Facebook, LinkedIn and Twitter are widely known OSNs, and cater to a wide array of professions and markets including healthcare. These OSNs offer ease of access compared to online communities dedicated to medical practitioners, which may be the reason why they were frequently used. ShareGP is a private GP community exclusive to RACGP members. The large RACGP membership among participants explains the popularity of ShareGP in the survey results.

6.4.3 Learning Interaction

Based on the respondents’ self-reported results of their online participation, we found an overall low level of interaction, with a small number of active users in OSNs. In this study, an active user was defined as a user who participated in an OSN daily or a few times a week. As shown in Figure 6.2, although all users read information, most users (approximately 65%) never participated in any other way or rarely did so (once or twice a month). Only a small number of respondents (15%) reported themselves as active users in OSNs. Twenty-two percent of respondents were identified as passive users, based on the fact that their only response to the four interaction items was that they read information in the OSNs. However, it is possible that most of the respondents were active participants in OSNs in other ways not asked about.
The patterns of interaction among those who ‘ask questions’, ‘answer questions’, and ‘contribute new information’ were similar. Examining the individual responses, we found that: 1) those who ask questions tend to answer questions and contribute new information; 2) those who do not ask questions do answer questions; 3) participants who do not answer questions or contribute new information generally do not ask questions. Most of these participants (72%) interacted with only a small number of people (< 10) in OSNs, and they generally interacted with people whom they had met in OSNs rather than anywhere else.

There was no consistent difference in their interaction over time. Thirty-two percent reported having increased interaction over time, 36% reported having decreased interaction and 32% believed their interaction neither increased nor decreased over time.

Most (60%) participants used OSNs to seek information, 14% engaged in case discussions, 11% were preparing for exams, and the remainder were benchmarking their practices or building connections. When the respondents were asked if the OSN was moderated, 66% indicated that the OSN was moderated, 16% believed that the OSN was not moderated, and 18% were unsure if the OSN was moderated.

In terms of their regular time and location in accessing OSNs for informal learning, 36% of the respondents were active in the evening, whereas 48% had no fixed time. Home
access accounted for 85% of the population, while others were mobile or accessing from a workplace.

6.4.4 Learning Interest and Topics

The respondents expressed having clinical areas of interest in dermatology, women’s health, chronic disease management, general medicine, diabetes, sports medicine, mental health, and medical education.

The respondents considered many topics discussed in OSNs were important to their informal learning, including both clinical and non-clinical topics. The clinical topics were quite diverse; 39 specialised topics were identified in total with 80% mentioned by only one or two respondents. Table 6.5 lists the topics that were mentioned most often (> 10 times).

Table 6.5. The clinical topics of discussion

<table>
<thead>
<tr>
<th>Clinical topics</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dermatology</td>
<td>24</td>
<td>16.1%</td>
</tr>
<tr>
<td>Mental health</td>
<td>14</td>
<td>9.4%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>12</td>
<td>8.1%</td>
</tr>
<tr>
<td>Chronic disease management</td>
<td>10</td>
<td>6.7%</td>
</tr>
<tr>
<td>Women’s health</td>
<td>9</td>
<td>6.0%</td>
</tr>
<tr>
<td>Addiction medicine</td>
<td>8</td>
<td>5.4%</td>
</tr>
<tr>
<td>Cardiology</td>
<td>9</td>
<td>6.0%</td>
</tr>
<tr>
<td>Emergency medicine</td>
<td>8</td>
<td>5.4%</td>
</tr>
<tr>
<td>Pain management</td>
<td>8</td>
<td>5.4%</td>
</tr>
</tbody>
</table>
Compared to clinical topics, non-clinical topics were more focused; 12 non-clinical topics were identified and summarised in Table 6.6.

**Table 6.6. The non-clinical topics of discussion**

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>%</th>
<th>Non-clinical topics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leisure</td>
<td>17</td>
<td>11.4%</td>
<td>Travel, movie, yoga, food, meditation, music, animal, church, fashion, gathering, nutrition</td>
</tr>
<tr>
<td>Policy</td>
<td>16</td>
<td>10.7%</td>
<td>Guidelines, Medicare, billing</td>
</tr>
<tr>
<td>Education and training</td>
<td>15</td>
<td>10.1%</td>
<td>Study, exam preparation, support and tips, pathways</td>
</tr>
<tr>
<td>Patient communication</td>
<td>15</td>
<td>10.1%</td>
<td>Handle difficult patients, satisfaction and complaints, empowerment, attitude on recalls</td>
</tr>
<tr>
<td>News</td>
<td>12</td>
<td>8.1%</td>
<td>College, funding, health systems, international practice, new medication, RACGP updates</td>
</tr>
<tr>
<td>Work experience</td>
<td>12</td>
<td>8.1%</td>
<td>Female GP, clinical data summarisation, frustration, doctor health, professional standards</td>
</tr>
<tr>
<td>Political issues</td>
<td>11</td>
<td>7.4%</td>
<td>Economic, international, war torn areas</td>
</tr>
<tr>
<td>Social media use</td>
<td>11</td>
<td>7.4%</td>
<td>General computer skills, online interaction, ethics</td>
</tr>
<tr>
<td>Workload management</td>
<td>12</td>
<td>8.1%</td>
<td>Leave arrangement for registrars, manage burnout, safe work, work availability, workplace stress</td>
</tr>
<tr>
<td>Peer support</td>
<td>10</td>
<td>6.7%</td>
<td>Benchmark, clinic networks, interact with other health workers</td>
</tr>
<tr>
<td>Family</td>
<td>8</td>
<td>5.4%</td>
<td>Commitment, problem, situation</td>
</tr>
<tr>
<td>Running a practice</td>
<td>8</td>
<td>5.4%</td>
<td>Practice design, ideas, setting up, financial</td>
</tr>
</tbody>
</table>

Slightly more respondents believed that clinical topics were more frequently discussed than non-clinical topics in OSNs, which marginally indicates that OSNs are used more for professional rather than leisure or social discussions.

**6.4.5 Learning Experience**

The respondents’ learning experience in OSNs was positive overall. As shown in Figure 6.3, most respondents (approximately 70%) agreed that they could participate easily in
any discussion and share knowledge in OSNs, although approximately 20% disagreed with this statement. By inspecting their individual responses, we found those 20% were passive users and inactive users (who never asked or answered any question nor contributed new information, or who did so only once or twice a month).

![Bar chart showing responses to various statements about OSNs](image)

**Figure 6.3. The learning experience in OSNs**

Sixty-five percent of respondents (27% of whom were passive users) believed that OSNs provided valuable information for their learning and contributed to solving challenges in their practice. This may indicate that OSNs were useful for their informal learning – they could search and obtain just-in-time information in OSNs and then apply it in practice to solve challenges. However, overall approximately 25% gave a neutral response, perhaps indicating that they not have found OSNs beneficial to their learning.

Regarding their attitudes on accrediting the interaction in OSNs for CPD purposes, 64% of respondents agreed with this approach, indicating that providing CPD credits for activities in OSNs may lead to its increased acceptance (Millar et al., 2016). However, 19% were not sure, and 17% disagreed. They were keen to “keep informal learning informal”, in the words of two respondents.

The participants were given an opportunity to provide general comments about their learning experience of using OSNs. They pointed out that the learning content in current OSNs is one of their major concerns. Two participants suggested that the online content should be made easier to translate to practice with the use of plain language. Three
participants wanted to filter the content so as to be able to cover large amounts of information in minimum timeframes. Three participants commented that the content needs to be targeted to specific aspects of their practice. Other comments concerned the usability and accessibility of technology, protecting patient confidentiality, and providing support to build useful connections.

6.5 Comparison of the Findings about Learning Process

This section compares the survey results reported above with the findings about learning interaction, learning content and learning context obtained from the analytic studies. It also discusses potential reasons for each instance where there were inconsistencies.

6.5.1 Learning Interaction

Table 6.7 provides the comparison of findings about learning interaction. In the table, we number each comparison, identify the analytic finding that corresponds to a given survey-based finding, and assess its consistency. As shown, the survey-based findings about learning interaction were mostly consistent with the analytic findings except the comparisons #2 (number of passive users) and #4 (pattern of decreased interaction). The survey-based finding in comparison #2 (number of passive users) suggests only a few passive users but this may be biased as it could be that people who were passive users of OSNs may not have been interested in responding to the survey either. If that was the case it could account for why only a small number of passive users were included in the total survey respondents. The survey-based finding in comparison #4 (pattern of decreased interaction) shows no consistent difference in participants’ interactions over time. This may be dependent on OSNs and/or when an individual’s circumstances change, so effectively this survey finding may not validate the corresponding analytic findings.
Table 6.7. Comparison of findings about learning interaction

<table>
<thead>
<tr>
<th>ID</th>
<th>Analytic findings</th>
<th>Survey Findings</th>
<th>Consistent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A small number of active users.</td>
<td>15% were active users who participated daily or a few times a week</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>A large number of passive users.</td>
<td>22% were passive users.</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>An overall low level of interaction.</td>
<td>15% participated daily or a few times a week, 20% participated once a week, 65%</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>never participated or rarely (once or twice a month).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A decreasing level of interaction over time.</td>
<td>32% had increased interaction, 36% had decreased interaction, 32% had neither</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>deceased nor increased interaction.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Discussions were dominated by moderators.</td>
<td>66% participated in a moderated OSN, 16% believed that the OSN was not moderated,</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>18% were unsure if the OSN was moderated.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Active users may easily interact with others.</td>
<td>Active users all agreed that they could easily share knowledge</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>There was little small group learning occurring in the form of case discussion.</td>
<td>14% were discussing cases</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>Tendency to seek information.</td>
<td>60% seeking information, 14% discussing cases, 8% benchmarking, 11% preparing</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>exam, 7% building connections</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.5.2 Learning Content

Table 6.8 provides the comparison of findings about learning content. As shown, the survey-based finding in comparison #1 (clinical topics) is not consistent with its corresponding analytic finding, which may suggest that the clinical topics identified in the case study forum through topic modeling techniques are not representative of other OSNs for medical practitioners. Indeed, different OSNs may have different clinical areas of focus and target user groups. The survey-based finding in comparison #2 (non-
clinical topics) is consistent to its corresponding analytic findings, which may suggest that medical practitioners in general are interested in similar non-clinical topics. The survey-based finding in comparison #3 (discussion frequency between clinical and non-clinical topics) suggests that slightly more clinical topics are frequently discussed than non-clinical topics, which is inconsistent with the corresponding analytic findings. The most plausible explanation for this inconsistency is that a small group of survey participants reported that they were using OSNs specifically to prepare for exams.

Table 6.8. Comparison of findings about learning content

<table>
<thead>
<tr>
<th>ID</th>
<th>Analytic Findings</th>
<th>Survey Findings</th>
<th>Consistent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Discussion of clinical topics includes: palliative care, rheumatology, evidence-based medicine, statins use, vitamin use, vaccines, women’s health, fibromyalgia.</td>
<td>Clinical topics most valued in discussion include: dermatology, mental health, diabetes, chronic disease management, women's health, addiction medicine, cardiology, emergency medicine, pain management.</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Discussion of non-clinical topics includes: patient fees, training, prescriptions, policy, and workload.</td>
<td>Non-clinical topics most valued in discussion include: leisure, policy, education and training, patient communication, news, work experience, political issues, social media use, workload management, peer support, family, running practice.</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Non-clinical topics were discussed more frequently than clinical topics.</td>
<td>44% disagreed that non-clinical topics were more discussed than clinical topics. 36% agreed that non-clinical topics were more frequently discussed than clinical topics. 20% neither agreed nor disagreed that non-clinical topics were more discussed than clinical topics.</td>
<td>No</td>
</tr>
</tbody>
</table>

6.5.3 Learning Context

Table 6.9 provides the comparison of findings about learning context. As shown, the survey-based finding in comparison #1 (age and practice location) is not consistent with its corresponding analytic finding, which may suggest that the demographics vary
across different OSNs. The survey-based findings in comparisons #2 (participation time and location) and #5 (clinical areas of interest) are consistent with their corresponding analytic findings, suggesting that the learning preferences (includes time, location and interests) of medical practitioners in OSNs are similar. The survey-based finding in comparison #3 (work experience) is not consistent with its corresponding analytic finding, suggesting that the participants from different OSNs may have different work experience and background. The survey-based finding in comparison #4 (educational background) is consistent with its corresponding analytic finding, suggesting that the educational backgrounds of the medical practitioners in OSNs are similar. However, this is less likely to be by chance and more likely to be predetermined by the OSN hosting arrangements and membership conditions; in this case it is an artefact of the choice of a validation group closely resembling the case study group.

**Table 6.9.** Comparison of findings about learning context

<table>
<thead>
<tr>
<th>ID</th>
<th>Analytic Findings</th>
<th>Survey Findings</th>
<th>Consistent?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Males aged 55+ years-old, practicing in a major city of populated state in Australia.</td>
<td>Only 27% were aged 55+ years-old. Practicing in a major city (41%), and inner regional (21%), and outer regional (33%) Australia.</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>Participating at home in the evening.</td>
<td>Participating in the evening (36%) + at no fixed time (48%) at home (85%)</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Highly experienced GPs with management responsibility. Some worked part-time and had portfolio career.</td>
<td>GPs with work experience &lt;10 years (34%), 10-19 years (35%), 31% work part-time, 24% had portfolio careers.</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>Graduated in Australia with a specialist qualification.</td>
<td>Graduated in Australia (62%) with specialist qualification FRACGP (66%).</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Clinical areas of interest: general practice, dermatology, women’s health, chronic disease management, general medicine, diabetes, and sports medicine.</td>
<td>Dermatology, women’s health, chronic disease management, general medicine, diabetes, sports medicine, mental health, medical education</td>
<td>Yes</td>
</tr>
</tbody>
</table>
6.6 Conclusions and Implications

Validation of findings from LA was sought through a cross-sectional survey with a similar group of medical practitioners who used OSNs for informal learning. Due to the limitations of the survey method, the result has validated the analytic findings about some but not all components of the learning process of medical practitioners in the case study forum.

Consistent with the analytic findings, the validation results have shown that most medical practitioners were not active in discussions in OSNs, and that they preferred to seek information in moderated OSNs. This implies that OSNs can be used to support medical practitioners to learn and interact with each other, which would allow practical medical knowledge to be shared and applied. However, we should provide support for their information seeking and help them to efficiently identify specialised resources for enabling their learning.

The validation results have also revealed that both clinical and non-clinical topics were presented in OSNs: non-clinical topics of interest were similar among medical practitioners, however clinical topics of interest appear to vary across different OSNs since different OSNs may have different clinical areas of focus and target user groups.

Aspects of the context of learning are markedly different in different OSNs (in particular learner characteristics). This suggests that learning context needs to be given the greatest consideration by people who wish to understand and support learning interaction and learning content in OSNs. We need to gain a clearer understanding of learning context to be able to better support the informal learning that occurs in OSNs for medical practitioners. Informal learning is significantly contextual (Kelly & Hager, 2015) and thus participation in it is different from formal learning, where usually there are clearly defined learning activities and goals targeting specific groups of learners.

More broadly, the findings of this study support the use of analytic methods to understand the learning processes of medical practitioners in OSNs. The overall detail and consistency of the findings of this study and the analytic studies supports the hypothesis that proposed LA methods (i.e. SNA, topic modeling and context analysis) are useful to gain an understanding of the learning process of medical practitioners in an
OSN. This set of methods can reveal what is occurring in one OSN, and also the similarities and differences among OSNs. It has the potential to provide summary reporting both at the level of the individual learner, and at the level of the network as a whole.

SNA can be used to identify the patterns of interaction and thus can contribute to understanding what medical practitioners are trying to achieve in an OSN. SNA can also be used to make inferences about participants’ learning experiences. Topic modeling can be used to quickly discover the learning topics discussed in a specific OSN, so as to discern their learning needs. Context analysis is useful to understand the mosaic of learning factors that characterises a specific OSN, so as to enhance the understanding of learning interaction and content there. However, given that only limited context data is currently collected in many OSNs, the extent of analysis may be limited and needs to involve manual collection of external context data.

Using a survey for validating the analytic findings has its own limitations. One is the inability to identify the exact number of passive users from the target survey population. In addition, the validation is based on self-reported data which may not be a true representation of the learning process. Lastly, the validation data is bound to be somewhat more removed from the case study than if we were able to survey the very same forum participants whose learning processes we had previously analysed.

Future work to address these limitations may involve further testing of the analytic findings (especially those that contradict the survey results) with other medical practitioners. Although it is not feasible to do so within the scope of this thesis, replicating the analytic methods on other similar datasets is ultimately the way to validate or disprove the findings reported here.

In the next chapter, we will discuss the key contributions of this thesis, which include the analytical framework proposed for understanding of medical practitioners’ informal learning in OSNs, and the intervention strategies proposed for improving their informal learning based on the findings identified in this research.
Chapter 7 Discussion and Conclusion

7.1 Introduction

This research arises from the need for understanding informal learning occurring in OSNs for medical practitioners. While OSNs are used increasingly by medical practitioners to share knowledge and discuss practice management challenges, there is a lack of understanding about how learning occurs in OSNs, making it difficult to design and facilitate this type of learning (Institute of Medicine, 2010).

Given known difficulties in applying traditional analysis methods to investigate informal learning in OSNs for medical practitioners, Learning Analytics (LA) has been considered as a potential approach to deal with this challenge (Sandars et al., 2014).

Thus, the aim of this thesis was to investigate how LA can be used as a methodology for understanding the process of informal learning in OSNs for medical practitioners. Based on a multi-method design, a set of LA methods was employed to investigate different components of the learning process (i.e. learning interaction, learning content and learning context) in a case study OSN. Chapter 3 illustrated how Social Network Analysis (SNA) can be used to understand learning interaction; Chapter 4 illustrated how topic modeling can be used to understand learning content; Chapter 5 illustrated how context analysis can be used to understand learning context; and Chapter 6 summarised the findings about the learning process from all these analytic studies, and it presented findings from a further study that was conducted to validate the previous findings.

This chapter discusses the key contributions of this thesis. Section 7.2 discusses the analytical framework for investigating informal learning in OSNs proposed in this thesis including how each of the analytic methods addressed the original objectives of the research. Section 7.3 presents intervention strategies to improve informal learning of medical practitioners in OSNs, that are based on the findings about the case study. Section 7.4 outlines opportunities for future work in terms of improving medical practitioners’ informal learning in OSNs. Finally, Section 7.5 discusses how LAs may assist better use of OSNs in future to support medical practitioners’ informal learning.
7.2 Analytical Framework for Analysing Medical Practitioners’ Informal Learning in OSNs

The literature review in Chapter 2 identified the lack of an analytical framework within which to make sense of the learning that occurs in OSNs for medical practitioners, and found that LA methods were considered a potentially useful basis for such a framework (Sandars et al., 2014). Thus, this thesis proposed an analytical framework by exploring a set of analytic methods including SNA, topic modeling, and context analysis to understand the process of learning in an online discussion forum for medical practitioners.

The De Laat and Schreurs (2013) multi-method research framework, combining SNA with content analysis and context analysis to produce a complete picture of the learning process in a social learning network, suggested a way forward for the research in this thesis. Their framework was applied to analyse teachers’ informal learning in the workplace, and the data collection was mostly manual and the method involved online surveys and interviews with the learners. Thus, it did not offer a scalable systematic approach for collecting data and analysing informal learning in multiple OSN settings. Such an approach was the goal of the research reported in this thesis.

In this thesis, we built on the De Laat and Schreurs (2013) multi-method research framework but proposed an alternative approach that worked with the raw data generated in an OSN, and provided strategies to use with these data to analyse different components of the learning process in that OSN (i.e. learning interaction, learning content and learning content). Table 7.1 summarises the proposed LA methods for analysing each of the components of the learning process and describes how each of the analytic methods can contribute to understanding the process of learning in an OSN. The proposed LA methods can serve as an analytical framework to be applied in any online informal learning setting to gain an understanding of the learning that has occurred.
Table 7.1. Proposed LA methods for analysing informal learning in OSNs

<table>
<thead>
<tr>
<th>Component</th>
<th>Method(s)</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning interaction</td>
<td>SNA individual-level measures</td>
<td>Methods for identifying the level of interaction for individuals</td>
</tr>
<tr>
<td></td>
<td>SNA network-level measures</td>
<td>Methods for identifying the level of overall interaction and the changes in the level of interaction</td>
</tr>
<tr>
<td></td>
<td>SNA visual analysis</td>
<td>Providing an overview of the patterns of interaction and insights into the changes in the patterns of interaction</td>
</tr>
<tr>
<td></td>
<td>Stochastic Actor-Oriented Models</td>
<td>Providing insights into the structural properties of interaction that explain interaction changes</td>
</tr>
<tr>
<td>Learning content</td>
<td>Topic modelling</td>
<td>Providing high-level insights into the content of interest to network participants</td>
</tr>
<tr>
<td>Learning context</td>
<td>Quantitative context analysis</td>
<td>Methods for identifying the context of learning including learner characteristics and learning environment</td>
</tr>
</tbody>
</table>

This thesis demonstrated the viability of these LA methods to investigate informal learning processes in OSNs for medical professionals, by leveraging the data collected from an online discussion forum for medical practitioners. The thesis showed that these methods could produce findings that were useful to gain an understanding of the learning process of those medical practitioners. This set of methods can reveal what has occurred in an OSN in an efficient manner, and so can enable the development of intervention strategies to improve that learning. It has the potential to provide summary reporting for this purpose both at the level of the individual learner, and at the level of the network as a whole. However, it is worth mentioning that none of methods applied are adequate alone to account for learning fully, but taken together they have been shown to have much more insightful than other existing methods to understand learning occurred in OSNs.

The following sub-sections summarise the overall contribution of the LA methods to the understanding of informal learning in the case study OSN.
7.2.1 Social Network Analysis

SNA was employed to understand learning-related interactions among medical practitioners in an online discussion forum. The analysis included using SNA individual-level measures, network-level measures, visual analysis, and network modelling techniques.

SNA individual-level measures, which included degree centrality, betweenness centrality, and closeness centrality, allowed us to identify the network position of individuals based on their mutual interaction; this further allowed us to deduce each one’s role in knowledge sharing and building processes. Using these measures, we mapped how individual users interacted with others in the network, and the role of the key users in the network (e.g. leaders, knowledge brokers or experts). We also identified the way that the moderators produce the bulk of discussion, and revealed that overall small number of users who were engaged in the network.

The SNA network-level measures, which included density, centralization, diameter and average path length, allowed us to understand the level of overall interaction in the forum. We found that the interaction in the forum was low in general and the network was highly centralized, which may imply that knowledge did not flow easily and the users found it difficult to collaborate. We detected the occurrence of small group learning through the measures of diameter and average path length. By applying the network-level measures to a set of longitudinal data, we further identified that over time there was a decreasing level of interaction and a change in the centralization status of the network.

SNA visual analysis is usually done by sociograms, which are visual representations of individuals and their relationships to others in a group. In this research, sociograms validated the findings about interaction from mathematical analysis. Visualization was especially useful to get an overview of the patterns of interaction, and provided further insights into the changes in patterns of interaction in the network. Through visualization, we clarified the potential reasons underlying the change of network centralization by observing the movements of the active users. For example, by observing the changes of
patterns of interaction over time, we noted a tendency for a network to become temporarily centralized if there is a change in the lead role.

The network modelling technique SAOMs are the statistical models for longitudinal networks that are used to evaluate associated structural network effects and individual characteristics that could explain these changes as features of the learning process (Snijders, 1996). As explained in Section 3.3.2.2, we used SAOMs to examine which micro structures might play a statistically significant role in interactions, to understand why an interaction changed. Specifically, the network effect transitivity and attribute effects homophily were considered. We found that the active users in the forum tended to seek information. We also found no pattern of interaction common to their gender or geographical location based on available user profile data.

In summary, the techniques of SNA used in this research were helpful to understand the interaction and engagement in the forum, as well as to get a sense of what the participating medical practitioners were trying to achieve there (e.g. seeking information) through structural analysis of the network. SNA was also useful to infer aspects of the learning experiences of these medical practitioners (e.g. their relationships with others) through understanding of their patterns of interaction.

7.2.2 Topic Modeling

Topic modeling was used to understand the learning content generated by the active users in the forum. Using topic modeling, we identified their topics of interest, which included both clinical and non-clinical topics. The clinical topics were related to their practices, sharing practical and experiential knowledge and providing benchmarks, and included ‘palliative care’, ‘rheumatology’, ‘evidence-based medicine’, ‘statins use’, ‘vitamin use’, ‘vaccines’, ‘women’s health check’, and ‘fibromyalgia’. The non-clinical topics were related to controversial topics (e.g. ‘policy’, ‘workload’ and ‘patient fees’), and news and information in the field (e.g. ‘training’).

The thesis thus showed that the topic modeling techniques tested in this research could be used to discover the learning topics discussed in a specific OSN, so as to discern the participants’ learning interests and learning needs.
7.2.3 Context Analysis

Quantitative context analysis was performed to understand the learning context of the active users in the forum. Context analysis was focused on investigating learner factors and environmental factors. Learner factors in our model included key demographics, work and educational background, and clinical areas of interest. Environmental factors in our model included time, location, activity, relations, recognition and application.

Based on these contextual factors, we identified typical characteristics of active users in the case study forum. This provided background to their informal learning activity, helped in understanding whether the forum’s discussion content (as determined by topic modeling) was responsive to their learning needs, and further supported interpreting the patterns of their learning interaction (as identified from SNA). However, the availability of environmental context data in the case study forum was rather limited. Potentially some context data (e.g. those relating to application, relations and recognition) could be elicited from participants and analysed by the operator of a forum fairly readily – using existing technologies, through surveys or via other user feedback data collection techniques. Therefore, a recommendation arising from this thesis is that context analysis using available data be given more weight in further research, so as to understand the learning context for each specific OSN. OSN operators should consider additional data collection techniques to gather missing context information that could be used to support the learning activities in an OSN.

7.3 Intervention Strategies for Improving Medical Practitioners’ Informal Learning in OSNs

Our analytic findings provided new evidence about medical practitioners’ use of OSNs for informal learning. However, the findings also indicated that current OSNs may not be configured to support this informal learning as fully as possible. OSNs may not support medical practitioners to seek information and benchmark practice efficiently, and they may not assist facilitators to scaffold learning and form learning groups insightfully. Therefore, more sophisticated services are required to better serve the needs of the learning and facilitation in OSNs for medical practitioners.
Since the ultimate goal of LA is improving learning and teaching, providing valid interventions based on analytic results is highly important (Ferguson, 2012). The findings from our case study have led us to formulate design strategies for interventions to improve OSNs as learning environments. We propose a list of strategies (in Table 7.2) aimed at helping improve and sustain the ongoing informal learning of medical practitioners in OSNs.

**Table 7.2. Proposed intervention strategies**

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Intervention</th>
<th>For</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeking information</td>
<td>Intervention 1. Organising topics of discussion</td>
<td>Learner</td>
</tr>
<tr>
<td>Benchmarking practice</td>
<td>Intervention 2. Recommending subject experts</td>
<td>Learner</td>
</tr>
<tr>
<td>Scaffolding learning</td>
<td>Intervention 3. Alerting special discussions</td>
<td>Facilitator</td>
</tr>
<tr>
<td>Small group learning</td>
<td>Intervention 4. Detecting learning groups</td>
<td>Facilitator</td>
</tr>
</tbody>
</table>

These intervention strategies are based on the theory of connectivism (Siemens, 2005). Connectivism has been particularly useful to interpret informal learning by medical practitioners in OSNs and guide the design of intervention strategies in this thesis. In the following sub-sections we elaborate these intervention strategies by illustrating how each of them was derived based on connectivist principles. We also describe proposed approaches to implement these interventions, and present relevant application scenarios.

**7.3.1 Intervention 1 – Organising Topics of Discussion**

*Requirement: Seeking Information*

Studies have been conducted (via survey and interviews) to understand how medical practitioners use OSNs (McGowan et al., 2012; Panahi et al., 2012; Usher, 2012). They found that medical practitioners were keen to share knowledge in OSNs because they could easily access their peers anywhere and anytime. Other studies (Casebeer et al., 2002; Singh et al., 2012) found medical practitioners were highly strategic online in seeking information to solve problems and build careers, and used OSNs as “Intensional Networks”, which were characterised by emergency, being used for a task at hand, and a sense of history, being based on known relationships and a shared experience of working on similar tasks.
We employed SAOMs statistical models in this research to investigate and understand the changes in interaction. By examining the network effect transitivity, we found that the active users tended to seek information. We inferred that learning was likely to occur when they sought information and subsequently were connected with the relevant content available in the network.

OSNs can be used to support medical practitioners to learn and interact with each other. These cumulative interactions could potentially serve as an information store that allows practical medical knowledge to be organised systematically to be shared and applied. However, current implementations of OSNs for medical practitioners mostly promote CoP for immediate knowledge sharing (Barnett et al., 2016), and provide little support for a knowledge management approach to seek information and get up-to-date know-how.

The principles of connectivism (Siemens, 2005) states that learning is a process of connecting specialised information sources. Therefore, to support medical practitioners’ information seeking needs and efficiently connect them to specialised resources for enabling their learning, the content of historical and trending discussions in OSNs could be organised and re-organised into specific themes, and visually presented to medical practitioners, in a way that allows them to easily locate information and drill down into the details of discussion if required.

An useful approach to identify the themes of topics is through an automated analysis of discussion content such as topic modeling (Portier et al., 2013). Another possible approach is to use semantic annotation (Cao et al., 2010). This method would invite users to annotate their content, so the themes of topics could be identified by analysing their personal annotations as well as those assigned by subject experts.

An alternative approach is automated semantic annotation that can be used to alleviate manual work by users. A large number of semantic annotation tools have been developed for the biomedical domain and have been applied in practice (Funk et al., 2014; Tseytlin et al., 2016). According to Tseytlin et al. (2016), who conducted a comprehensive empirical study that included state-of-the-art semantic annotators and compared them based on the execution time and standard annotation performance
metrics (precision, recall, F1-measure), the most widely used semantic annotation tools include cTAKES\textsuperscript{10}, MetaMap\textsuperscript{11}, ConceptMapper\textsuperscript{12}, and NOBLE Coder\textsuperscript{13}.

\textit{Application Scenario}

Dr A had to make decision about the management of an elderly patient with diabetes. After searching a few online resources, he was wondering if an online professional community he regularly visited might provide useful information. The online community had several forums for different clinical and non-clinical topics, however, from his entry view of this community it was not easy to identify a particular topic such as management of diabetes for elderly patients. To support Dr A to identify information on this specific topic, the system had the facility to organise the topics being discussed and visually present them to him. Thus, Dr A quickly located the information he was looking for.

\textbf{7.3.2 Intervention 2 – Recommending Subject Experts}

\textit{Requirement: Benchmarking Practice}

Information and opinion are essential for making sense of the complex problems that medical practitioners face in everyday clinical practice (Wallace & May, 2016). OSNs offer the opportunity for medical practitioners to link with a vast amount of information and opinion, and as well as to obtain validation of their actions. In this research, through the analysis of learning content (i.e. discussion posts), we identified that the OSN was used for practice validation. In particular, content analysis suggested that users validated their practices around women’s health checks, fibromyalgia, the use of statins, vaccines, and vitamins. Further, through analysis of interaction, we identified that there was no pattern of interaction associated with gender or specific geographical location, suggesting that medical practitioners had no preference of this kind when validating their practices. The survey results indicated that medical practitioners mostly interact in OSNs with people whom they do not know, suggesting that they were willing to interact with people whom they do not know when validating their practices.

\textsuperscript{10} http://ctakes.apache.org/

\textsuperscript{11} https://metamap.nlm.nih.gov/

\textsuperscript{12} https://uima.apache.org/sandbox.html#concept.mapper.annotator

\textsuperscript{13} http://noble-tools.dbmi.pitt.edu/
The connectivism theory emphasises the importance of finding new connections; the
learning process begins by establishing new connections between people and learning
resources (Siemens, 2005). Further, learning connection is required for benchmarking
practice and facilitating continual learning. Therefore, it is important to support medical
practitioners to build and manage their learning connections in a network so that they
can target their interactions to further exploit learning opportunities.

Couros (2009) suggests that a connectivist approach to instructional design promotes
building and managing learning connections in a network through the development of
Personal Learning Networks (PLNs). A PLN is understood as an intentionally created
network of people set up by an individual specifically for improving their learning
(Siemens, 2005). Tools and services for developing and managing one’s PLN, such as
services for recommending subject experts, are already available (Klamma, 2013). In
particular, recommending subject experts who can provide guidance to deal with
practical challenges in specialised clinical areas could contribute greatly to medical
practitioners’ interest in and engagement with OSNs, as it would allow for seamlessly
finding other professionals with whom they can build knowledge collaboratively, and
improve practices and solve challenges together.

Subject experts can be identified by tracking users’ contributions, in other words, based
on analytics (Anderson, Huttenlocher, Kleinberg, & Leskovec, 2012). For example, a
user could be acknowledged as a subject expert in a specialised area once they had
contributed a certain number of relevant discussion posts (e.g. ten or more). An
alternative might be to provide users with the ability to rate posts. So, when a
contributor had a certain number of posts rated by others (e.g. five or more), that
contributor could be recognised as a subject expert in the area. Yet another possibility is
that identifying subject experts could be done through peer endorsement (Bhatia et al.,
2015); the idea is similar to the functionality of LinkedIn where users are asked to
endorse each other for specific skills and expertise. In a medical professional’s
community, users could be requested to endorse each other’s clinical expertise or skills.
For example, a user could be recognised as a subject expert in a specialised area when
the user had been endorsed for a specialised skill by ten or more people.
Application Scenario

Dr B had a patient with very low vitamin D. She knew that a low dose of vitamin D may raise the risk of fractures, so she prescribed a replacement regimen to the patient. However, she was interested in hearing different opinions on this treatment in an online community she regularly visited. She was unsure whom to ask and who might be the most qualified to answer, so she described this case in an online post. The system had the facility to recommend and notify a few subject experts based on the topic of the case he described. Five subject experts replied to her posts with their opinions. She evaluated their responses and was convinced to follow a suggestion given by Dr C, who recommended a compounded formulation. Dr C thus became one of Dr A’s learning connections.

7.3.3 Intervention 3 – Alerting Special Discussions

Requirement: Scaffolding Learning

Previous studies have recognised that facilitators played an important role to draw all participants in and guide and support their learning activities in an online community for medical practitioners (Barnett et al., 2014). However, questions remain about how to facilitate learning in a specific learning setting and what kind and level of impact to expect from different facilitation approaches. In the research reported in this thesis, the analysis of learning interaction identified that most online discussions were led by facilitators, and there was an overall low level of interaction and a decreasing level of interaction over time. This suggests that a facilitator-led OSN may sustain interactivity only for a short time and in the early stage of the learning process. Over time, the network may stop functioning if the facilitators leave (Chung, Piraveenan, Levula, & Uddin, 2013). The theory of connectivism runs against this leadership-oriented way of facilitation as it offers little support to the learners for developing their abilities to recognise and assess connections. Therefore, a different approach to facilitation in OSNs needs to be considered.

Connectivism suggests that facilitators should play the role of coaching and mentoring – moving from facilitation of pre-determined discussion to ‘facilitated and supported enquiry’. This is seen as scaffolding (Vygotsky, 1978), which involves providing
assistance to learners on an as-needed basis and phasing out the assistance as learner competence increases (Salmon, 2004). Therefore, facilitators should support the learners to develop their ability to recognise and assess connections by scaffolding learning activities (e.g. online discussions initiated by enquirers). Scaffolding of learning activities enables learners to solve a problem or carry out a task that would be beyond what they could accomplish independently (Vygotsky, 1978). A similar approach is described as facilitation in the style of “a guide on the side” (Collison, 2000) to support learning and communication among the participants themselves, instead of one or two designated prompters leading and staying in the centre.

To enable facilitators to scaffold learning activities in an OSN, tools could be developed that serve as a support system by alerting them when a special discussion requires intervention. The facilitators are alerted for special discussions so that they can give timely support to learners to develop their ability to hold meaningful discussions. The facilitators do not only include formally appointed moderators but also include the recognised subject experts in an online community, so that learners in OSNs are given more opportunities to take ownership of their own learning.

Special discussions can be identified via analytics by tracking the data relating to the number of responses to discussions, in real time. For example, a discussion that received very little response could be considered as a special discussion that might require assistance in fostering the discussion; conversely one that attracted a great deal of attention might offer opportunities for critical input to discussion. Another possible approach to identify special discussions is through sentiment analysis of discussion content; that could provide indications of need for assistance by understanding learners’ emotions in their conversations about a particular topic.

*Application Scenario*

Dr D was a trained facilitator responsible for facilitating discussions in an online professional community for medical practitioners. As part of his role, he initiated many topics for discussion and tried to participate in as many discussions as he could. However, he found that the overall participation in the community was low and most conversations stopped when he left the discussion. He felt it was difficult to keep himself up-to-date on all discussions because of other administrative activities he had to
do in the community. To support Dr D to track discussions occurring in the community, the system had facility to send alerts to him on a weekly basis about those discussions that might benefit from his intervention. Based on other topics under discussion, the system also sent these alerts to identified subject experts in specialised areas, which allowed those subject experts to facilitate discussions on their specialised topics. This functionality in the OSN shared the responsibility and knowledge more widely through the community, and streamlined the workload of the facilitator Dr D.

7.3.4 Intervention 4 – Detecting Learning Groups

*Requirement: Small Group Learning*

Small group learning is one of the most effective learning strategies used in formal CPD to support on-going learning (Lindsay et al., 2016). It is considered more effective than other traditional formats, such as lectures, because it promotes self-directed learning and allows deeper learning (Williams, 2005). Small group learning in the form of case-based learning (CBL) is increasingly conducted online (Thistlethwaite et al., 2012), and is considered particularly useful for medical practitioners who are located in sites of clinical training and practice that are geographically separated (Mann & Sargeant, 2013). CBL is defined as a form of inquiry-based learning, that aims to prepare learners for clinical practice through the application of knowledge to authentic clinical cases (Thistlethwaite et al., 2012).

The research in this thesis found that CBL was supported through the online discussion forum, but was not sustainable. In the forum we studied, CBL was conducted by two trained facilitators, who prepared and presented cases on a fortnightly basis. Once a case was presented, any user could choose to engage in discussion of the case by giving their opinions about the management of the clinical problem. The case topics were specialised clinical topics and some attracted more attention than others. Conducting CBL in OSNs is seen as more challenging for facilitators; unlike in face-to-face CBL, online learning groups generally are not formed prior to the discussion of cases (Thistlethwaite et al., 2012). Further, in informal learning processes, learning goals and learning activities are not fixed (Klamma, 2013), so the topics of interest are constantly changing in OSNs, which makes it hard for CBL facilitators to identify topics that will stimulate the learners’ interest.
Small group learning could be used as an effective learning practice in OSNs, in line with the connectivist view that learning and knowledge rest in diversity of opinion. To effectively facilitate CBL in OSNs, facilitators need analytics-based tools to identify learning groups with similar interests as well as case topics that stimulate interest. The tools that help identify potential learning groups or communities could be developed by implementing community detection algorithms to analyse discussion content and interactions (Fortunato, 2010; Newman & Girvan, 2004). Another way to group relevant individuals for this purpose could be by applying analytics to individuals’ professional profiles (Kritikou et al., 2008).

**Application Scenario**

Dr E was a trained facilitator responsible for organising CBL in an online professional community for medical practitioners. She ran CBL fortnightly in the community to promote conversation. However, Dr E noted decreasing participation in the discussion of the cases she presented. To resolve this issue, she wanted to send invitations to some regular users in the community, but she was unsure who should be receiving invitations. To help Dr E overcome this challenge, the system had the facility to deliver potential learning interest groups (e.g. Mental Health, Cardiology) via email to her on a weekly basis. She received an email that included information about potential interest groups with a list of users in each group. She reviewed these interest groups and identified the major current topics of interest in the community. Based on this, she was able to propose more meaningful cases for more targeted groups and to invite specific users to participate in these case discussions. Participation was improved, with benefits for all participants; Dr E had better insights into the interest groups and experts in the community and could give more focus to preparing cases accordingly.

### 7.4 Opportunities for Future Work

After reviewing the details of our understanding of the informal learning in an OSN for medical practitioners, we can identify four main future directions towards using an analytics approach to improve medical practitioners’ informal learning in OSNs.

The first, is to investigate approaches and methods for effective context capturing and modeling in OSNs for medical practitioners. In this research, we analysed the very
limited context data that we could collect from the system. However, the context of learning is different in every OSN, and is critical to understand informal learning and support personalised interventions in OSNs (Bicans, 2015). Therefore, future work should investigate approaches and methods to capture learning context; this would include information related to personal profiles of medical practitioners, learning activities and objectives, as well as the environment in which they work and learn. Learning context has already been widely researched under the area of lifelong learner modeling (Kay & Kummerfeld, 2012). The six most popular and useful features in learner modeling include: learner’s knowledge, interests, goals, background, individual traits and context. Thus, Chatti, Brandt, and Schroeder (2015) proposed an updated version of this context model by incorporating learner interest information into the lifelong learner model. Further, tracking of learner context via mobile technologies, such as geospatial location apps, could also be a useful approach. The Learning Layers Project\(^{14}\) was a project about informal mobile learning in professional communities. Last but not least, MedBiquitous\(^{15}\) focuses on developing technology standards for healthcare education, proposing a ‘Professional Profile’ data standard in XML format that provides a common format for exchanging clinician contact, education, training, certification, and membership information. This could potentially help the exchange of learner context information, if the learner is given the ability to update their profile and import and manage their profile. All these approaches and methods show potential ways of collecting, managing and exchanging context information of medical practitioners in OSN settings, subject to careful evaluation of their potential applications and implications, for instance for privacy and security.

Second, there would be value in investigating the learning process in OSNs in more depth using other novel analytic methods. While the research in this thesis was being undertaken, other researchers were working towards the application of LA methods in online formal learning settings. One notable work is the large research project being jointly done by Srecko Joksimovic and Vitomir Kovanović at the University of Edinburgh to investigate learning processes within networked learning. They have employed a variety of analytic methods (including SNA, linguistic analysis, and others)

---

\(^{14}\) [http://www.learning-layers.eu](http://www.learning-layers.eu)

\(^{15}\) [https://www.medbiq.org](https://www.medbiq.org)
to assess effects of interaction and quality of discussions, based on analysing data collected in online Learning Management Systems (Kovanović et al., 2016; Skrypnyk, Joksimović, Kovanović, Gašević, & Dawson, 2015). Future research may evaluate the application of these methods in informal professional learning settings. For example, using discourse analysis to examine the quality of discussion, and sentiment analysis to understand participants’ opinions about different discussion topics, would provide more in-depth insight into networked and online learning.

Third, it is worthwhile to investigate the learning processes of passive learners to determine the reasons for their lack of engagement; examine the ways to motivate them to engage, as well as to determine the usefulness of OSNs to this segment of the medical profession. While developing interactive learning is important because of its known effectiveness, the value of passive learning should be acknowledged and adequate support should be provided to such learning activities. In this research, we identified a very large number of passive users (81%) in the case study forum. This finding is consistent with previous studies (Ikioda et al., 2013; Stewart & Abidi, 2013), and suggests that passive users tend to make up the majority of members within any online community for medical practitioners. Therefore, without insight into the learning processes of these passive users, our understanding of learning processes for medical professional in OSNs is incomplete.

Understanding of the learning behaviour of passive learners requires tracking their learning activities. However, none of the activities of the passive users were tracked in the case study forum in this thesis so there is no way of knowing their learning behaviours (e.g. what types of discussion they most visited and spent most time reading). If such data can be obtained, which is technically possible, further study could help to identify how they behave as learners and why and thus to understand what value they may get from the community through passive engagement.

Finally, based on the proposed intervention strategies, developing learning interventions (as proposed in Section 7.3) that provide different forms of facilitating and scaffolding learning to improve medical practitioners’ online learning experience would be useful and informative. Timely feedback based on data from LA approaches is desirable. Since the ultimate aim of LA is to support learning by making use of the data generated by
learners’ online activities, we need to improve ways to make the findings about the learning process visible at the most opportune times to different system users including learners, learning groups, and system administrators or operators, together with ensuing recommendations that spark and support learning.

7.5 Concluding Thoughts

There is more information available now than any one medical professional can manage. Medical practitioners must commit to continuous learning through CPD to maintain up-to-date knowledge and skills. OSNs present an opportunity for health CPD to take place in a less formally structured way than ever before, but their full benefit for learning is yet to be realised.

In this thesis, we have explored how medical practitioners used an online discussion forum for their informal learning. We found that most participants tended to search for information and to benchmark their professional practice. Although social network technologies like online discussion forums can be designed for information seeking and knowledge sharing, the full potential of these technologies for supporting medical practitioners’ learning and development may not be realised in current implementations.

Medical practitioners need to engage in lifelong learning. In our opinion, OSNs have the potential to advance their lifelong learning by being the ‘master’ of learning. An OSN can serve as the engine of learning connection, that prompts medical practitioners and connects them to various forms of learning activities or resources based on their personalised learning contexts. In other words, OSN-enabled informal learning may sit in the centre of a learning environment, in which many formal learning resources (e.g. online learning modules, broadcasts, webinars) may sit on the sides as added support for learners to fulfil their own learning needs. LA could play a critical role in OSNs to bring this about. In the following paragraph, we describe how LA-based interventions (in Section 7.3) may help medical practitioners to navigate to many forms of formal or informal learning activities or resources, and create the most ideal ways for learning to occur in an OSN.

When a doctor who is interested in mental health visits an online community to seek information or share knowledge, the software platform that supports the online
community will first record her/his profile, then present organised information (Intervention 1 – Organising topics of discussion), and track all the interactions. When an online webinar for mental health is scheduled, the community system will prompt this doctor about this upcoming event based on this interest. If the doctor attends the webinar, s/he will not only be awarded CPD credits for webinar attendance, but also will be informed that there is an interest group on mental health organised in the online community (Intervention 4 – Detecting learning groups). If the doctor joins the interest group for mental health, s/he will have opportunity to share experiences and opinions with recommended subject experts and learning peers (Intervention 2 – Recommending subject experts). Through informal online discussion with these others, this participant subsequently may identify a gap in his/her knowledge and be made aware of an online learning module on mental health (Intervention 3 – Alerting special discussions). If the doctor undertakes the online learning module, that activity will be awarded CPD credits and that knowledge gap may be filled.

When an OSN is designed in the way describe in this scenario, we believe that the OSN can better support medical practitioners’ learning and ultimately contribute to their CPD. It is possible that informal learning could be recognised (even be awarded CPD points in its own right) and integrated with formal learning approaches. Making a separation between informal and formal learning is not a fruitful way to think about professional learning, as Jarvis (2004)’s model of adult learning suggests. An important question is how to find the right balance between formality and informality in any given learning situation in a way that is responsive to learner needs and available resources. Medical practitioners’ formal CPD and informal lifelong learning can be supported well when driven by their practice needs, and we believe that OSNs can support such learning needs by providing personalised learning and discovering relevant learning opportunities across the Internet. We believe that more systematic and concerted use of standard analytic approaches to OSN activity will contribute to achieving this level of functionality in these OSNs in future.

This thesis has added new knowledge to the field of health informatics, specifically to refining the way that information technologies can be understood to enhance learning in the health professions. It has provided a novel way of explaining how medical
practitioners' informal learning occurs in OSNs. It has demonstrated how the capability of LA can be leveraged to interpret the expanding amount of data that are being produced in OSNs for medical practitioners. It thus has contributed to the development of a methodology that has an important place in an emerging form of technology – supported learning in medicine.
Bibliography


mixed-methods case study. *Journal of Medical Internet Research, 16*(3), e83.


Griffiths (Eds.), *Complexity and healthcare: An introduction* (pp. 149-166). London, UK: Radcliffe Medical Press.


Ikioda, F., Kendall, S., Brooks, F., De Liddo, A., & Buckingham Shum, S. (2013). Factors that influence healthcare professionals’ online interaction in a virtual
doi:10.4236/sn.2013.24017

Institute of Medicine. (2010). Redesigning continuing education in the health

development (CPD) for quality care: Context, mechanisms, outcome and impact.
England, UK: Canterbury Christ Church University.


Routledge.

Retrieved from https://www.learning-theories.com/


Australian doctors aged 65 years and older. Australian Health Review, 39(5),
582-587.

Lesgold (Eds.), Adaptive technologies for training and education (pp. 140-164).
Melbourne, Australia: Cambridge University Press.

influential discussions among physicians: A social network analysis of a primary
care practice. Journal of General Internal Medicine, 22(6), 794-798.

health care simulation. Clinical Simulation in Nursing, 11(8), 376-382.

of Psychiatry, 27(2), 124-132.


Klamma, R. (2013, October 6-9). Community learning analytics – challenges and
opportunities. Paper presented at the 12th International Conference on Web-
Based Learning, Kenting, Taiwan.


social networks: Theory, methods, and applications (pp. 9-14). Cambridge, UK: Cambridge University Press.


Schostak, J., Davis, M., Hanson, J., Schostak, J., Brown, T., Driscoll, P., . . . Jenkins, N.


### Appendix A. Online Social Networks for Health Professionals

<table>
<thead>
<tr>
<th>OSN name</th>
<th>Site url</th>
<th>Origin</th>
<th>Target user</th>
<th># Users</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMA forum</td>
<td><a href="https://www.ama.com.au">https://www.ama.com.au</a></td>
<td>Australia</td>
<td>Medical practitioners</td>
<td>N/A</td>
<td>Online community for registered medical practitioners and medical students</td>
</tr>
<tr>
<td>CME (LinkedIn)</td>
<td><a href="https://www.linkedin.com">https://www.linkedin.com</a></td>
<td>US</td>
<td>Health professionals</td>
<td>11k+</td>
<td>Discussing continuing medical education</td>
</tr>
<tr>
<td>Doctors.net.uk</td>
<td><a href="https://doctors.net.uk">https://doctors.net.uk</a></td>
<td>UK</td>
<td>Medical practitioners</td>
<td>222k+</td>
<td>Clinical problem solving and general case discussion</td>
</tr>
<tr>
<td>DoctorsHangout</td>
<td><a href="https://www.doctorshangout.com">https://www.doctorshangout.com</a></td>
<td>India</td>
<td>Medical practitioners</td>
<td>N/A</td>
<td>Medical knowledge sharing and building connections</td>
</tr>
<tr>
<td>Doximity</td>
<td><a href="https://www.doximity.com">https://www.doximity.com</a></td>
<td>US</td>
<td>Health professionals</td>
<td>500k+</td>
<td>Developing professional connections</td>
</tr>
<tr>
<td>Figure1</td>
<td><a href="https://www.figure1.com">https://www.figure1.com</a></td>
<td>US</td>
<td>Health professionals</td>
<td>500k+</td>
<td>Sharing and discussing medical cases with anonymous images</td>
</tr>
<tr>
<td>GPDU (Facebook)</td>
<td><a href="https://www.facebook.com/groups/gpsdownunder">https://www.facebook.com/groups/gpsdownunder</a></td>
<td>Australia</td>
<td>Medical practitioners</td>
<td>4k+</td>
<td>Sharing experience through case discussion</td>
</tr>
<tr>
<td>MedTalk (Google+)</td>
<td><a href="https://plus.google.com">https://plus.google.com</a></td>
<td>US</td>
<td>Medical practitioners</td>
<td>10k+</td>
<td>Helping international physicians connect with one another</td>
</tr>
<tr>
<td>Medical Observer</td>
<td><a href="https://www.medicalobserver.com.au">https://www.medicalobserver.com.au</a></td>
<td>Australia</td>
<td>Medical practitioners</td>
<td>N/A</td>
<td>Medical news, and practice-shaping opinions</td>
</tr>
<tr>
<td>OZeGP (Mailgroup)</td>
<td><a href="https://www.gp.org.au/Ozegpmailgroup.html">https://www.gp.org.au/Ozegpmailgroup.html</a></td>
<td>Australia</td>
<td>Medical practitioners</td>
<td>N/A</td>
<td>Sharing knowledge and experience</td>
</tr>
<tr>
<td>Ozmosis</td>
<td><a href="https://www.ozmosis.com.au">https://www.ozmosis.com.au</a></td>
<td>US</td>
<td>Medical practitioners</td>
<td>N/A</td>
<td>Exchange clinical, practice management and</td>
</tr>
<tr>
<td>OSN name</td>
<td>Site url</td>
<td>Origin</td>
<td>Target user</td>
<td># Users</td>
<td>Purpose</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------</td>
<td>--------</td>
<td>---------------------</td>
<td>---------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td>QuantiaMD</td>
<td><a href="https://www.quantiamd.com">https://www.quantiamd.com</a></td>
<td>US</td>
<td>Medical practitioners</td>
<td>200k+</td>
<td>Social learning and collaboration</td>
</tr>
<tr>
<td>Sermo</td>
<td><a href="https://www.sermo.com">https://www.sermo.com</a></td>
<td>US</td>
<td>Medical practitioners</td>
<td>550k+</td>
<td>Medical crowd-sourcing</td>
</tr>
<tr>
<td>ShareGP</td>
<td><a href="https://www.racgp.org.au/sharegp">https://www.racgp.org.au/sharegp</a></td>
<td>Australia</td>
<td>Medical practitioners</td>
<td>33k+</td>
<td>Collaboration, sharing and building connections</td>
</tr>
<tr>
<td>ThinkGP</td>
<td><a href="https://thinkgp.com.au">https://thinkgp.com.au</a></td>
<td>Australia</td>
<td>Health professionals</td>
<td>38k+</td>
<td>News and education for medical practitioners, including online learning module and forum</td>
</tr>
<tr>
<td>UpToDate</td>
<td><a href="https://www.uptodate.com">https://www.uptodate.com</a></td>
<td>US</td>
<td>Health professionals</td>
<td>6k+</td>
<td>Provide clinical decision support resources</td>
</tr>
</tbody>
</table>
Appendix B. Network Modeling Using RSiena

```r
# load RSiena library
require(RSiena)

# load interaction data (2012, 2013, and 2014) from csv files
user_2012 <- read.csv("data/2012.csv", header=FALSE, check.names=FALSE)
user_2013 <- read.csv("data/2013.csv", header=FALSE, check.names=FALSE)
user_2014 <- read.csv("data/2014.csv", header=FALSE, check.names=FALSE)

# convert dataframe to matrix
muser_2012 <- as.matrix(user_2012)
muser_2013 <- as.matrix(user_2013)
muser_2014 <- as.matrix(user_2014)

# load user gender, state(location) from csvs
state <- as.matrix(read.csv("data/state.csv", header=FALSE))
gender <- as.matrix(read.csv("data/gender.csv", header=FALSE))

# define RSiena data structures:
interactionData <- array(c(muser_2012, muser_2013, muser_2014),
                          dim = c(48, 48, 3))

# create objects for the dependent variable
interaction <- sienaDependent(interactionData)

# construct objects for the explanatory (independent) variables
state <- ccCovar(state[, 1])
gender <- ccCovar(gender[, 1])

# Combine dependent and independent variables to define the dataset.
dat.1 <- sienaDataCreate(interaction, gender, state)

# Obtain the basic effects object, create effects structure
eff.1 <- getEffects(dat.1)

# print to inspect basic effects objects
print01Report(dat.1, eff.1)
effectsDocumentation(eff.1)

# Specify model, define the effects to include in the model
eff.1 <- includeEffects(eff.1, transTriads, include=T)
eff.1 <- includeEffects(eff.1, nbrDist2, include=T)

# Include homophily effect for the constant covariate
eff.1 <- includeEffects(eff.1, sameX, interaction1 = "state")
eff.1 <- includeEffects(eff.1, sameX, interaction1 = "gender")

# Define the algorithm settings
myAlgorithm <- sienaAlgorithmCreate(projname = 'interaction')

# Estimate the model
ans <- siena07(myAlgorithm, data = dat.1, effects = eff.1, batch=TRUE, prevAns=ans)

# obtain estimation results
summary(ans)
siena.table(ans, type="html")
```

151
Appendix C. Topic Modeling Using MALLET

```r
# load mallet library
require(mallet)

# load post content from csv
forumtext <- read.csv(file="data/post.csv", stringsAsFactors = F)

documents <- data.frame(text = forumtext$text,
                         id = forumtext$id,
                         stringsAsFactors=FALSE)

mallet.instances <-
mallet.import(as.character(documents$id),as.character(documents$text))

## Create a topic object.
number.topics <- 20

topic.model <- MalletLDA(number.topics)

## Load documents

topic.model$loadDocuments(mallet.instances)

## Optimize hyperparameters every 20 iterations, after 50 burn-in iterations.
topic.model$setAlphaOptimization(20, 50)

## Train model - 800 iterations
topic.model$train(800)

## pick the best topic for each token
topic.model$maximize(10)

## Get the probability of topics in documents and the probability of words in topics.
document.topics <- mallet.doc.topics(topic.model, normalized=T)
topic.words <- mallet.topic.words(topic.model, normalized=T)

## normalize and transpose document topics
document.topics <- t(document.topics)
topic.documents <- topic.documents / rowSums(topic.documents)
write.csv(topic.documents, "topic-docs.csv")

## Get short names for the topics
topics.labels <- rep("", number.topics)
for (topic in 1:number.topics) topics.labels[topic] <-
paste(mallet.top.words(topic.model, topic.words[topic,],
                       num.top.words=10)$words, collapse=" ")

# return the top keywords for each topic
write.csv(topics.labels, "topics-labels.csv")
```
Appendix D. Online Questionnaire Used for Study 5

Confidential

Health Professionals' Online Social Network Use for Informal Learning

Welcome!

You have been invited to complete a short questionnaire about your use of online social networks for your informal learning.

Informal learning is unofficial, unscheduled, and impromptu. It does not count toward an educational qualification and is not done as part of a training program organised by your employer or professional association.

Online social networks refer to online community platforms that allow you to connect, share interests, and interact with each other. Generic examples include but are not limited to: Discussion Forum, Facebook, LinkedIn, and Twitter.

This study is being carried out at Health and Biomedical Informatics Research Centre, University of Melbourne. If you would like to read detailed information about this study, please click here to view the full participant information sheet.

The survey will take about 5 minutes to complete.

Your answers are important in helping shape the future of online social network for health professionals’ learning!

If you would like to participate, please click ‘Next Page’ to start the survey.

Please complete this survey.

What is your gender?
- Female
- Male

What is your age?
- 25-34
- 35-44
- 45-54
- 55-64
- 65-74
- 75+

Where is your practice located?
- Australian Capital Territory
- New South Wales
- Northern Territory
- Queensland
- South Australia
- Tasmania
- Victoria
- Western Australia
- Other

Other (please specify)

In what type of area is your practice located?
- Major cities
- Inner regional
- Outer regional
- Remote
- Very remote
- Don’t know

Was your first medical qualification completed in Australia?
- Yes
- No
What is your professional qualification? (Please select all that apply)
- FRACGP
- FA/FRM
- FRACP
- Other (please specify)
- NONE

Other (please specify)

What are your clinical areas of interest?

What is your current status? (Please select all that apply)
- Working full-time
- Working part-time
- Retired
- Doctor in training
- Doing unpaid work/Voluntary activities
- Other (please specify)

Other (please specify)

What is your main job title?

How many years have you worked in this role?
- < 10
- 10-19
- 20-29
- 30+

Do you have any other job? If yes, please provide details.

* Please list up to THREE online social networks by frequency that you have used for informal learning.

The following questions relate specifically to the online social network you have used most often for your informal learning (i.e. the FIRST online social network you listed above).

What is your main purpose in using this online social network?
- Seek information
- Build connections
- Benchmark practice
- Discuss specific cases
- Other (please specify)

Other (please specify)

How long have you been using this online social network?
- Less than 6 months
- 1-2 years
- 2-3 years
- 3+ years
How many different people have you interacted with in this online social network since you started using it?
- 0
- 1-2
- 2-5
- 5-10
- 10+

Of those people you have interacted with, how many did you interact outside of this online social network? (e.g. in your workplace)
- 0
- 1-2
- 2-5
- 5-10
- 10+

When do you usually use this online social network?
- Morning (6am-12pm)
- Afternoon (12pm-6pm)
- Evening (6pm-12am)
- Night (12am-6am)
- No fixed time

Where do you usually use this online social network?
- Home
- Workplace
- Internet café
- When I am mobile / moving from place to place
- Other (please specify)
- No fixed location

Other (please specify)

<table>
<thead>
<tr>
<th></th>
<th>Daily</th>
<th>A few Times a Week</th>
<th>Once a Week</th>
<th>Once or twice per month</th>
<th>Never</th>
</tr>
</thead>
</table>

How often do you read information in this online social network?
- ○

How often do you ask questions in this online social network?
- ○

How often do you answer questions in this online social network?
- ○

How often do you contribute new information or ideas to this online social network?
- ○

In your interactions in this online social network, please list up to THREE Clinical topics that have been most important to your learning.

In your interactions in this online social network, please list up to THREE Non-Clinical topics that have been most important to your learning.
Confidential

<table>
<thead>
<tr>
<th>Question</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>My interactions in this online social network are more often about Non-Clinical topics than Clinical topics</td>
<td>○ Yes</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>My interactions in this online social network have decreased over time.</td>
<td>○ Yes</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Are your discussions in this online social network facilitated?</td>
<td>○ All</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

The online social network has provided valuable information for my learning.

The online social network has contributed to my ability to address challenges in professional practice.

The online social network allows me to easily share knowledge with others.

I can easily participate any discussion topic in the online social network.

It would be valuable to me to have my participation in this online social network accredited towards my Continuing Professional Development (CPD) points.

Please enter any comments you would like to make about ways that using this online social network has improved your informal learning or might improve it in future.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Li, Xin

Title:
Learning analytics as a methodology for understanding medical practitioners' informal learning in online social networks

Date:
2017

Persistent Link:
http://hdl.handle.net/11343/194686

File Description:
Thesis

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.