Summer 6-27-2016

A CROWD-SOURCED KNOWLEDGE MANAGEMENT APPROACH TO LANGUAGE PRESERVATION AND REVITALIZATION: THE CASE OF TE REO MĀORI

Asfahaan Mirza  
*University of Auckland*, a.mirza@auckland.ac.nz

David Sundaram  
*University of Auckland*, d.sundaram@auckland.ac.nz

Follow this and additional works at: [http://aisel.aisnet.org/pacis2016](http://aisel.aisnet.org/pacis2016)

Recommended Citation
A CROWD-SOURCED KNOWLEDGE MANAGEMENT APPROACH TO LANGUAGE PRESERVATION AND REVITALIZATION: THE CASE OF TE REO MĀORI

Asfahaan Mirza, Department of Information Systems and Operations Management, University of Auckland, New Zealand, a.mirza@auckland.ac.nz
David Sundaram, Department of Information Systems and Operations Management, University of Auckland, New Zealand, d.sundaram@auckland.ac.nz

Abstract

Many linguists claim as many as half of the world’s nearly 7,105 languages spoken today could disappear by the end of this century. When a language becomes extinct, communities lose their cultural identity and practices tied to a language and intellectual wealth. Preservation of endangered languages is critical, but a challenging effort. A language is not preserved and revitalized by just documenting, archiving and developing shared resources. The revitalization is highly dependent on reviving the language among the youth.

Te Reo Māori is the native language of the indigenous Polynesian population of New Zealand, who are known as Māori. Māori widely believe that Te Reo Māori is the cornerstone of their culture and identity. Te Reo Māori is considered an endangered language. The continuous decline of active Māori speakers motivates this study to adopt new mediums in order to reverse this language shift and revitalize Te Reo Māori.

Current language revitalization initiatives focus predominately on language learning rather than issues of preservation and curation. Moreover, there is limited research on models, frameworks, and/or architectures for the design and implementation of a holistic crowd sourced knowledge management system to revitalize endangered languages. The main objective of this research is to propose and implement concepts, models, processes, frameworks, and architecture for the holistic capturing, curating, accessing and learning (use) of endangered languages with a particular focus on Te Reo Māori. This research is based on a design science multi-methodological approach, which consists of multiple phases - observation, theory building, system development and experimentation. The design science artifacts will be generalised to help revitalize other endangered languages including non-Roman script based languages.

Keywords: Language Revitalization, Social Media, Knowledge Management, Crowd Sourced, Endangered Languages, Mobile Apps
1 INTRODUCTION

Te Reo Māori is the indigenous language of New Zealand. It is a living national taonga/treasure for all New Zealanders. Before the European settlers arrived (1800s), Māori was the only spoken language of New Zealand. The dramatic decline of Māori speakers was observed in 1980s, less than 20% of Māori population were fluent enough to be classified as native speakers (Ministry for Culture and Heritage, 2012). The most recent Statistics New Zealand (2013) census report on speakers of Te Reo Māori indicates that approximately 25% speakers could speak Māori during 1996 and 2001. However, the proportion of Māori speakers further declined in 2006. The continuous decline of active Māori speakers motivates this study to adopt new mediums in order to reverse this language shift. Therefore, there is a clear motivation to revitalize Te Reo Māori.

Language revitalization is to reverse the decline of a language from becoming extinct or endangered. Te Reo Māori is considered as an endangered language and could become extinct, if no language revitalization efforts are made. Past research suggests that total immersion schools and classrooms have been very successful for Te Reo Māori revitalization (Grenoble & Whaley, 2005). Over the years, many publications emerged (books, dictionaries, audiotapes, videos, games, websites, applications and CD/DVD-ROMS) that were intended primarily for learning and reference purposes. However, there are limited initiatives aimed at capturing and curating Te Reo Māori. Combining emerging ubiquitous systems, social media and knowledge management principles provides capabilities that were not available before which gives birth to new concepts towards language revitalization (Mirza & Sundaram, 2016a).

1.1 Practical Problems identified:

• Half of the world’s nearly 7,100 languages will become extinct in this century (Krauss, 1992; Romaine, 2007).
• Te Reo Māori is considered to be an endangered language.
• The number of Māori speakers is declining (Statistics New Zealand, 2013).
• The 2006 census shows that approximately only 4% of all those living in NZ, and 24% of Māori, speak Māori (Statistics New Zealand, 2011).
• With the extinction of a language, significant parts of a culture, identity and knowledge are lost (Hinton & Hale, 2001; National Geographic, 2013).
• Current Te Reo Māori applications are limited to learning or reference tools.
• Limited initiatives aimed at preservation of Te Reo Māori.

1.2 Research problems identified:

• Limited research on models, frameworks, and/or architectures for the design and implementation of Te Reo Māori revitalization system.
• Limited research on the usage of ubiquitous systems and knowledge management for Te Reo Māori preservation.
• Limited research on holistic approaches for language capturing and curating Te Reo Māori using not only experts but also the wisdom of the crowd.
• Lack of applications that offer crowd-sourced approaches to the capture and curation of Te Reo Māori.
• Current systems offer support for standard but limited vocabulary of words and phrases and do not facilitate the capture and curation of diverse dialects and media such as imagery, poetry, proverbs and idioms that are specific to a particular Iwi, Hapū or even Whānau (Mirza & Sundaram, 2016b).

Figure 1 illustrates the research problems. The research process in the past has been in the design and implementation of traditional systems to capture and curate Te Reo Māori by experts in limited context.
focused on standard vocabulary and media. This research (red zone) tries to address these research problems by exploring crowd sourced and knowledge management approaches using ubiquitous systems to capture, curate, discover and learn linguistics diversity and richness anytime anywhere.

Figure 1: Research Dimensions

2 RESEARCH METHODOLOGY

The primary aim of this research is to design and implement a system. The word “Design” means "to create, fashion, execute, or construct according to plan" (Merriam-Webster Dictionary, 2013). Therefore, it is best to discover through design (Baskerville, 2008) and adapt a multi-methodological approach to conduct this design science research. For this study, Nunamaker Jr, Chen, & Purdin (1990) multi-methodological approach for information systems research (ISR) will be adapted to propose and develop various artifacts. Moreover, the criteria for the design science artifacts proposed by both Nunamaker et al. (1990) and Hevner, March, Park, & Ram (2004) will be followed throughout the study.

The adapted multi-methodological approach is a practical way of designing and implementing a system. It consists of four research strategies/phases - observation, theory building, systems development and experimentation. The phases are not in any particular order but they are all mutually connected to support creation and validation of a system with multiple iterations. As this research focuses mainly on design and implementation of a system, the proposed approach will follow the sequence of observation, theory building, system development, and experimentation.

2.1 Research Questions and Methods

The identified dimensions, practical and research problems lead us to propose the following research questions:

RQ1: Observation - What are the concepts, models, processes and frameworks in current academic literature (knowledge management, social media, ubiquitous systems, and language revitalization) and industrial systems to support crowd sourced approaches towards language revitalization?

The observations will help bring clarity to the research domain – to know what is currently taking place. We will examine existing academic literature on language revitalization and review existing applications that are available for indigenous languages. The outcomes of RQ1 are: comparison of existing applications and synthesis on existing concepts, models, processes and frameworks.
RQ2: Theory Building – How can we synthesize and adapt social media, knowledge management, ubiquitous systems and language revitalization concepts, models, processes and frameworks to support Te Reo Māori revitalization?

Theory building phase will consist of adapting and developing ideas and concepts, creation of conceptual models, processes and frameworks. The proposed theories will help conceptualize a generic system that supports a crowd sourced approach towards language revitalization including Te Reo Māori revitalization. The outcomes of RQ2 are: conceptual concepts, models, processes, frameworks and architectures for crowd sourced knowledge management driven approach towards language revitalization.

RQ3: Systems Development - How do we design and implement a holistic crowd sourced knowledge management system to capture, curate, discover and learn the diversity and richness of endangered languages especially Te Reo Māori?

The proposed concepts, models, processes, and frameworks will enable us to design and implement a holistic crowd sourced knowledge management system to capture, curate, discover and learn Te Reo Māori which supports dialect variations and media such as words, phrases, imagery, poetry, proverbs and idioms that are common as well as specific to a particular Iwi, Hapū or even Whānau. The development of the Te Reo Māori revitalization system will help demonstrate the feasibility of the system for other endangered languages. The outcomes of RQ3 are: Crowd sourced knowledge management system to revitalize Te Reo Māori, and a refined architecture and implementation.

RQ4: Evaluation – How to conduct laboratory or field experiments, computer simulations, surveys or interviews to validate and refine purposes theories (concepts, models, processes, frameworks and architectures) and to enhance the system.

Once the system is developed, we will conduct controlled lab experiments in the evaluation phase. Functionality, Usability, Reliability, Performance and Supportability (FURPS) quality testing model will be employed to evaluate the system. Moreover, portability will also be assessed as it is missing from the FURPS model (Samadhiya, Wang, & Chen, 2010). Development is an iterative process and the issues identified during experimentation (RQ4) will lead to further refinement or creation of design artifacts.

RQ5: Generalisation – How can we generalize the concepts, models, processes, frameworks, architectures and the system to be applicable to all endangered languages?

This research question (RQ5) will be answered by aggregating results from RQ1,2,3 and 4. The concepts, models, processes, frameworks, architectures and the system will be refined to meet the needs of other languages such as Hawaiian, aboriginal languages of Australia, and non-Roman script based languages such as Arabic, Chinese, Hindi, Urdu, Tamil and others. The Framework of Common Design Elements as shown in Figure 2 will be language independent and will be able to adapt for majority of endangered and mainstream languages.

Figure 2: Framework of Common Design Elements

In summary, the research methods illustrated in Figure 3 will begin with observation (answering RQ1), and this observation informs the building of theory (RQ2). The results obtained by answering RQ1 and RQ2 informs the design and implementation of the crowd sourced knowledge management system for Te Reo Māori (RQ3). The system is evaluated (RQ4) to refine and generalise research artifacts for other endangered languages (RQ5).
3 THEORETICAL FOUNDATIONS

In the following sections we will be looking at the theoretical underpinnings of this research namely: knowledge management, social media, ubiquitous systems, and language revitalization.

3.1 Language Revitalization

The linguistic diversity is an important part of human diversity. Language defines a culture, which cannot be translated or replaced by another language. The loss of a language is a loss of a whole culture and knowledge system including environmental, philosophical, music, oral literacy, medical, cultural practices and artistic skills (Hinton & Hale, 2001). Language revitalization is to reverse the decline of a language from becoming extinct or endangered. Krauss (1992) estimates that 90% of the world’s 7105 languages would become endangered or extinct by the end of this century, if no language revitalization efforts are made. Crystal (2002) and Nettle & Romaine (2000) have an optimistic estimate that only half of the languages will become extinct. Languages are at greater risk of extinction than species of animals and plants (Sutherland, 2003).

There are two levels of language revitalization; first level is the development of programs that results in full re-establishment of language in the community (e.g. Hebrew). Second level is reversing the decline of a language, which has first language speakers and used at home for communication (Hinton & Hale, 2001). Linguists have proposed various models for language revitalization (Grenoble & Whaley, 2005; Hinton & Hale, 2001). The language revitalization models include school-based programs (total and partial immersion), children's programs outside the school, adult language programs, documentation and materials development, home-based programs, and language reclamation. In the subsequent section we explore the fundamentals of social media and knowledge management systems.

3.2 Social Media

Social media refers to an interaction among individuals and groups where the participants are involved in the creation, sharing and exchange of information, ideas and data in a virtual community as well as networks. The role played by individuals communicating in the past was not dynamic; the consumer audience and communicator were distinct groups. At present the consumers actively create, publish, produce and broadcast in the autonomous form that a platform facilitates (Lewis, 2009).

Social media relies on defining characteristics of interactions mediated by online channels that have become a key tool to reach individuals and masses alike separated by geographical and ideological
divides (Kaplan & Haenlein, 2010; Matthews, 2010). Consequently, social media is profiled as the optimal tool to apply in modern communications and interactions where private characteristics are used as critical variables. The key strength of social media is in its interactivity and simplicity to generate content. Having defined social media, we will now move on to discuss the key aspects of knowledge management.

3.3 Knowledge Management

Knowledge can be classified as personal, shared, embedded in artifacts, or be part of an organization’s or society’s makeup and culture. Information becomes knowledge once it is processed by people in their minds. Consequently, knowledge becomes information once it is stored in the form of text, graphics, words, or other symbolic forms (Alavi & Leidner, 2001). The people are society’s basic knowledge agents.

Knowledge Management (KM) is defined as creating value by leveraging intangible assets (Setiawan Assegaff & Dahlan, 2013). KM is underpinned by the management of interaction among knowledge agents and information (Hansen, Nohria, & Tierney, 1999). Knowledge continues to grow and it is virtually inexhaustible. It is acquired through social means by leveraging the management process through information technology (Sharratt & Usoro, 2003). Knowledge Management Systems (KMS) are used to support creation, transfer, and application of knowledge in organizations and communities (Alavi & Leidner, 2001). Furthermore, according to Qwaider (2011), KMS facilitate key processes and paradigms that assist in creating, capturing, sharing and utilizing of knowledge, skills or expertise. While social media is used to create content, it is rarely integrated into the different facets of KMS. In the following section, we briefly look at the advantages of using ubiquitous systems for language revitalization.

3.4 Ubiquitous Systems for Language Revitalization

Many indigenous communities have started to adopt Ubiquitous Information Systems and Devices (UIS&D) to preserve, maintain and revitalize their language and cultural practices (Galla, 2009; Holton, 2011). UIS&D refer to systems and devices (tabs, pads, or boards) that are available abundantly without boundaries (Sorensen, Yoo, Lyytinen, & DeGross, 2005; Weisser, 1991). There is a significant rise in adoption of UIS&D among everyone; both digital natives and digital immigrants. Using this platform for language revitalization will have higher impact over a widespread population because they can collaborate, contribute and access information. Todays trend especially among digital natives “always connected” helps them to practically use the language rather than having to learn the language.

This research aims to design and implement the systems core modules, which allows users to remotely contribute and collaborate towards capturing, curating, discovering and learning Te Reo Māori. Furthermore, the design artifacts be generalised for other endangered languages.

4 CURRENT STAGE OF THE RESEARCH

The research is currently at the completion of third stage (RQ3) of the proposed research methodology in section 2. As part of this research, we were able to publish a peer-reviewed book chapter (Mirza & Sundaram, 2016a) and a conference paper (Mirza & Sundaram, 2016b) to disseminate and validate the design artifacts and prototype implementation developed. This section highlights some of the outputs of each research question that has been completed.

RQ1: Observation and RQ2: Theory Building – the synthesis of academic literature on social media, knowledge management, ubiquitous systems and language revitalization and analysis of current language systems led to the creation of concepts, models, processes, framework (Figure 2) and architecture (Alavi & Leidner, 2001; Hinton & Hale, 2001; Lévy, 1997; Malone, Laubacher, & Dellarocas, 2009; Nonaka, Takeuchi, & Umemoto, 1996; Romaine, 2007).
To assist the design of the system, a high-level use case model shown in Figure 5. This presents a graphical overview of the key functions (use cases) provided by each of the systems.

![High level use case diagram](image)

**RQ3: Systems Development** – the prototypical implementation of the crowd sourced knowledge management system was developed to revitalize Te Reo Māori. The system is built for mobile devices and we have named it as “Save Lingo” (Figure 6). Save Lingo extends upon concepts and processes that were derived from answering RQ1 and RQ2. We created a highly interactive platform through which individuals and communities capture, co-create, modify, and discover user-generated content to help revitalize endangered languages, especially Te Reo Māori. Furthermore, we also developed two learning apps that use the content captured and curated in Save Lingo core application. The learning apps are called Learn Lingo: Flash Card (Figure 7) and Learn Lingo: Hangman (Figure 8). We have generalised our research artifacts and Save Lingo system to support multiple languages including Vietnamese, Hawaiian, and non-Roman script based languages such as Arabic, Hindi, and Urdu.

The research commenced in 2012 and expected to complete by end of 2016.

### 5 CONCLUSION

The rapid disappearance of vital knowledge and culture embedded within languages, as well as the limitation of current systems and approaches motivates this research to design and implement a holistic crowd sourced knowledge management approach to revitalise endangered languages. The primary contributions of this research are towards endangered language revitalization in general and Te Reo
Māori revitalization in particular. However, the concepts, models and processes proposed could potentially contribute to closely related disciplines such as education, linguistics, Māori studies, computer science and information systems.

Figure 6:  Save Lingo app – Crowd sourced knowledge management system

Figure 7:  Learn Lingo: Flash card app workflow

Figure 8:  Learn Lingo: Hangman app workflow
References


