An exploration into the impact of neural machine translation in the service language domain

ROBERT CALLAGHAN

MASTER OF BUSINESS ADMINISTRATION

This thesis is submitted to the School of Business at the National College of Ireland in partial fulfilment of the requirements for the degree of Master in Business Administration

Submitted to the National College of Ireland (August, 2018)

Abstract

While there is a large body of literature on machine translation there is very little known about the impact of the neural machine translation (NMT) model in the context of the service language domain. Through a inductive, exploratory research approach, using in-depth qualitative interviews with five machine translation experts from both academia and the business environment this small-scale research project aimed to address this gap in the literature.

The rich data garnered from the interviews highlighted the impact of disruption within the machine translation industry. The data also indicated how technological advances have allowed the machine translation industry to adapt and evolve rapidly. There has been a convergence of translation business models which has led to increased business benefits of the contemporary machine translation model. As the NMT model has gained traction within the Service Language Domain, the data gathered also underlined the importance of looking to the future of NMT models and how businesses can best benefit through their use.

Findings from this study provide up-to-date real world insight into NMT models and their impact in the service language domain. They also point to potential for further research in this area such how the NMT model measures quality output for certain language domains.

Keywords: neural machine translation, statistical machine translation, artificial intelligence

Submission of Thesis and Dissertation

National College of Ireland

Research Students Declaration Form

(Thesis/Author Declaration Form)

Name: Robert Callaghan

Student Number: 16111958

Degree for which thesis is submitted: Master of Business Administration

Material submitted for award

(a) I declare that the work has been composed by myself.

(b) I declare that all verbatim extracts contained in the thesis have been distinguished by quotation marks and the sources of information specifically acknowledged.

(c) My thesis will be included in electronic format in the College Institutional Repository TRAP (thesis reports and projects)

(d) ***Either*** *I declare that no material contained in the thesis has been used in any other submission for an academic award.

Or *I declare that the following material contained in the thesis formed part of a submission for the award of Degree 9, Master of Business Administration

Signature of research student: ______________________________

Date: 29th August 2018

Acknowledgements

Firstly, I would like to thank the machine translation professionals who gave me their valuable time and were willing to share their knowledge and expertise.

I would also like to thank David Hurley, my supervisor for his guidance and advice not only throughout this thesis but also over the last two years, thank you.

In addition, I'd like to express my gratitude to friends, family and classmates who supported me throughout this MBA endeavor. I'm thankful for their guidance, invaluable constructive criticism and encouragement.

Furthermore I'd like to say a special thanks to my parents, Paul and Claire, who have always had my back and supported me in all my pursuits.

Above all, I'd like to thank my partner Suzanne for her patience, encouragement and inspiration.

Table of Contents

List of Figures

Chapter 1: Introduction

1.0 Introduction

Machine Translation models have developed at a rapid pace in recent years. Most recently there has been a shift from the use of Statistical Machine Translation (SMT) to Neural Machine Translation (NMT) in various language domains (Bentivogli et al., 2016). While there is a large body of literature on machine translation including SMT and NMT (May and Way, 2009, Costa-jussa et al, 2017, Peris, Domingo and Casacuberta, 2017) there is very little known about the impact of the neural machine translation (NMT) model in the context of the service language domain (Castilho et al., 2017). Even though this is a contemporary industry, it is growing rapidly therefore this research study will focus solely on the service language domain. Using a qualitative approach, this research study aims to address the dearth in the literature by exploring the topic , through semi-structured interviews, from the perspective of experts in business and academia in this field. This chapter will provide justification for the study; outline its aims and objectives; discuss the study's scope and limitations; and finally, it will provide details of the structure of this thesis.

1.1 Justification for Research

Some studies analysed the role of the NMT model with some studies confirming that the NMT model has led to improved effectiveness stating that the new NMT model outperformed the SMT model in the language domains that they researched (Bahdanau Cho and Bengio, 2015, Bentivogli et al., 2016, Luong, Pham and Manning, 2015, He et al. 2016).

However, some of the research suggests that particular aspects of the NMT model need further analysis especially in understanding certain language domains (Bentivogli et. al 2016, Castilho et al., 2017). Castilho et al., (2017) argued in their findings that even though the NMT model has shown improvements for some language domains there is still need for further research before broad generalisations can be made that the NMT model outperforms the SMT model. In fact, they indicated that depending on the language domain, the NMT model did not always provide the best results in comparison to the SMT model for the same language domain.

Costa-jussa et al., (2017) also posit that while the NMT model has shown great results in certain translation tasks within a short time period there is a need for further research. They also note that it took ten years for SMT model to become an established and widespread solution for machine translation. However, within a short period of time the NMT model has called into question the continued validity of the SMT model. They suggest the NMT model is starting to be established as the new and unified framework for machine translation which may make the SMT obsolete in the near future (ibid).

Recent research into NMT clearly shows there is a gap in the academic research to understand the impact of NMT model in some language domains. In order to provide a broader understanding of the impact of neural machine translation this study will explore the impact of NMT model in the service language domain through qualitative interviews with professionals in the area of machine translation.

1.2 Research Aims and Objectives

The aim of this research project is to understand the impact of neural machine translation in the service language domain. Specifically its objectives area as follows:

- Investigate if the NMT model is a suitable model for the service language domain.
- Investigate what is driving the improvement in machine translation specifically in neural machine translation models and how they compare to the SMT models.
- How can NMT be used to improve translation quality for the service language domain.
- What are the key areas that one would need to develop a suitable NMT engine for a service language domain.

1.3 Scope and Limitations

The study was designed to gain an in-depth understanding of how an NMT model may perform in a service language domain but due to the timeframe completion limitations it can only provide a snapshot. Ideally following a longitudinal approach without restricted time constraints would lead to more in-depth understanding of the performance of NMT model in a service language domain over a period of time. Furthermore, this study will not include any analysis of the impact of NMT model for other language domains as is often the case with qualitative data findings are not generalisable.

1.4 Thesis Structure

The first chapter provides an introduction to the subject. The second chapter in this study looks at the literature relating to the topic under study in order to establish the context in which the study was carried out. Subsequently, the third chapter describes the methodological approach employed to conduct the study and includes details on the research design, data collection, data analysis and study limitations. The fourth chapter describes and analyses the findings of this exploratory study. The fifth chapter discusses the findings of the study in the context of the literature. Finally the sixth chapter provides a brief conclusion and recommendations for further research. The next chapter is the literature review chapter. The goal of this chapter is to establish the significance of the field of study and identify a place where a new contribution could be made to the research field.

Chapter 2: Literature Review

2.0 Introduction

As previously stated the aim of this study is to explore the impact of machine translation in particular the NMT model, within the service language domain. Research studies have been undertaken to measure their impact of the NMT model but only with certain language domains. (Bentivogli et al., 2016). This review will discuss the literature and highlight opportunities to build on the current research.

This chapter will start with an overview of machine translation will be given followed by a brief explanation of translation domains and techniques used to evaluate quality. From there, this chapter will look into the evolution of machine translation models and artificial intelligence and its impact. The chapter will conclude with details of research in the area of the contemporary machine translation model: the NMT model.

2.1 An introduction to machine translation

Machines are assessed for their ability to emulate human behaviour and they are said to demonstrate intelligent behaviour if they have the ability to pass the Turing test. (Turing, 1950). The test is considered passed if the output performed by the machine and by a human, cannot be distinguished by the evaluating judge (ibid).

A machine translation engine is an example of machine that emulates human behaviour as it performs tasks usually undertaken by human translators. Machine translation is essentially automated translation and in general terms it is a process in which computer software is used to translate text from one language (source language) to another language (target language) (Peris, Domingo and Casacuberta, 2017).

Currently, machine translation engines are extremely popular and used extensively and are now easily accessible to everyone with an internet connection (Barrett, 2015, Madhavan, 2018). For example, within a click of a few buttons, internet users can translate web pages or snippets of web content (Lardinois, 2018).

As a result, there has been interest in machine translation (MT), its output and whether it is indistinguishable from what a human translator produces (Costa-jussa et al, 2017).

In terms of when and where machine translation started, one can go back almost 70 years. The 1950's saw the introduction of the first translation of words by machine. The first time a sentence was translated by a computer was when an IBM computer translated sixty Russian sentences into English. The IBM 701 Data Processing Machine made grammatical and semantic decisions that mimicked the work of a bilingual human (Gao, 2018).

While the intervening years from the 1950's to the 1980's saw little progress in this area, in the 1980's the first rules-based machine translation model was introduced. Since the 1980's the model has changed from rules-based to statistical and now today, to the Neural Machine Translation (NMT) model. Costa-jussa et al., (2017) state that the rapid transformation has only come about in the last 25 years with the introduction of the Statistical Machine Translation (SMT) model in the 90s due to the evolution of machine learning. However in the last few years there has been a shift to the new contemporary NMT model which has been enabled by the rapid evolution of artificial intelligence (Peris, Domingo and Casacuberta, 2017).

2.2 Translation domains

In machine translation, the term 'domain' refers to the origin of the training data (Bisazza et. al, 2011, Foster and Kuhn, 2007). This essentially refers to the language in one data set (domain) may be different to that in another domain. Chen et al. (2013) posit that "*the best translation practice differs widely across genres, topics, and dialects*". A combination of all these factors represent a domain. Difference of words and grammars between language can also signify a domain (Pecina et al. 2012). Whereas, Hasler (2015) posits domains as the thematic content in the training data such as language dialects or topics within the data. It is important to recognise that many studies have tried to determine the meaning of a 'domain' but the data indicates that is still no firm agreement amongst academics.

SMT and NMT models are active research topics in the field of Natural Language Processing (NLP). The process of creating translation models is very data dependent and translation quality is strongly influenced by the quality and quantity of the training data (Bertoldi and Federico 2009, Haddow and Koehn 2012, Ma and Way, 2009, Luong, Pham and Manning, 2015).

As translation quality is a crucial factor for machine translation, the source of the training data is just as crucial as the training data needs to be drawn from the same domain as the testing data (Zhang, 2017). The quantity factor draws from the need for the training data to be large enough to cover all language possibilities of the language domain. As a result, for a machine translation models such SMT or NMT a prerequisite is to capture as much high quality training data as possible. This is critical in order to gain high quality translation performance and quality (Luong, Pham and Manning, 2015).

The human language is a very complex system therefore capturing and collecting high quality clean data can be extremely difficult particularly for certain domain and language pairs. Thus, it is necessary to supplement the scarce training data with training data from other domains (Haddow and Koehn 2012, Zhang, 2017).

In this thesis we will look at the training data needed to train a machine translation system for a service language domain and if the training data is in fact available and of high quality and quantity.

2.3 Machine Translation Metrics

The machine translation models have two evaluation metrics called the Bilingual Evaluation Understudy (BLEU) (Papineni et.al., 2002) and the Translation Error Rate (TER) (Snover et.al., 2006). Both evaluation metrics take into account the number of edits needed to convert the raw translation into usable text and estimates the post-edition effort (LaGarda et al, 2009).

The human evaluation step consists of professional translation experts evaluation of the accuracy and quality of certain phrases within the raw translation. A translation is considered suitable if it can be manually edited with minimum effort. If the evaluator deems the raw translation as unsuitable, the translated text is modified and added to the translation memory (ibid). The literature points to the importance of these post-editing metrics as they are key to measure the quality of the output. This is a key driver in their adoption of the models. If the human evaluators don't like the output the perceived quality of the model is put into question.

The below sections will now explain the early machine translation model, i.e. rules based machine translation followed by the framework of statistical machine translation (SMT) which include various models. This section will then examine the introduction of artificial intelligence and it's impact on machine translation. This section will then conclude with an introduction to neural machine translation (NMT) including relevant models. Finally, the content of this chapter will then be summarised in section 2.6.2.

2.2 Rule Based Machine Translation

As mentioned previously for 30 years, the progress of machine translation was very slowly until the 1980s with the introduction of Rule-based machine translation (RBMT). RBMT was the first commercial machine translation system and was based on rules being defined by language experts and computer programmers that worked extensively to map the rules between a source and a target language (LaGarda et al., 2009).

The RBMT model can achieve positive results however in order to reach these results it takes high development costs and large amounts of training data. Furthermore, the data indicated that the maintainability of an RBMT can be extremely costly as translation quality is hard to maintain. As more and more language variations occur rules need to be updated and maintained (ibid).

Research literature also indicates that fluency has been an issue with RBMT models, with translated text reading very much 'machine like'. This means that the text relies heavily on post-editing. This post-editing phase increases the costs of using this machine translation model (ibid). While in use in the 80s and 90s, RBMT model were replaced by the first SMT model in the late 1990s (Brown et. al, 1993).

2.3 Statistical Machine Translation

Ma and Way (2009) concluded that the difference between the previous RBMT model and the SMT model is that it is deeply rooted in machine learning technology. Furthermore, SMT utilises statistical translation models created from the evaluation of monolingual and bilingual training data (ibid).

The SMT model uses high levels of computing data to build these models to translate one source language to another. Essentially, the translation is chosen from the training data using algorithms to select the most appropriate words or phrases (Ma and Way, 2009).

Peris, Domingo and Casacuberta, (2017) found that the process of building a statistical machine translation system is straightforward. The process involves uploading training data to train the engine for a specific language pair or domain. There is a minimum of two million words needed to train a translation engine for a specific language pair or domain (ibid).

Ma and Way (2009) also noted that SMT models rely heavily on bilingual language training data such as glossaries and translation memories to train the engine on the language patterns within language domains. Furthermore they stated that SMT models will have a higher quality output if trained with domain specific training data (ibid). The key point here is the importance and reliance on domain specific training data to build a successful SMT model.

SMT models require an extensive hardware configuration to run translation models at acceptable performance levels (De Bie, Turchi, and Cristianini, 2008). Therefore, cloud-based systems are recommended, so that the model can scale to meet the demands of businesses without the businesses having to invest heavily in hardware and software costs (ibid).

The pace of change within the last 20 years in machine translation due to machine learning has been staggering with the rise of statistical machine translation, leading to it becoming an industry norm. (Peris, Domingo and Casacuberta, 2017). However, even though the SMT models are straightforward to build, the data has indicated that this models needs an extensive amount of domain specific training data to achieve acceptable quality levels. Castilho et al., (2017) concluded that this is an absolute requirement for both human translators and machine translation specialists as low quality output led to expensive post-editing tasks.

2.4 Word-based models

Various SMT machine translation models were developed in order to meet business requirements of higher quality output and remove costly human post translation editing(LaGarda et al, 2009).

Statistical models were developed to combat issues of low accuracy and poor fluency. These models developed on the back of analysis of bilingual texts (ibid). The first model to combat these issues was the Word-based model developed by Brown et.al. (1993). The Word based model was developed through a combination of several existing models into log-linear fashion (ibid). The word-based model included several features that evaluated area such as language accuracy, lexicons and fluency (LaGarda et. al, 2009).

In order to evaluate these models, evaluation steps were included within the models to measure performance in terms of accuracy and fluency of translation outputs (ibid). There are two steps on the evaluation process automatic evaluation and human evaluation. The automatic step includes automation metrics to give a first glance of the quality of the translations.

In addition, Foster, Isabelle and Plamondon (1997) introduced the concept of interactive – predictive machine translation, which is an interactive process between the software and the human agent which involves the human correcting a word and the system reacting to the changes and ensuring the next translation will be better the previous. (LaGarda et al., 2009) concluded that although these models improved performance, there is still significant levels of post-editing needed. The next section will introduce artificial intelligence and how it has had an impact on machine translation models.

2.5 Artificial Intelligence and its impact on machine translation

Artificial intelligence is a sophisticated software that has the ability to perform activities that we associate with human thinking such as decision-making, problem solving and learning (Frey and Osbourne, 2013). Artificial intelligence has already had an impact on society and has become commonplace in our everyday lives (Ford, 2013). Artificial intelligence is embedded in many of the applications we use every day from GPS systems to internet search engines. Artificial intelligence and machine learning have reached a level at which cognitive human functions can be easily replicated and replaced. Moreover, these developments are happening in a relatively short space of time (ibid).

Indeed, just over 10 years ago, Levy and Murnane, (2005) stated that automation and artificial intelligence would not be able to replicate human perception and therefore driving would be insusceptible to automation. Six years later in 2010, Google announced that it had modified several cars to be fully autonomous (Frey and Osbourne, 2013).

The pace of change within the last 20 years in machine translation was firstly down to machine learning and now the recent introduction of artificial intelligence has had a significant impact on the machine translation industry (Peris, Domingo and Casacuberta, 2017). Despite this advancement in AI technology, machine translation models are still not error free and the outputs from machine translation models must be corrected by a human agent in a post editing phase (ibid). Machine translation quality is still based on human – computer collaboration. The key point here and highlighted in the data is that the advancement of artificial intelligence has been an enabler to significant reduction in post-editing tasks due to increases in fluency and accuracy (Bentivogli et al., 2016).

In the studies performed by Bentivogli et al., (2016), their analysis stated that the influence of the artificial intelligence has enabled the NMT model to become the most suitable translation model for the translation industry. Their study indicated the NMT outputs were considerably lower in terms of post – editing effort compared to SMT. In fact, they concluded that NMT outperformed phrase-based SMT on all error types that they investigated. Furthermore, they stated that they saw translation improvements for NMT improvements in six language combinations. Bentivogli et.al

(2016) posit that quality evaluators performing the post translation edits found that NMT errors were more difficult to identify than in SMT. This demonstrates the increased positive impact of the NMT model in terms of translation fluency and accuracy.

Furthermore, Bentivogli et al, (2016), state that some aspects of neural machine translation need further analysis to understand effectiveness in certain language domains. In the paper by Castilho et.al (2017) the authors argued that even though NMT has shown improvements for some language pairs and domains there is still need for further research to conducted before broad generalisations can be made about NMT outperforming SMT.

Castilho et.al, (2017) also posit that depending on the language domain with various language pairs neural machine translation did not always provide the best results in comparison to the phrase-based statistical machine translation for the same language domain. However according to the authors when they evaluated the NMT model in comparison to the SMT model in three language domains, (i) e-commerce product listings (ii) medical patent domain (iii) educational domain the NMT output was much more effective.

Costa-jussa et.al, (2017) also posit that even though NMT has shown great results in certain translation tasks within a short time period there is a need for further research. They also note that it took ten years for SMT to become an established and widespread solution for machine translation. However, within a short period of time NMT has brought into question the continued validity of SMT. NMT is starting to establish itself as the new and unified framework for machine translation which may make SMT obsolete in the near future (ibid).

2.6 Neural Machine Translation

Research in the field of machine translation has pointed to the fact that the use of neural networks models have led to significant improvements in machine translation (Bahdanau Cho and Bengio, 2015, Luong, Pham and Manning, 2015, He et al. 2016, Tu et.al. 2016).

As stated previously, the emergence of artificial intelligence has enabled the NMT model to become the state of art technology with studies confirming that NMT models were outperforming SMT models for many different language pairs (Bentivogli et al., 2016). Another key enabler highlighted in the data for the shift to NMT models is that it used to be to expensive to run these models and companies had to invest a lot of computer hardware resources in order to see significant benefits in comparison to SMT. However in recent years this has changed with computational costs coming down allowing companies to run these models (ibid).

Tu et. al, (2016) in their research highlight the advantages of NMT over SMT are firstly that the NMT models are based on sequences (sentences) and can be trained on full constituted sentences. Furthermore, due to the introduction of artificial intelligence the NMT models can learn inter-word relationships between sentences which eliminates many of the problems associated with SMT in terms of lack of fluency.

2.6.1 The shift to neural machine translation models

2.6.1.1 Encoder-Decoder Model

The recent increase in usage of the NMT model can be attributed to the fact that the models can be quickly developed and trained in comparison to statistical machine translation models which take considerable amounts of training data and computational memory (Gulcehre et al, 2017).

Importantly NMT models are based on RNN Encoder-Decoder models which were developed by Cho, van Merrienboer and Bahdanau (2014). The RNN Encoder-Decoder models are a sequence to sequence deep learning model which is, simply put, sentence to sentence translations which remove inherent issues with word ordering and incorrect word placements (ibid).

Another key attribute of the Encoder-Decoder model is the introduction of deep learning into the model. Once the source sentence is encoded and decoded into the target language the model learns from the output and translation bi-directionally (ibid). The benefit of this is the increase in productivity of the model compared to SMT models which only translate from source language to target language, i.e. English to French and not bi-directionally. The importance of this and argued by the authors is that these models lead to increased fluency as they don't translate at statistical level but at a sequence level leading to increased natural language.

2.6.1.2 Attention Based Model

The Attention based machine translation model developed by (Bahdanau Cho and Bengio, 2015) is at the cutting edge of machine translation. The model has gained popularity as it goes beyond just translation text (ibid). The Attention-based models can be trained to learn alignments between different modules such as images and speech (Luong, Pham and Manning, 2015). The next section will conclude this chapter.

2.7 Conclusion

In conclusion the data indicates that machine translation has evolved exponentially in recent years due to the introduction of new models that leverage artificial intelligence. However, there is a large requirement on having vast amounts of training data and computational memory to generate acceptable quality machine translation outputs.

Furthermore, the literature indicates that results obtained in previous studies have been mixed with some studies highlighting the superior results of the NMT model. Other studies have found further research is needed to measure impact of the NMT model for certain language domains (Bentivogli et. al 2016, Cohn et. al, 2017, Castilho et. al, 2017, Costa-jussa et. al, 2017).

Although previous research studies provide some insight into the impact of NMT for some language domains, there is a need for continued research to understand the impact of NMT. The following chapter will explain the research methodology undertaken by the researcher to explore the impact of NMT in a service language domain.

Chapter 3: Research Methodology and Methods

3.0 Introduction

Saunders, Lewis and Thornhill (2012, p5) define research as something people begin in order to figure things out in a systematic way to increase their knowledge of a given subject.

This chapter provides an account of the research design and methods used to explore the impact of NMT in a service language domain. It starts by outlining, the research question and the aims and objectives of the study. This is followed by a detailed discussion of the approach to research. The limitations of the study are also considered and the paper will look into them in detail.

3.1 Research aim and objectives

This exploratory study aims to examine the impact of the NMT model in the service language domain.

Although much has been written about the impact of NMT, there is a gap in the literature on its impact in a service language domain. This research aims to address this gap in knowledge and to conduct a detailed exploration to determine the impact of NMT model in a service organisation/domain in comparison to the effectiveness of the statistical machine translation.

Specifically, the objectives of this study are as follows:

- Investigate if the NMT model is a suitable model for the service language domain.
- Investigate what is driving the improvement in machine translation specifically in neural machine translation models and how they compare to the SMT models.
- How can NMT models be used to improve translation quality for the service language domain.
- What are the key areas that one would need to develop a suitable NMT engine for a service language domain.

3.2 Research Methodology

The aim of this study is to explore using in a qualitative approach the impact of NMT in the context of the service language domain. The research approach selected for this thesis was taken from the Research Onion (Saunders, Lewis, Thornhill, 2012, p.128). The Research Onion allows the researcher to follow a structured framework throughout the research process.

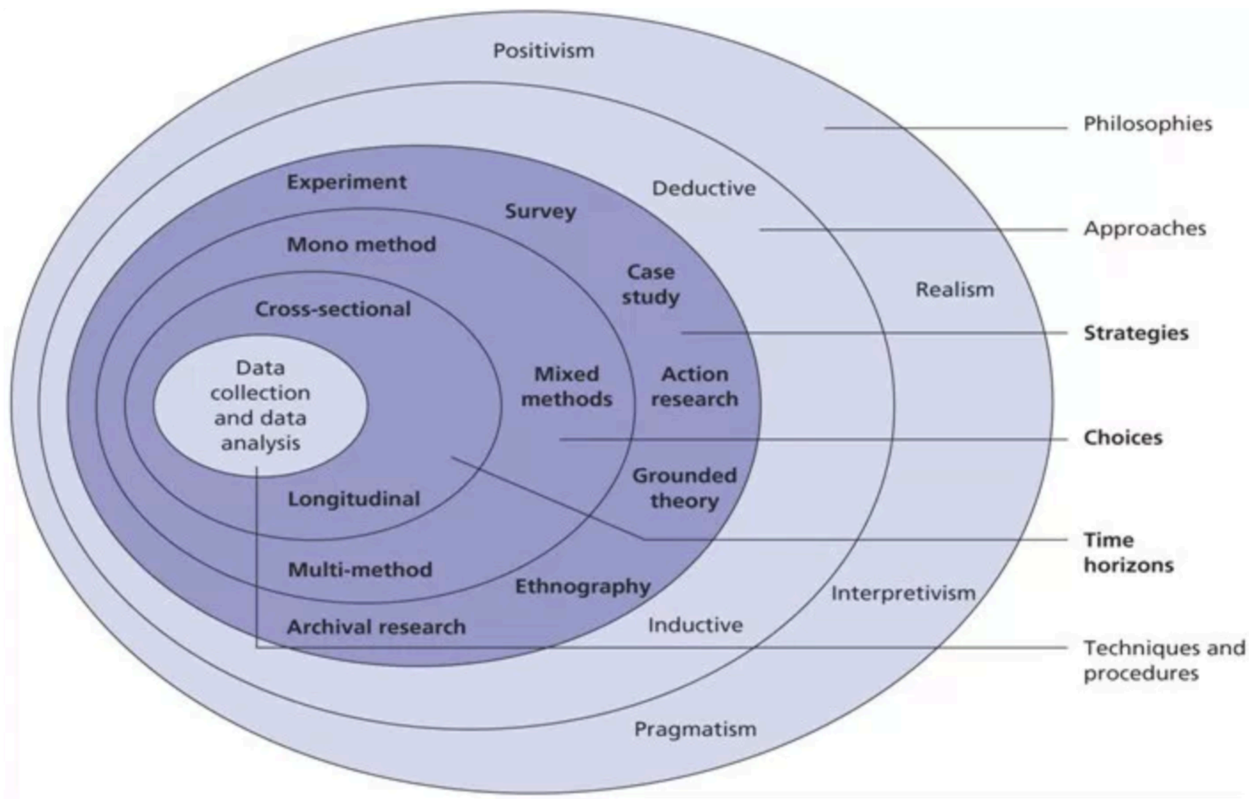

Figure 1 Research Onion Framework

Essentially the framework can be broken into three stages. The first two outer layers of the Research Onion consist of the research philosophies and approaches. The second stage and inner layers consist of research strategies, choices and time frame. Whilst the third and final inner later consists of data collection and analysis.

3.3 Research Philosophy

Saunders, Lewis and Thornhill (2012, p128) point to the fact that in research philosophy there are three research paradigms. Each of them influences the way the researcher considers not only the goals and objectives of their research, but also to try and understand the nature of the knowledge that is trying to be discovered. The three research paradigms are ontology, epistemology and axiology. (Saunders, Lewis, Thornhill, 2009, p140).

- Ontology - the researcher's view of the nature of reality or being;
- Epistemology - the researcher's view regarding what constitutes acceptable knowledge
- Axiology - the researcher's view of the role of values in research.

Understanding ontological and epistemological standpoints will provide clarity to the data that is uncovered, and help to frame the overall aims and nature of the research. For this thesis, the researcher has adopted a epistemological approach.

The initial step in devising a research methodology is to first define the research philosophy of the project. As Saunders, Lewis and Thornhill state, "The research philosophy you adopt contains important assumptions about the way in which you view the world. These assumptions will underpin your research strategy and the methods you choose as part of that strategy" (Saunders, Lewis, Thornhill, 2009, p108). The four philosophies that Saunder's highlights in the research onion pragmatism, positivism, realism and interpretivism are discussed below.

Pragmatism - the pragmatist philosophical view is focused on the belief that there are many different ways to view the world therefore when undertaking research one cannot take a single point of view as there are multiple realities (Saunders, Lewis, Thornhill, 2009, p130). Furthermore, the pragmatist's view is that if one approach or methodology doesn't present itself as the definitive way of answering the research question at hand, then a multiple method is not only possible, but also highly appropriate (ibid).

Positivism - Positivists believe that in the world that we live, there are universal truths and that research goals should discover the laws of the universe relating to these universal truths (Horn, 2009, p109). To summarise, this approach looks at society as the primary focus for research and by observing the researcher can understand how and why individuals behave as they do.

Furthermore, positivist research is highly structured and relies heavily on quantitative data and observations contribute to statistical analysis (Saunders, Lewis, Thornhill, 2009, p135)

Realism - The realism philosophical position is related to the nature of reality and what we sense is reality, is independent of human mind and very similar to positivism in terms of the scientific approach to the development of knowledge (Saunders, Lewis, Thornhill, 2009, p136). Furthermore, there are two schools of thought within in realism, direct and critical realism.

- Direct realism - the direct realism approach is very much grounded in the perspective that what you see is what you get and what one experiences through ones senses portrays the world accurately (ibid).
- Critical realism argues that there are two steps what we experience and the sensation we feel from the events themselves.

Interpretivism - interpretivism is at the complete opposite end of the research philosophical spectrum to positivism and argues that positivism doesn't adequately provide a real understanding of the complex interrelationships in society and between individuals (Walliman, 2005, p204). Furthermore, interpretivism is best applied to research which may be too "complex to lend itself to theorising by definite 'laws' in the same way as the physical sciences" (Saunders, Lewis, Thornhill, 2009, p137).

Moreover, Saunders et al, posit that following the interpretivism philosophy is highly suitable to business and management context as "they are a function of a particular set of circumstances and individuals coming together at a specific time" (Saunders, Lewis, Thornhill, 2009, p137).

By reviewing the research questions, objectives of the research and the research philosophies the researcher can start to determine the most appropriate philosophy. By taking the literature review into account, the 'best' use of NMT in the context of a service domain is still undecided. Therefore, in critically evaluating each of the research philosophies it would be difficult to apply the positivist or realist philosophical approach to the research at hand as the findings would be subjective. As a result of this, the researcher has selected interpretivism as the research philosophy most appropriate to answer the research question put forward.

3.4 Research Approach

The succeeding layer of the research onion is the research approach. According to Ghauri and Gronhaug (2005, p10) the research approach is critical as it lays down a set of systematic, disciplined procedures to answer the research question at hand.

Saunders, Lewis and Thornhill (2012) posit that there are two approaches to consider, the inductive approach and the deductive approach. The inductive approach draws general conclusions from observations and findings which in turn can be incorporated back into the exiting theories therefore building on the body of knowledge (Ghauri and Gronhaug, 2005, p15).

Saunders, Lewis, Thornhill (2012, p.145) state that the inductive approach is when 'data collection is used to explore the phenomenon, identify themes and patterns and create conceptual framework'. As such, inductive research can often focus on a smaller sample size, with a focus on qualitative data.

Deduction by comparison starts with a focus on theory and the research design is to test the theory. Collis and Hussey (2009, p8) define deductive research as "a study in which a conceptual and theoretical structure is developed which is then tested by empirical observation; thus particular instances are deducted from general inferences". The deductive approach is very much the dominant research approach for social sciences and this approach to research is the one that people typically associate with scientific investigation (Saunders, Lewis, Thornhill, 2012, p.145).

Due to its highly structured approach one key characteristic of the deductive approach is the ability to explain causal relationships between concepts and variables (ibid).

The researcher reflected on research philosophies and epistemologies from the above section – deductive research believes that there is only one truth, that’s validated by observable phenomena, whereas inductive research believes that there is no single truth, and that it's subjective and subject to change. Hence, it's unlikely that there is a single true answer to the research question and more importantly any that could be measured in a reliable sense. Due to the research question an inductive approach was considered appropriate to this thesis in order to explore the impact of NMT in the service language domain. The researcher felt obtaining contextual qualitative data should allow the researcher to draw conclusions which can help him provide answers to the research question.

3.5 Research Strategy

In this section the focus will be on the next layer of the research onion which is the strategy layer. The research strategy is defined by Saunders, Lewis, Thornhill (2012, p.173) as the plan in which the researcher will answer their research question. The strategies available to the researcher are: experiment, survey, archival research, case study, ethnography, action research, grounded theory and narrative inquiry.

Saunders, Lewis, Thornhill (2012, p.173) assert that the researcher's choice of a strategy needs to be led predominately by the research question, their objectives and a reasonable level consistency in order to create a link from research philosophy, to approach and purpose. Importantly, Saunders, Lewis, Thornhill (2012) note that when researchers are considering their research strategies they must consider pragmatic concerns and take into account the extent of existing knowledge, availability of and access to potential participants, and amount of time to complete their plan.

In terms of which strategies suit which design, experiment and survey strategies are exclusively linked to quantitative research design, whereas archival research and case studies can be either quantitative and/or qualitative.

The remaining four research strategies ethnography, action research, grounded theory and narrative inquiry are exclusively linked to a qualitative research design. The research strategy applied in this thesis will be narrative inquiry through the use of semi-structured interviews. This strategy will be discussed further in the qualitative data primary collection section.

3.6 Research Choice

When it comes methodological choice Saunders, Lewis, Thornhill (2012) posit that the first choice is between mono method and a multiple methods research design.

There are three possible research paradigms available to the researcher which are qualitative, quantitative and mixed methods. Qualitative research explores "*the meaning individuals or groups ascribe to a social or human problem*" while quantitative research is concerned more with testing theories and relationships between variables (Creswell, 2014, p4).

The below figure provides an illustration of the methodological choices and possible methods. (Saunders, Lewis, Thornhill 2012, p165)

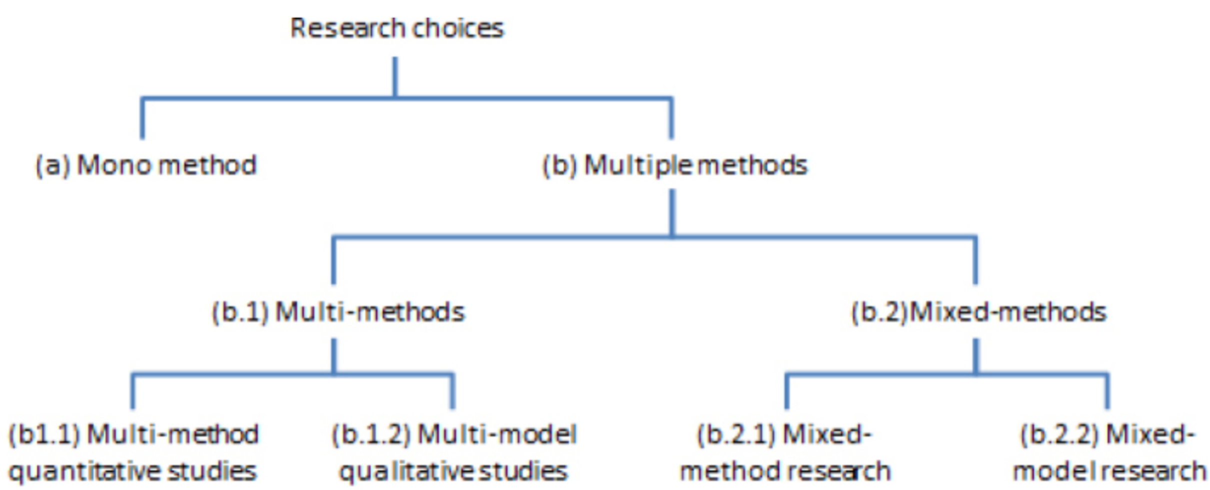

Figure 2 Multiple Research Methods

As this thesis wasn't concerned with testing theories and relationships between variables and focused on experiences of individuals a qualitative method was chosen. Semi-structured interviews were carried out with machine translation experts in various locations around the world with the prospect of gaining in-depth information on the impact of the NMT model in a service language domain.

3.7 Time Horizon

The next layer of the research onion and final section in research design is related to time horizon in which there are two choices cross-sectional or longitudinal. Saunders, Lewis, Thornhill (2012, p.190) state that this is a critical point in research design as the researcher needs to ask themselves will the research be a snapshot at a particular time (cross-sectional) or a series of snapshots over a given period (longitudinal).

For the thesis, the cross-sectional approach was taken. The key factor to choosing cross-sectional is down to practicality. It would be of interest to follow a longitudinal approach to measure the impact of neural machine translation in a service language domain over an extended period however the research project has a set timeframe for completion therefore following a cross-sectional is the best approach.

3.8 Secondary Data Collection

The final inner layer of the research onion is data collection and data analysis layer. For secondary data collection Ghauri and Gronhaug (2005, p91) state that secondary data information is critical in order to not just solve the researchers research problem but also to understand and explain the research problem at hand.

Furthermore Saunders, Lewis, Thornhill (2012, p.307) state that secondary data fits into three main sections, documentary, surveys and multiple sources. The below figure illustrates the three main sections in detail.

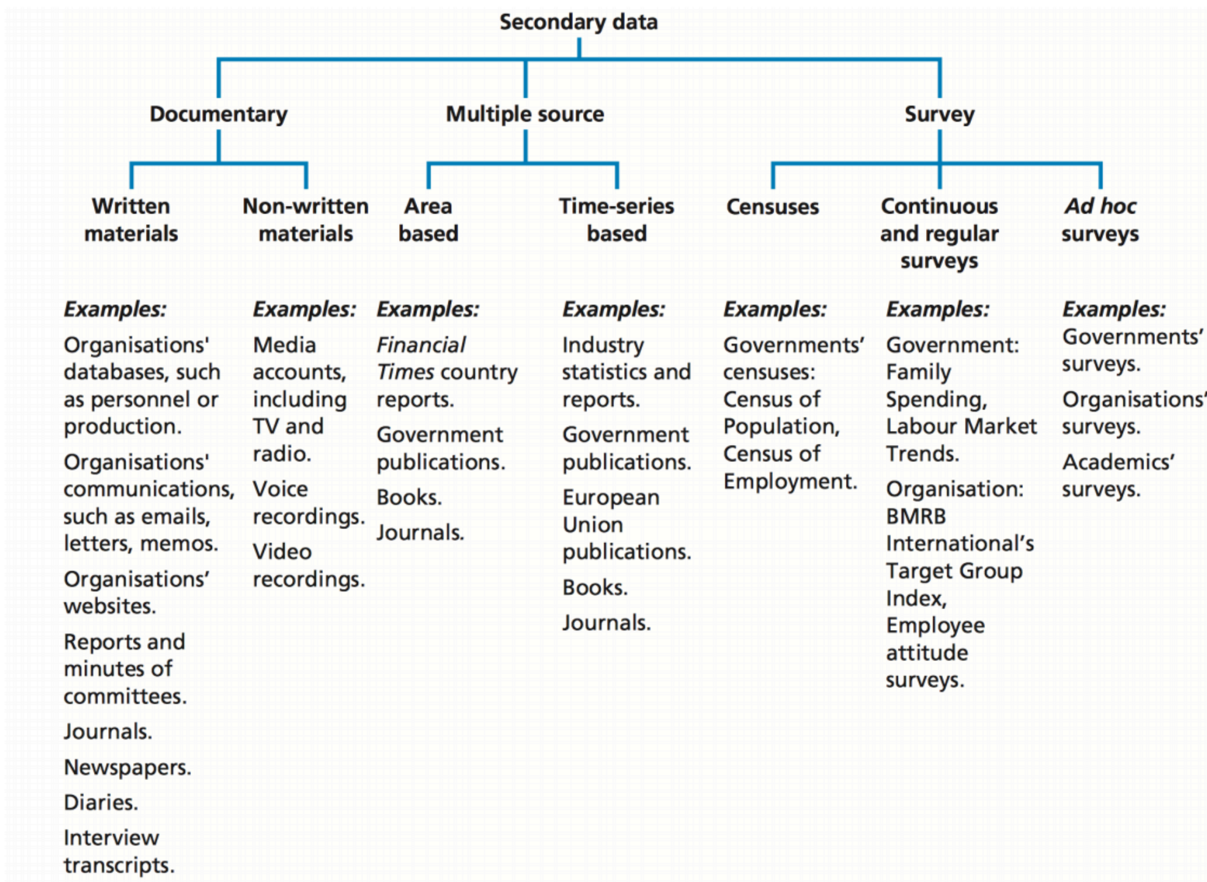

Figure 3 Types of Secondary Data

In terms of this thesis the researcher focused mainly on the multiple sources section and specifically academic journals and books.

3.9 Qualitative Primary Data Collection

In terms of qualitative primary data collection there were numerous options available. Due to the inductive research design, qualitative research interviews would suit the research problem. Saunders, Lewis and Thornhill (2012, p375) categorise interviews into two distinct areas standardised and non-standardised. The figure below will illustrate how these two areas are broken down and the process in which the interviews should take place such as face-to-face or over the telephone.

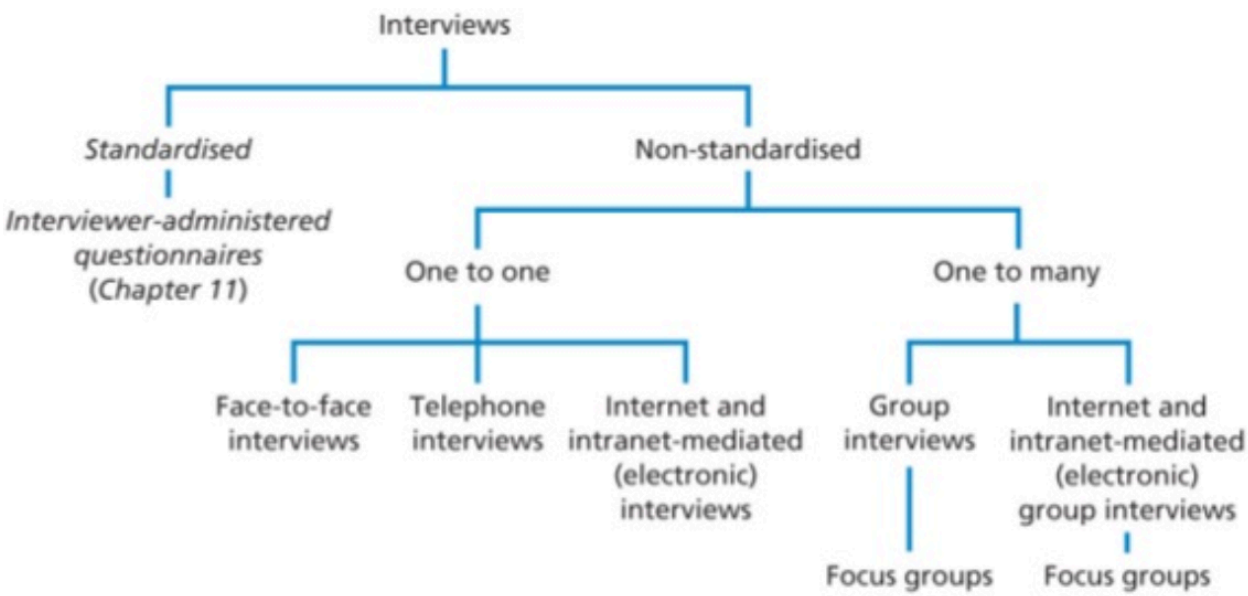

Figure 4 Forms of Interviews

Due to the exploratory nature of the inductive research design the researcher felt that in-depth semi-structured interviews would be most suitable to the research problem. As stated by Ghauri and Gronhaug (2005, p133) and Saunders, Lewis and Thornhill,(2012, p377) this method of primary data collection is highly suitable for exploratory and inductive types of study as the main advantage of interviews is that it allows the researcher to gain an accurate picture of the interview respondent's thoughts and behaviour.

Furthermore, the semi-structured interview approach gave the researcher the chance to dig deeper when the respondents answered questions. Saunders, Lewis and Thornhill (2012, p378) posit that this is critical when the researcher has adopted an interpretivist epistemology. The authors also point out that with the non-standardization of the semi-structured approach may raise questions of reliability, bias, generalisability and validity (Saunders, Lewis and Thornhill 2012, p381).

In order to combat these concerns with the semi-structured approach the questions outlined under each topic were intended as prompts for the interviewer should they be necessary.

In terms of the type of questions that should be asked the four topic areas, based on the aims and objectives of the study, were identified as relevant to the research project.

Under these topic headings in the interview schedule, the interviewer listed possible questions that could be asked during the interview. Although perhaps it could be argued that an in-depth semi-structured interview ought not be prescriptive, it was important that the interviewer was cognizant of his own status within the translation industry which may cause bias. Thus a semi-structured interviewed ensured that the interview remained focused on the research questions.

Furthermore, Saunders, Lewis and Thornhill (2012, p391-393) posit that researchers should follow an appropriate use of different types of questions. The three different types that the authors call out are open questions, probing questions and closing questions. Furthermore, the authors posit that following this approach of using all three types of questions will increase the success of the semi-structured in-depth interview process. This approach can be seen in the interview schedule in section 3.12.

3.10 Population

Participants were accessed through the researchers business network. General purposive sampling was used to select interviewees.

The researcher interviewed five experts in the machine translation industry via telephone interviews.

- Participant one - Machine Translation Lead - Translation Vendor.
- Participant two - CEO – Translation Vendor and published academic.
- Participant three - Client Services – Translation Vendor.
- Participant four - Translation Lead – Large multinational.
- Participant five – Machine Translation Specialist - Large multinational.

3.11 Interview Schedule

Below is an outline of the questions posed in the semi-structured interviews.

Open Questions

- What are your thoughts on SMT?
- What is your opinion of neural machine translation?
- What major changes have you seen in recent years in relation to machine translation?
- Can you tell me a bit about your experience of machine translation?
- What do you think are the shortcomings of NMT in comparison to other machine translation models/engines?
- Have you been involved in successful NMT engine releases and if so what domains were they?
- For you what are the key differences between NMT and previous machine translation models?
- What are the key factors in building an NMT engine/model?
- What are the challenges in machine translation industry?
- What are the opportunities in the machine translation industry?

Probing Questions

- Elaboration Probe - Could you tell more me about that?
- Clarification Probe - When you said the project was successful? What do you mean by success?

Closing Questions

- What excites you about the possible changes in NMT models can bring to machine translation in the future?
- On a scale of one to ten, how significant do you think NMT models are to machine translation?
- What do you think MT will look like in 3-5 years and what are the key factors in the changes?

3.12 Analysing Qualitative Data

One of the key challenges in qualitative research is performing meaningful analysis especially due to the non-standardisation of the data (Saunders, Lewis and Thornhill, 2012, p546). For the purposes of this thesis, the researcher will be utilising thematic analysis, as described by Braun and Clarke. "Thematic analysis is a method for identifying, analysing, and reporting patterns (themes) within data" (Braun and Clarke, 2006). The thematic analysis framework is very useful as it allows the researcher to proactively and systematically discover themes and patterns within the data, and synthesise the data in a way that allows for meaningful insights. Furthermore, the thematic analysis framework is extremely flexible which can provide a rich and detailed account of data (ibid).

Clarke and Braun (2006) believe that certain considerations and decisions need to be made at the start of the analysis phase in order to ensure consistency through the process. These decisions revolve around the nature and definition of a "theme", as well as how that theme is reported and quantified**.** With that in mind and for the purposes of this research, a theme will be defined as an idea or concept which reoccurs across either individual or multiple interviews and that has a direct correlation to:

- The impact of NMT in the service language domain.
- The interviewees viewpoints on the impact of the NMT models and their future in the machine translation industry.

Furthermore, Clarke and Braun (2006) believe that it is essential to decide whether an analysis will be semantic, (i.e. carried out at an explicit level, focusing only on what the subject says) or latent (i.e. implicit, going beyond what is said and beginning to examine underlying ideas and concepts). The analysis of this report will focus on the semantic, rather than the latent – looking for patterns in what interviewees say, and then trying to interpret and synthesise their responses.

3.12.1 Data familiarisation

In order to familiarise himself with the data the researcher transcribed the interviews verbatim and read them numerous times to become familiar with the depth and breadth of the content that had been transcribed. This repeated reading of the data allowed the researcher to start to see patterns in the data. Braun and Clarke (2006) argued that this phase is critical and acts as the foundation for the entire analysis.

3.12.2 Generating initial codes

This phase began after the researcher had read and familiarized himself with the data. The researcher used MaxQDA, a qualitative data analysis software, to help code all five qualitative interviews. The software was extremely useful in dividing the interview text into segments by assigning codes. This involved naming sections of data with specific codes. Furthermore, the researcher worked systematically through the entire data set, giving each interview equal attention and trying to identify interesting aspects of the data that may form key themes or patterns across all interviews.

3.12.3 Search for themes

Clarke and Braun (2006) state after generating initial codes that it is critical to re-focus the analysis at a broader level and start to group similar codes into themes. Drawing up a thematic mind map assisted the researcher in searching for key themes in the qualitative data. At this point, the researcher can start to combine codes and start to look for overarching themes and relationships between themes. The below image is a mind map codes generated from the semi-structured interviews using MAXQDA.

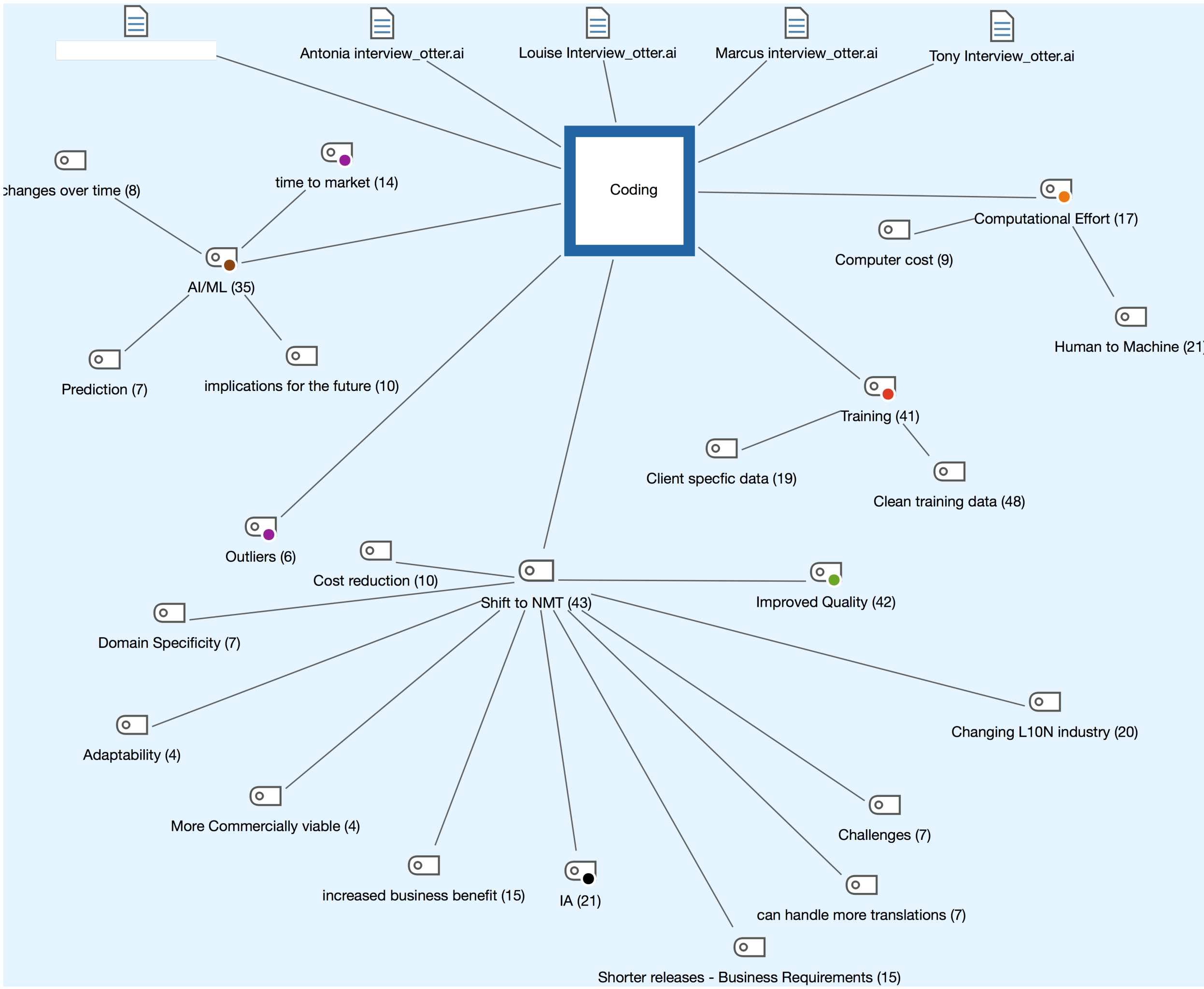

Figure 5 Mind map of codes

3.12.4 Reviewing and naming the themes

This stage involved reviewing the themes to see if they fit into the overall research objectives. The researcher also ensured that there a clear and identifiable distinctions between themes.

3.12.5 Producing the report

In this final part of the data analysis the researcher used the themes that had been identified to produce the report discussing his findings.

3.13 Ethical Issues

Before commencing the research study, ethical approval was sought from the Department of Business at the National College of Ireland as part of the research proposal in January 2018.

All participants were given a detailed written and verbal account of the nature and aims of the study following which participants were asked to sign an Informed Consent form. While reiterating the points outlined on the Informed Consent form, the researcher also assured the participant of anonymity and confidentiality and their right to withdraw from the interview at any stage without any negative repercussions.

As with most studies involving the primary collection of data, there are always potential risks for the researcher and steps were taken to ensure that these were minimal. All interviews took place by telephone at a time convenient for the participants.

The audio files and transcribed interviews are currently stored on a password protected computer. The audio files will be destroyed after the thesis has been examined.

3.14 Limitations of Research

As mentioned previously in the time horizon section (3.8), the research project had a set timeframe for completion and therefore the researcher could only interview a small number of participants. Following a longitudinal approach without restricted time constraints would lead to more in-depth understanding of the performance of NMT in a service domain over a period of time. This small-scale qualitative study aimed to address the dearth in the literature by exploring the impact of the neural machine translation model in the service language domain. The study was designed to gain in-depth understanding of how NMT may perform in a service domain but as it often the case in qualitative studies, findings are not generalisable.

Chapter 4 - Research Analysis and Findings

4.0 Introduction

As outlined in the previous chapter, this study aimed to explore the impact of the NMT model in the service language domain using the thematic analysis framework outlined by Braun and Clarke (2006). It has produced rich narrative data on the subject under study. This chapter will outline the study findings.

From the interviews four key themes emerged as essential to the understanding of the subject: Disruption in the machine translation industry, technological advancements, convergence of business models and the future of machine translation.

The researcher hopes that his findings can encourage machine translation experts and academics to conduct further studies examining and understanding the impact of NMT in the service language domain.

4.1 Analysis and Findings

The analysis section of this research utilised the Braun and Clarke (2006) thematic analysis framework to assess the research data. This chapter will outline and discuss key themes identified through this exploratory, inductive research study namely disruption, technological advances, convergence and the future state of machine translation.

4.1.1 Disruption

The first theme that will be discussed will be disruption. There have been high levels of disruption in the translation industry and there are some key levers that have been driving this disruption which are discussed below.

4.1.1.1 - Business Requirements and rising expectations

From the outset, the majority of interviewees highlighted the importance of recognising the high levels of disruption in the translation industry. The critical point highlighted through the data is that businesses and the way they operate are driving the disruption in the translation industry.

Businesses are no longer releasing products every eighteen months to two years but are now following agile development cycles with new products released or variations of products of every two to three weeks. The time to market has contracted for businesses and they are required to move more quickly and effectively to ensure they meet market demands.

'Companies are no longer releasing a new product every two years, a new product effectively or a distribution of a product is now every two, three weeks, because of the agile development techniques that companies are now using. The translation industry has to dovetail into that agile software development delivery'. (Participant two)

There was a common consistent tread emerging from the data from all interviewees. The significant reduction in time to market has reduced significantly due to the business requirements of many of their clients. Expectations of businesses have risen significantly in terms of expected timeframes for the delivery of high quality translation.

'It's a big challenge. The timeframe for translations are getting smaller, and the cost factor is getting smaller. So it's almost like we're being asked to do more faster. And that's really what's happening in the industry'. (Participant two)

4.1.1.3 - Globalisation

What became apparent through the research was that businesses were not just releasing products every few weeks to one particular market but across all the markets in which they operate. Several of the interviewees pointed out that as companies continue to grow into other geographical markets the amount of information that needs to be translated on a regular basis has grown exponentially. This globalisation factor has driven the need for accelerating improvement in machine translation models.

'The amount of content that is produced every day is just unbelievable and growing all the time and because of its sheer size, it cannot be done by human translators. The only option is to use machine translation'. (Participant one)

4.1.1.4 - Traditional Translation Industry Business Model

The traditional translation industry business model is based on cost per words translated. Every word translated by a human translator, the translation company can charge up to 20-25 cent per word. As a result, a ten thousand word document could cost up to $2,500 for one language, multiply across multiple languages and costs start to skyrocket. Participants highlighted the inadequacy of this model and the challenge it posed. For example, one participant discussing work he was doing with large multinational felt were it not for the NMT model the translations would never have materialised due to amount of content to be translated in a short period of time.

The adoption of machine translation into more language domains, such as the service language domain, is highlighting non-traditional translation requirements which require a non-traditional translation business model. The data suggests that companies are eager to move forward with new translation models, specifically the NMT model as it meets their translation requirements.

'*I think that's one of the biggest challenges for the translation industry, a lack of clear a business model. It doesn't really work when basically their business model doesn't fit our business model anymore'.* (Participant two)

The data also highlights a clear theme that businesses are eager to invest in machine translation. As one of the interviewees, Participant three stated *'NMT become a viable option for companies, clients are coming to us and saying, really want to get started with this, we know it's good quality. There's definitely a lot more confidence in the whole process'.*

The data indicates that businesses are now understanding this need to move away from the traditional translation business model.

'NMT has brought a major change in the translation industry. By the end of the year NMT is going to be surely the main paradigm in machine translation' (Participant one).

'There's a convergence of business models and we (as machine translation vendors) can now offer a business model that gives confidence and predictability to businesses (Participant two)'.

This move is being driven by the paradigm shift in the translation industry which in its essence is essentially a move away from the SMT model to the highly effective NMT model.

4.1.2 Technological Advances

The second theme that will be discussed are technological advances. The rapid technological advances in recent years have been an enabler and removed some of the barriers to entry that many of the machine translation companies have been facing. Below are the key enablers and their impact.

4.1.2.1 Computational Advances

It is important to recognise that many studies have pointed out that its only recently that we have seen significant improvements in machine translation (Bahdanau Cho and Bengio, 2015, Luong, Pham and Manning, 2015, He et al. 2016, Tu et al. 2016).

In fact the main reason was that NMT was too computationally costly and businesses had to invest significant resources in order to see benefits in comparison to SMT, however in recent years this has changed (Bentivogli et al., 2016).

In the interviews, the majority of respondents noted a significant reduction in costs of building NMT engines which has resulted in a massive increase in adoption. It was clear from the interviews that a reduction in the computational cost has helped with NMT adoption in the business environment. Several of the interviewees stated that the ability to access CPUs (central processing units) via the cloud such as Amazon AWS or Microsoft Azure have allowed companies to run these NMT models.

'What's frightening, these algorithms were around in the 1970s, we just couldn't run them. We're now able to run these models today, because we're running supercomputers, CPU based machines on these models. (Participant two)

'The main change here is the availability of computing power, the computing power is there, and it's cheaper, so we can now run these NMT models' (Participant three).

The data suggests that even though the set-up of NMT is expensive, companies are willing to invest in translation machines due to the reduction in time to market and assurance of quality output. These factors are driving the move towards NMT and its adoption in multiple language domains.

'There is a reduction in cost over time, the decrease time to market has been critical to the increased adoption of NMT model (Participant three)'.

4.1.2.2 Training Data

Emerging from the data was the importance of training data for the success of the NMT model. The availability of training data at the start of the machine build process is essential to building a solid NMT engine. As stated by all of the interviewees, the difference between SMT and NMT is the fact that with the NMT engine, getting as much clean training as possible is critical. Furthermore, clean training data appears to make the difference between a good and a great NMT engine.

'We are really in the age of garbage in garbage out. Because a lot of our algorithms today are non-deterministic, clean data is critical and getting as much clean data as possible, and then continually adding to that, having clean data is a key corporate advantage' (Participant two).

However, all interviewees stated that getting large amounts of clean training data is one of the biggest challenges to building an effective NMT engine. However, emerging from the data was the implication that for NMT engines the training data doesn't have to be domain specific. In fact as long as the data is clean the engine will benefit. This is a key finding for this thesis as the literature stated that having domain specific training was critical to the success of building an NMT engine.

'For an NMT engine, we have found out that the domain specificity is not important anymore, especially as it was for SMT engines (Participant one)'.

The findings from the data also suggest that training an NMT engine with domain specific data causes it to become confused. Restricting training data leads to a lack of fluency and reduction of quality in terms of output.

Another key theme that emerged clearly from the data was the importance of having client specific training data, such as extensive translation memories and terminology. The aforementioned augments the NMT engines performance with exceptional quality therefore reducing post-editing resources and costs.

'The results really can be improved by the addition of client specific training data that really adds the polish because what you're introducing there is the right terminology use, you're introducing the vernacular and translation style that you guys have grown used to seeing in your localization efforts' (Participant two).

4.1.2.3 Humans to Machines

A key attribute of the NMT model is the introduction of deep learning into the model which enables the model to move past the statistical restrictions of the SMT model. The neural networks of the NMT are modelled on human brains and progress has been staggering and has led to rapid change in the human's role in the translation process.

'Around 2000-2001, we had our first commercial statistical systems on the market and we've been using those for the last 16 years. But in the last 18 months, neural MT has pretty much surpassed them. It's frightening to think that we've been able to replace nearly two decades of research within 20 months to 24 months' (Participant two).

'Instead of using and the sweat of humans to do the work the raw CPU power of computers generate those rules and building an engine went from two years to a few days' (Participant two).

This compares with the two years that computer engineers could spend on writing rules for the SMT models. The key learning here is that all this computation is occurring behind the scenes with little human involvement and tasks have moved from humans to machines. This is driven by the increased fluency and accuracy that the NMT model is bringing.

'I think the neural engines are providing a much better output. So then that brings us to the conversation about what is the role of the human translator or what is going to be the role of human translator, in the future. The human translator role will evolve probably towards something more technical or more project managing' (Participant one).

In fact one of the interviewees has gone through this process and moved from a human translator into a role of solely doing post-edits.

'[…] I always worked as a normal translator, but now with new opportunities, I've also changed my job. I've become a post-editor of the machine output' (Participant five).

4.1.2.4 Artificial Intelligence

Machine translation has been in existence for over 20 years but its only in the last few years that the pace of innovation has accelerated with the evolution of artificial intelligence and machine learning (Peris, Domingo and Casacuberta, 2017). Findings from this study indicated that AI has had a big impact on the world of machine translation and led to the increased pace of adoption of the NMT model.

'AI is staggering. It's frightening and is having a big impact, I think AI is going to pair us with all sorts of different applications, different user scenarios in the future (Participant two).

This, it could be argued, has accelerated the move from human to machine translation with humans playing an ever smaller role in the translation process.

4.1.3 Convergence

The third theme that will be discussed is convergence. This section will look at the reasons why businesses are starting to take notice of NMT model and converge with the new technology.

4.1.3.1 Shift to NMT

Another theme that became apparent in the interviews was the shift from using SMT to the new NMT models and the resistance to change from certain parts of the translation industry. As discussed previously in section 4.1.1, there was significant disruption in the translation industry caused by shifting customer business models and their rising expectations.

All interviewees acknowledged that the translation industry has started to change to meet their customers' needs but it has been slow to adopt the NMT model. The interviewees believe we are at an inflection point as the industry needs to change rapidly and this rapid change requires a convergence of business models.

'The emergence of neural machine translation, shatters the industry to pieces. And this is why the industry I think, has been slow to embrace NMT and are somewhat skeptical about it, you know, and in some cases are quite resistant to talking about the advances of machine translation because it impacts their business model' (Participant two).

Another interesting sub-theme emerging was the convergence of machine translation and digital marketing due to the fact that companies now need to be consistent with their messaging across all media platforms instantaneously.

'I was at a conference a couple of months ago about translation and localisation [translation] and digital marketing are converging with each other, because localization itself, you know, cannot be separated anymore from digital marketing'. (Participant one).

4.1.3.2 Increased Business Benefit

The participants indicated that the NMT model provides increased business benefit to clients and to the business itself in particular because of the increased quality output that the NMT models provide.

'The neural engines or neural networks are providing a much better output. The main advantages of NMT models is the quality of the output (Participant one)'.

Participant four is on the client side and recently performed a test comparing the SMT model to the NMT model recently for service language domain and the results were very impressive.

'The outcome of our testing for the NMT engine for German customer service content, the NMT engine performed so much better than the SMT engine. So yes, I've seen proof of it being much superior than the previous SMT model'. (Participant four)

Another key element that arose under increased business benefit was the cost reduction for businesses. All translated content including human translations, SMT and NMT go through a post-editing cycle. The key performance indicators for businesses is post-editing, the need for post editing is reduced by using NMT versus other forms of translation. One of the interviewees highlighted that you need to compare the overall cost to ensure there is a reduction in cost. For example, the amount of post-editing hours needed indicates overall cost reduction. This convergence is happening with machine translation being recognised as the first step in the translation process before post-editing tasks take place. One of the respondents performed an NMT test of 500 content segments and asked a post-editor to review.

'The post-editor hadn't realised until halfway through the project that he was reviewing NMT machine translation he thought he was reviewing a two human translators work, it gives an indication as to the quality of the NMT engine that to that point he could not tell the difference between the machine translation segments on a on a human transcript segment' (Participant three).

This accuracy and quality point again to the business benefit of NMT models in terms of cost reduction.

Another theme that emerged from the interviews was the adaptability of the NMT engines and that once trained they can be tuned to increase performance via the addition of new terminology as it becomes available. Several interviewees pointed out that this is another key difference in an SMT engine as the NMT model learns through constant use.

This leads to reduced costs in terms of maintainability. This theme of cost reduction was consistent across the interviews. The fact that businesses don't have to continually build new engines means that NMT engines are much more scalable and commercially feasible.

‘Companies and academic institutions are looking at NMT from a commercial perspective, and coming up with ways to reduce build times. NMT is more usable, user friendly and the idea here it's now established. But now it's a question of making a more commercially viable’ (Participant three).

4.1.4 Future state of machine translation

The fourth and final theme that will be discussed is the future state of machine translation. This section will look at the reasons why businesses will continue to invest in machine translation and specifically NMT models.

4.1.4.1 Return on investment

One of the key themes emerging from the data was what the future held for the translation industry. The data stated that there was a key dependency on the confidence level of their clients that machine translation would provide return on investment. They felt confident that the NMT engines that they build are providing noticeable return on investment leading to increased business.

Several of the interviewees also stated that the combination of increased computational power and influence of artificial intelligence will allow them to build and train engines even faster. They felt that this was going to increase the suitability and adoption of the NMT model.

4.1.4.2 Development and integration

Data indicated that AI will have more significant implications for the future of the translation industry as businesses will start to integrate NMT models into other applications. Respondents believe that as time to market reduces they will be able to integrate NMT engines into other applications increasing their effectiveness and increasing overall business value.

'I think integration is the absolute be all and end all. We're doing some really good stuff with integration and adapting the engine to get to market quicker. So I think integration is something that we looking at very carefully and trying to improve' (Participant three).

4.1.4.3 Quality development and prediction

One of the participants went on to state that AI is allowing them to build models that not only have high levels of quality at a sentence level but will now be able to predict context and add context to the translation output. He stated that they will be switching over to the new improved model shortly.

'The next step is context-based machine translation, which we will be shortly switching over to that translating at paragraph level and will truly encapsulate the subject matter and be knowledgeable of the subject matter that it's translating, that's going to be a key improvements going forward' (Participant two).

The fact that their new improved model will introduce prediction through predictive analytics will allow them to predict the accuracy and quality of the translation before it goes through the engine. He states that this has been a big challenge for the machine translation industry as they will now be able to give businesses increased confidence in the NMT model..

'We'll be able to detect translations anomalies, and that's a big, big challenge. And to me, it's probably one of the biggest challenges of the industry today, how can we predict the quality of the output. In September/October, we'll be able to tell you how good the translation will be at a percentage level (Participant two)'

This participant believes that this will build confidence in the industry around using NMT and help drive the machine translation business model.

'By detecting translation anomalies we can build the industry's confidence around using NMT and getting them to the point where they can use that to drive our business models with confidence and predictability' (Participant two).

4.2 Conclusion

Data garnered from the interviews provide rich insight into the area of machine translation from the viewpoint of experts in the field. There are key trends in the industry which are driving change and leading to the increased adoption of the NMT model across more language domains. The following section will discuss the findings and their relevance in the context of the objectives of this research study.

Chapter 5: Discussion

The research findings were derived by conducting semi-structured interviews with four machine translation professionals with substantial expertise in machine translation from both academic and business perspectives. The following discusses the findings of the study in the context of the research objectives.

5.1 - Objective 1 Research findings

Investigate if the NMT model is a suitable model for the service language domain

Although previous research investigations provided some insight into the effectiveness of neural machine translation (Bahdanau, Cho and Bengio, 2015, Luong, Pham and Manning, 2015, He et al. 2016, Tu et al. 2016). There are conflicting opinions within academia as to the overall effectiveness of the new NMT model. Costa-jussa et al., (2017) posited that even though NMT had shown improvements in certain language domains within a short time period there is a need for further research.

Furthermore, Castilho et. al, (2017) concluded that the NMT proved to be a suitable model for the three language domains that they evaluated in their research,(i) e-commerce product listings (ii) medical patent domain (iii) educational domain but highlighted that their research was only limited to three language domains.

In addition to the study described above, Bentivogli et. al, (2016) found that some aspects of NMT need further analysis especially in understanding its performance in certain language domains. This shows that there is a lack of research to clearly demonstrate that neural machine translation was a suitable model for all language domains.

Findings from this study have shown that the neural translation model has proved itself as an suitable machine translation model for the service language domain. Having carried out interviews with machine translation experts it became evident that the NMT model has shown to be a more sophisticated model than the SMT model. Commonalties could be found across the narratives of participants that the NMT model brought improved quality in terms of levels of accuracy and fluency. One of the participants even stated that in one of her tests of 500 content segments the human evaluator couldn't tell that he was evaluating NMT translated content. This clearly shows the high level of fluency of the model.

Another of the participants stated that they had performed an extensive test of the NMT model in comparison SMT for customer service content and found the results of the NMT model to be far superior than the SMT model with higher levels of accuracy and fluency.

These statements bear close proximity to the findings in the literature that state that the NMT model has shown significant performance when compared to the SMT model (Bahdanau, Cho and Bengio, 2015, Luong, Pham and Manning, 2015, He et al. 2016).

5.2 - Objective 2 Research Findings

Investigate what is driving the improvement in machine translation specifically in neural machine translation models and how they compare to the SMT models

Cho, van Merrienboer and Bahdanau (2014) stated that the recent increase in neural machine translation model adoption was attributed to the fact that the models can be quickly developed and trained in comparison to statistical machine translation engines due to increased computational power. The findings by this researcher confirmed that computational power is a key driver in the improvement in machine translation and specifically neural machine translation.

The research highlighted that previously building an NMT engine for a specific domain might take up to two months but timeframe has now been reduced to four days . The respondents stated that this was due to the fact that computational cost to build an NMT engine had reduced significantly. The participants also confirmed this improvement is a far distance from the SMT models that could take up to a year to develop.

This corresponds with the findings of previous research that NMT models were much quicker to develop than SMT models (Bahdanau Cho and Bengio, 2015, Cho, van Merrienboer and Bahdanau, 2014).

In the case of artificial intelligence, the literature has shown that AI has had significant impact on the machine translation industry (Bentivogli et al., 2016, Peris, Domingo and Casacuberta, 2017). Their analysis stated the influence of the artificial intelligence has enabled the NMT model to become the most suitable translation model due to the fact that AI is an enabler to significant reduction in post-editing tasks due to increases fluency and accuracy.

The findings from this study agreed with literature with participants confirming that AI has had significant impact on the performance of the NMT models that they are building.

5.3 - Objective 3 Research Findings

How can NMT be used to improve translation quality for the service language domain

The literature shows that in order to gain high levels of translation quality the source of the training data is paramount. Moreover, it has been found that the training from the same domain improved the quality of the translation output (Bertoldi and Federico 2009, Haddow and Koehn 2012, Ma and Way, 2009, Luong Pham and Manning, 2015, Bentivogli et al., 2016, Peris, Domingo and Casacuberta, 2017). It has also been found that for a machine translation model such as SMT or NMT a prerequisite is to capture as much high quality domain specific training data as possible to ensure high quality translation performance and quality (ibid).

The primary research conducted by this researcher partially agreed with the above statements. All of the respondents agreed that training data was critical to NMT engine and the availability of training data at the start was also crucial. However, the key finding from this study that goes beyond the current literature is that domain specificity isn't a crucial factor for the NMT model.

In fact, participants suggest that the broader the training data the better NMT engines will perform in terms of fluency and accuracy. Furthermore, the training data doesn't have to be domain specific, as long as the data is clean the model will benefit.

One could argue that this indicates that the current secondary research data is now redundant and needs to be updated or reviewed in its entirety.

5.4 - Objective 4 Research Findings

What are the key areas that one would need to develop a suitable NMT engine for a service language domain.

As mentioned previously access to clean training data was a key area for developing a NMT engine (Gulcehre et al, 2017, Peris, Domingo and Casacuberta, 2017, Luong, Pham and Manning, 2015).

However, a number of the participants stated that access to clean was important but having access to client specific training data was critical in turning a good NMT engine into a high performing NMT engine.

The implications for businesses is that in order to get the best results out of a NMT engine it is critical that they invest in building client specific training data sets through the creation and maintenance of translation memories. Having this will give them a key corporate advantage as their NMT engines will be high performing and cost effective.

The literature has also stated the importance of improvements NMT models have brought about, specifically improvements in areas such as cost effectiveness and return on investment in comparison to the SMT models (Bahdanau Cho and Bengio, 2015, Luong, Pham and Manning, 2015).

Commonalties emerging from the participants was there has been a reluctance of many organisations to shift to NMT because it took a lot of resources to implement and maintain an SMT models. The participants indicated that the NMT model provided increased business benefit to clients and they have seen increased adoption of the NMT models. A key area that organisations should develop is to build partnerships with machine translation vendors. Having this expertise at hand will help reduce the time to market for products they wish to release into multiple regions.

The section has discussed the key findings in the context of the objectives of the study. The next chapter will discuss conclusions and recommendations.

Chapter 6: Conclusions and Recommendations

6.0 Introduction

This was a small-scale qualitative study that explored the impact of the NMT model in the service language domain. Whilst the findings are not generalisable, the rich data collected in the study has pointed to the importance of understanding the impact that NMT models has had on the service language domain.

6.1 Conclusions

6.1.1 Suitability of the NMT model

The advancements in the NMT model has led to increased fluency and accuracy of translated text for increasingly more language domains. This finding was very much in line with the research of Bentivogli et. al, (2016) which found increased fluency and accuracy of the NMT model in the domains that they researched. This increased suitability of the NMT model has led to increased business benefit for companies using the model but also companies looking to invest in developing the NMT engines for various language domains.

6.1.2 Driving forces in improvement

The technological advances have been enablers for the NMT model to outpace the previous SMT models. Moreover, these advances have allowed for increased adoption of the NMT model as companies can now build NMT engines more quickly but also more cost effectively. These key enablers such as increased computational power and artificial intelligence has allowed the translation industry to develop NMT model which has become extremely attractive to businesses that operate in various language domains.

6.1.3 Training data

As the NMT model continues to be adopted by more businesses it’s critical that companies invest heavily in building clean and client specific training data sets. Building these training data sets is essential to ensure that the translation quality of the NMT models are extremely high but also adaptable to various language domains.

6.1.4 Disruption

Artificial intelligence has had a big impact on the translation industry with some participants stating that the deep learning in the NMT models have made the business models of the translation industry redundant due to increased fluency, accuracy and cost effectiveness of the models. As the NMT models develop further they will continue to reduce the time to market for translation projects by the continued reduction in post editing tasks. The NMT model has not just outpaced the SMT model it has left human translators at risk due to its increased business benefit. It is possible that this disruption will have a profound impact on translators working in the industry with their roles completely changing or becoming redundant.

6.2 Recommendations for further research

Although this research provided some insight into the impact of the NMT model for a specific language domain further research should be conducted to investigate if the NMT is an effective machine translation model for other language domains.

Further research is also needed to understand how to measure the quality of the translation output as more and more companies adopt the model. One of the participants believes that the metric to measure quality the Bilingual Evaluation Understanding (BLEU) developed by Papineni et. al (2002) has become redundant with the NMT model. Further research in this area would be interesting to pursue.

6.3 Recommendations for businesses

The research pointed to the importance of having clean and client specific training data. This is critical to the build and maintenance of a NMT model therefore one recommendation for businesses is to ensure that their translation glossaries and translation are kept up to date.

Furthermore some the participants discussed how they have started to integrate the NMT model into other applications. From a service language domain perspective, it would be interesting to see how businesses would integrate the NMT model into their customer service tools. The use case would be the model being the model being able to translate content in real time for customer service agents as they handle chat conversations in various languages.

7. References

Bahdanau, D., Cho, K. and Bengio, Y. (2015). Neural Machine Translation by Jointly Learning to Align and Translate. *3rd International Conference on Learning Representations*, pp.1-15.

Barrett, L. (2015). Embedding the Google Translate Widget on Taylor & Francis Online. *Editors' Bulletin*, 10(3), pp.48-50.

Bentivogli, L., Bisazza, A., Cettolo, M. and Federico, M. (2016). Neural versus phrase-based machine translation quality: a case study. *Empirical Methods in Natural Language Processing*.

Bertoldi, N. and Federico, M. (2009). Domain Adaptation for Statistical Machine Translation with Monolingual Resources. *Association for Computational Linguistics*, pp.182–189.

Bisazza, A., Pighin, D. and Federico, M. (2011). Chunk-lattices for verb reordering in Arabic–English statistical machine translation. *Machine Translation*, 26(1-2), pp.85-103.

Blumberg, B., Cooper, D. and Schindler, P. (2014). *Business research methods*. London: McGraw-Hill Education.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), pp.77-101.

Brown, P., Della Pietra, S., Della Pietra, V. and Mercer, R. (1993). The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*, 19(2), pp.263-311.

Brownlee, J. (2018). *A Gentle Introduction to Neural Machine Translation*. [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/introduction-neural-machine-translation/ [Accessed 27 Aug. 2018].

Castilho, S., Moorkens, J., Gaspari, F., Calixto, I., Tinsley, J. and Way, A. (2017). Is neural machine translation the new state of art?. *The Prague Bulletin of Mathematical Linguistics,*, 108(1), pp.109-120.

Cho, K., van Merrienboer, B. and Bahdanau, D. (2014). On the Properties of Neural Machine Translation: Encoder–Decoder Approaches.

Chowdhury, H. (2017). *Faster and more accurate: machine translation is nearly here | Financial Times*. [online] Ft.com. Available at: https://www.ft.com/content/d87f6e7a-c560-11e7-b30e-a7c1c7c13aab [Accessed 27 Aug. 2018].

Cohn, T., Duy Vu Hoang, C., Vymolova, E., Kaisheng, Y., Dyer, C. and Haffari, G. (2016). Incorporating Structural Alignment Biases into an Attentional Neural Translation Model.

Collis, J. and Hussey, R. (2009). *Business research*. Basingstoke: Palgrave Macmillan.

Costa-jussà, M., Allauzen, A., Barrault, L., Cho, K. and Schwenk, H. (2017). Introduction to the special issue on deep learning approaches for machine translation. *Computer Speech & Language*, 46, pp.367-373.

Creswell, J. (2014). *Research design*. Thousand Oaks, California: SAGE Publications.

De Bie, T., Turchi, M. and Cristianini, N. (2008). Learning Performance of a Machine Translation System: a Statistical and Computational Analysis. *Association for Computational Linguistics*, pp.35–43.

Ford, M. (2013). Could artificial intelligence create an unemployment crisis?. *Communications of the ACM*, 56(7), p.37.

Foster, G., Isabelle, P. and Plamondon, P. (1997). Target-text mediated interactive machine translation. *Machine Translation*, 12(1/2), pp.175-194.

Frey, C. and Osborne, M. (2013). The future of employment: How susceptible are jobs to computerisation?. *Technological Forecasting and Social Change*, 114, pp.254-280.

Gao, Y. (2018). *IBM100 - Pioneering Machine-Aided Translation*. [online] Www-03.ibm.com. Available at: http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/translation/ [Accessed 24 Aug. 2018].

Ghauri, P. and Gronhaug, K. (2005). *Research methods in business studies*. Harlow: Financial Times Prentice Hall.

Gulcehre, C., Firat, O., Xu, K., Cho, K. and Bengio, Y. (2017). On integrating a language model into neural machine translation. *Computer Speech & Language*, 45, pp.137-148.

Haddow, B. and Koehn, P. (2012). Analysing the effect of out-of-domain data on SMT systems. *Association for Computational Linguistics*, Proceedings of the Third Workshop on Statistical Machine Translation, pp.422–432.

Hasler, E. (2015). *Dynamic Topic Adaptation for Improved Contextual Modelling in Statistical Machine Translation*. Ph.D.

He, W., Zhongjun, H., Hua, W. and Haifeng, W. (2016). Improved Neural Machine Translation with SMT Features. *In Proc. of the Thirtieth AAAI Conference on Artificial Intelligence*, pp.pages 151–157.

Horn, R. (2009). *Researching and Writing Dissertations: A Complete Guide for Business Students*. 1st ed. Chartered Institute of Personnel and Development, p.109.

Koehn, P. (2014). *Statistical machine translation*. 1st ed. Cambridge: Cambridge University Press.

LaGarda, A., Alabau, V., Casacuberta, F., Silva, R. and Dıaz-de-Liano, E. (2009). Statistical Post-Editing of a Rule-Based Machine Translation System. *Association for Computational Linguistics*.

Lardinois, F. (2018). *Google brings offline neural machine translations for 59 languages to its Translate app*. [online] TechCrunch. Available at: https://techcrunch.com/2018/06/12/google-brings-offline-neural-machine-translation-for-59-languages-to-its-translate-app/?guccounter=1 [Accessed 27 Aug. 2018].

Levy, F. and Murnane, R. (2005). *The new division of labor: How computers are creating the next job market*.

Lewis-Kraus, G. (2018). *The Great A.I. Awakening*. [online] Nytimes.com. Available at: https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html [Accessed 27 Aug. 2018].

Luong, T., Pham, H. and Manning, C. (2015). Effective approaches to attention- based neural machine translation. *Empirical Methods in Natural Language Processing*.

Ma, Y. and Way, A. (2009). Bilingually Motivated Word Segmentation for Statistical Machine Translation. *ACM Transactions on Asian Language Information Processing*, 8(2), pp.1-24.

Madhavan, R. (2018). *Machine Translation - 14 Current Applications and Services*. [online] TechEmergence. Available at: https://www.techemergence.com/machine-translation-14-current-applications-and-services/ [Accessed 24 Aug. 2018].

Peris, Á., Domingo, M. and Casacuberta, F. (2017). Interactive neural machine translation. *Computer Speech & Language*, 45, pp.201-220.

Saunders, M., Lewis, P. and Thornhill, A. (2012). *Research methods for business students*. Harlow (Essex): Pearson.

Thirgood, J. and Johal, S. (2017). Digital Disruption. *Economic Development Journal*, 16(2), pp.25-32.

TURING, A. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, LIX(236), pp.433-460.

Walliman, N. (2005). *Your research project*. London: Sage Publications.

Zhang, J. (2017). *Domain Adaptation for Statistical Machine Translation and Neural Machine Translation*. Ph.D. Dublin City University.

8. Appendices

Appendix 1: Participant Consent Form

Study Title: An exploration into the impact of NMT on a service language domain.

Consent Form

Thank you for reading the information sheet about the interview sub-study. If you are happy to participate then please complete and sign the form below. Please tick the boxes below to confirm that you agree with each statement:

Please Tick box:

I confirm that I have read and understood the information sheet dated __________ and have had the opportunity to ask questions. ☐

I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question or questions, I am free to decline. ☐

I understand that my responses will be kept strictly confidential. I understand that my name will not be linked with the research materials, and will not be identified or identifiable in the report or reports that result from the research. ☐

I agree for this interview to be audio-recorded. I understand that the audio recording made of this interview will be used only for analysis and that extracts from the interview, from which I would not be personally identified, may be used in any conference presentation, report or journal article developed as a result of the research. I understand that no other use will be made of the recording without my written permission, and that no one outside the research team will be allowed access to the original recording. ☐

I agree that my anonymised data will be kept for future research purposes such as publications related to this study after the completion of the study. ☐

I agree to take part in this interview. ☐

______________________	______________	__________________
Name of participant	Date	Signature
______________________	______________	__________________
Principal Investigator	Date	Signature

To be counter-signed and dated electronically for telephone interviews or in the presence of the participant for face to face interviews

Copies: *Once this has been signed by all parties the participant should receive a copy of the signed and dated participant consent form, and the information sheet. A copy of the signed and dated consent form should be placed in the main project file which must be kept in a secure location.*

Appendix 2: Participant information sheet

PARTICIPANT INFORMATION SHEET

Study Title: Research Thesis

Aim of the study: To understand the impact of NMT in a service language domain.

What will you have to do?
If you agree to take part in the study you will be asked to participate in one face-to-face or telephone interview with the researcher as named below. The interview will take approximately 45 minutes.

What are the risks to you? Taking part carries no obvious risks to you. If you do want to answer any of the you are free to decline to answer.

What if I do not want to take part?

There will be no negative consequences to you should you decline to take part.

What if I change my mind during the study?

There are no negative consequences if you decide that you don't want to finish answering the questions – just let the researcher know.

What happens to the data? The names of all participants will be strictly confidential. In the writing up of the study all names will be changed to ensure anonymity. This study will be submitted for evaluation to the Director of the School of Business in National College of Ireland and to an external examiner as part of the researcher's course work. It may also be chosen for publication in a journal. The names of all participants will remain strictly confidential at all times.

Where can I get further information? If you have any further questions about the study or if you want to opt out of the study now or at any time in the future, please contact:

Robert Callaghan (Researcher)
Tel: 086 00225126

Or

David Hurley - School of Business Centre, NCI, Dublin
David.hurley@staff.ncirl.ie