

Towards Analysis of Biblical Entities and Names using Deep Learning

Mikolaj Martinjak^{1*}, Davor Lauc², Ines Skelac³

Faculty of Philosophy and Religious Studies, University of Zagreb, Zagreb, Croatia^{1,3}
Faculty of Humanities and Social Sciences, University of Zagreb, Croatia²

Abstract—Scholars from various fields have studied the translations of the Bible in different languages to understand the changes that have occurred over time. Taking into account recent advances in deep learning, there is an opportunity to improve the understanding of these texts and conduct analyses that were previously unattainable. This study used deep learning techniques of NLP to analyze the distribution and appearance of names in the Polish, Croatian, and English translations of the Gospel of Mark. Within the scope of social network analysis (SNA), various centrality metrics were used to determine the importance of different entities (names) within the gospel. Degree Centrality, Closeness Centrality, and Betweenness Centrality were leveraged, given their capacity to provide unique insights into the network structure. The findings of this study demonstrate that deep learning could help uncover interesting connections between individuals who may have initially been considered less important. It also highlighted the critical role of onomastic sciences and the philosophy of language in analyzing the richness and importance of human and other proper names in biblical texts. Further research should be conducted to produce more relevant language resources, improve parallel multilingual corpora and annotated data sets for the major languages of the Bible, and develop an accurate end-to-end deep neural model that facilitates joint entity recognition and resolution.

Keywords—Bible; deep learning; gospel of mark; natural language processing; social network analysis

I. INTRODUCTION

The Bible is a significant religious text that has been translated into many languages and is one of the world's most widely studied ancient texts. However, the translations of the Bible exhibit substantial variation across the languages into which they have been translated, and these translations have undergone numerous revisions over the years. As a result, studying the translations of the Bible in different languages and understanding the changes that have taken place over time are essential for scholars in several fields, including theology, linguistics, religious studies, and history.

With recent advances in deep learning, there is an opportunity to improve understanding of these texts and conduct previously impossible analyses. Deep learning techniques are still not commonly used in this field of research, but they can help scholars analyze and interpret biblical text by enabling them to process large amounts of data and identify connections and information that cannot be identified using traditional methods. In this context, one application of deep learning is the use of natural language

processing (NLP) to analyze the language used in the text. Natural language processing can be used to identify patterns in the use of words and phrases, as well as in the syntax and grammar of the text.

Recent research has explored the use of SNA to study language and culture. Social network analysis is a useful tool for identifying patterns in large data sets, and has been applied to various domains, including linguistics, anthropology, and sociology. It is also a useful technique that can be used to uncover patterns and relationships between Bible translations by analyzing the frequency of words and phrases in different translations.

In this case study, the use of deep learning techniques is explored to analyze the distribution and appearance of names in the Polish, Croatian, and English translations of the Bible. For this research, data from the 100-language parallel corpora created by Christos Christodouloupoulos and Mark Steedman are used [1]. When multiple versions of the Bible were available, the creators of the corpora preferred to use the oldest translation available. For example, the King James Version was used for the English language.

Furthermore, this analysis uses the methodology of distant reading [2]. It aims to analyze large collections of texts using computational tools and techniques, as this allows for the generalization and identification of patterns that would otherwise be impossible to detect. The present case study uses SNA techniques to investigate the relationships and interactions within a network of entities in the selected book of the Bible. When SNA techniques are applied to large data sets of translations, insights can be gained into the relationships between translations and the religious and linguistic factors that influence the translation process.

This study aims to strengthen the robustness of biblical text analysis across multiple languages; therefore, it should be viewed as a first step or pilot research. As a starting point, the Gospel of Mark is chosen. The Gospel of Mark's Footprint will be expanded to other languages and subsequently to the entire New Testament and later to the entire Bible. The goal is to enhance understanding of the distribution of names in the biblical text and provide a framework for analyzing other interdisciplinary aspects, including theological, linguistic, and philosophical ones. Finally, during this analysis, different philosophical questions about the status of the names appeared and have been discussed using a framework from the analytical philosophical tradition.

In this paper, after the review of existing research and available language resources, the methods and results of entity recognition and resolution in a selected Gospel will be presented. Finally, the obtained results and the possibilities of future research will be discussed.

II. LITERATURE OVERVIEW

A. *About the Gospel of Mark*

The Gospel of Mark was written in the first century for a general audience of Christians, although scholars debate the exact audience and purpose of the gospel [3]. It is the shortest of the four canonical gospels and is believed to have been written between AD 65 and 70. Hence, it is one of the earliest written accounts of Jesus Christ's life and teachings.

In the Gospel of Mark, Jesus is presented as the Son of God and the long-awaited Messiah (Christ) of Jewish prophecy. Throughout the gospel, Jesus performs miracles and teaches with authority, attracting crowds of followers while challenging the religious and political authorities of his time. This gospel emphasizes Jesus' role as a suffering servant who will be rejected, betrayed, and crucified but will also rise from the dead and bring salvation to humanity. Specifically, Mark portrays Jesus as a human being with a divine mission who is uniquely qualified to fulfill the promises and prophecies of the Hebrew Bible. Jesus' identity is gradually revealed throughout the gospel as he heals the sick, forgives sins, calms storms, feeds the multitudes, and confronts demons. At Jesus' baptism, a voice from heaven proclaims, "...Thou art my beloved Son, in whom I am well pleased." (Mk 1:11) [4], confirming the divine identity of Jesus. Despite Jesus' powerful teachings and miraculous deeds, many people rejected him, including the religious leaders of his day, who saw him as a threat to their authority. This leads to Jesus' arrest, trial, and crucifixion, but also sets the stage for his ultimate triumph over death and evil powers.

The Gospel of Mark contains a distinctive theology that emphasizes certain key themes and ideas. There are a few theological themes found in this Gospel. For this research, the most prominent theme is the Messianic Secret. The Gospel of Mark portrays Jesus as the Messiah, while emphasizing that his identity is often hidden or a secret. This is because Jesus repeatedly commands those he has healed or exorcised to tell no one about what he has done, and he often speaks in parables that are difficult for his disciples and others to understand.

The reference for all names mentioned in the text is not obvious from the text. Sometimes it relates to our knowledge of the Bible, especially other synoptic gospels [5]. The Gospel of Mark gives us names such as Jesus, Simon, John, James, etc. However, who was James? It should be concluded that from the text itself. Another challenge is that there is more than one James, and, in some cases, it is not so evident to which person the text refers. The same thing happened with the name John, where it can distinguish at least John the Baptist and John the Apostle, and with the other names as well. The solution for distinguishing is mentioned in the Results and Discussion section.

B. *The Analysis of NLP Resources for Bible Studies*

There are various applications of NLP resources for Bible studies thus far. Natural language processing tools can be used to analyze the language and style of the Bible, which can provide insights into the authorship and composition of different biblical texts. However, using NLP tools for biblical studies has several challenges and limitations. For example, the complexity and diversity of biblical texts makes it difficult to create NLP resources that can capture all the nuances and subtleties of the language. In addition, the cultural and historical contexts in which the Bible was written can be difficult to capture using NLP tools alone and may require additional theological expertise. Therefore, NLP tools should be used in conjunction with other methods and expertise, not in place of close reading and interpretation of texts.

Büchler and Mellerin [6] discuss the use of NLP resources for Bible studies. This article provides an overview of several NLP tools and resources designed specifically for analyzing biblical texts, such as morphological analyzers, part-of-speech taggers, syntactic parsers, and named entity recognition systems. These tools enable scholars to identify patterns of language use, track the evolution of specific words and phrases, and uncover hidden connections between different parts of the Bible. The authors also discuss the potential benefits of using NLP tools for biblical studies. For example, they propose that NLP can help scholars identify allusions and quotations from other texts in the Bible, which can shed light on the cultural and historical contexts in which the Bible was written.

Another topic that can be investigated using NLP is authorship attribution. According to [7], attribution of authorship involves analyzing various linguistic features of a text to determine the probability that a given author wrote the text. The authors discuss challenges that arise when applying authorship attribution to the Bible, such as the use of pseudonyms and the difficulty of obtaining enough training data for accurate analysis and provide examples of how authorship attribution has been used to analyze different parts of the Bible, such as the Psalms and the Pauline epistles. Despite these challenges, authorship attribution can provide valuable insights into authorship and the historical context of biblical texts. Thus, by identifying the authors of different parts of the Bible, scholars and researchers could gain a deeper understanding of the historical and cultural contexts in which these texts were written.

In addition, several studies have been conducted to compare Bible translations in different languages. Using SNA techniques, a study examined the frequency of words in the English and German translations of the Bible. The study found that some words were used more frequently in one translation than in the other, which could be attributed to differences in language structure, culture, and translation choices [8]. By comparing translations in different languages, scholars can better understand the choices made by translators and the cultural and linguistic factors that influence those choices.

As stated previously, NLP tools are used for various analyses of the biblical text; however, none of the existing studies conducted a comparative analysis of personal names in

different languages. This analysis can have several outcomes: 1) It can show linguistic (phonetic) differences between Bible translation and different languages, allowing for a language distance analysis; 2) It can provide a parallel philosophical and theological perspective to the name analysis; 3) It can use SNA to show differences in social connections between different translations of the Bible, which can help in determining some ambiguous places in understanding personal relations in the Bible.

III. METHODS

A. Distant Learning Method of Literary Analysis

In recent years, distant reading, a literary analysis method that involves analyzing large bodies of texts using computational techniques, has been applied to biblical studies. This method has been used to identify patterns and trends in the use of language and themes across multiple biblical texts, shedding new light on the historical and cultural contexts in which these texts were written. One of the key aspects of distant reading is the use of quantitative methods to measure and compare literary phenomena. Moretti [2] suggests that we should examine the “big picture” of literary history by analyzing large data sets and mapping out the evolution of genres, themes, and styles over time. A general goal of distant reading is to uncover broad patterns and trends in literature that can reveal new insights and understandings about literary history, genres, and themes [9].

In doing so, patterns that may not be visible at the level of individual texts or authors can be identified. Another important aspect of distant reading is the use of computational tools to process and analyze textual data. Moretti argues that we should go beyond traditional methods of literary analysis, such as close reading and interpretation, and embrace the potential of digital technologies to handle large amounts of data [2]. By using tools such as text mining, topic modeling, and network analysis, hidden patterns and connections in the literature that would otherwise be impossible to discern can be uncovered.

However, distant reading also faces several challenges and limitations. One of the main criticisms of distant reading is that it prioritizes quantitative analysis over qualitative interpretation and reduces literary texts to mere data points. Despite this limitation in this study, the results obtained using this methodology are always discussed and verified using traditional theological and linguistic methodologies.

B. Construction of the Entity Graph

The entity graph is built in two phases. In the first phase, named entities are recognized and classified using the ByT5 neural model [10] fine-tuned in the relevant parts of the multilingual named entity recognition (NER) data set. The results were manually evaluated and corrected, and the data set will be used to fine-tune the NER model in the future. The second phase was manual disambiguation and classification of the recognized named entities. The graph (social network) was constructed based on counts of co-occurrence of the entities of type person in the verses.

C. Social Network Analysis Metrics

Within SNA, various centrality metrics are used to determine the importance of network entities. In this research, degree centrality, closeness centrality, and betweenness centrality were used, each providing unique insights into the structure of the network.

Degree Centrality is a simple metric that counts the number of direct connections (edges) an entity has with other entities in the graph. Higher-degree centrality values indicate a higher number of connections. This metric can be used to identify nodes that are well connected and potentially influential within the network. Closeness centrality measures how close a node is to all other nodes in the network. It is calculated by taking the inverse of the sum of the shortest path distances between a node and all other nodes in the network. A higher closeness centrality value indicates that an entity is more central and can reach other nodes in the graph more quickly. Betweenness centrality, a related metric, measures the extent to which an entity acts as a bridge to other entities. It indicates how often a node lies on the shortest path between pairs of other nodes [11].

IV. RESULTS

An entity graph (a social network) created by the described method reveals interesting insights into relationships among people mentioned in the gospel. The selected centrality measures – degree centrality (D), closeness centrality (C), and betweenness centrality (B) – for the English, Croatian, and Polish languages are presented in Table I.

The difference between languages and the apparent difference in numbers on individuals and their relationship or closeness demonstrated in Table I is in significant part due to the use of pronouns. Each language will use a different number of pronouns, corresponding to the development and flow of the language. SNA metrics show that Jesus is a central figure, but also give some surprising results, such as that Mary Magdalene is closer in a relationship and connections than Mary, mother of Jesus. Yet, it is also due to how the gospel was written.

Differences and similarities between centrality measures are easier to interpret when visualized using a spring graph layout, as shown in Fig. 1, Fig. 2, and Fig. 3.

The graphs in the three languages show that Jesus is a central figure of the gospel. Obviously, many connections led to Jesus, especially those closest to him – for example, John the Baptist, Simon Peter, John and James Zebedee, and others. When the names of Timaeus and Bartimaeus are looked at, it is noticed that they are closer to Jesus in the Croatian language than in others. In that sense, they are not directly connected to Jesus in English or Polish but are in the Croatian version of the graph. The main reason for this occurrence is that Jesus was specifically mentioned by name in the Croatian version of the text. Another intriguing part of the graph is the name of Simon, who is in the text distinguished as a Cyrenian and his two sons, who have no connection to Jesus in the graph. However, according to knowledge of other gospels and reading of the Gospel of Mark, it is known that Simon was the one who assisted Jesus in carrying his cross. Mary Magdalene,

Mary the mother of James and Joses, and Salome all experienced similar events. There are no direct connections to Jesus in the graphs, even though it is known from other resources and standard interpretations of the Bible that there is a connection. Aside from the obvious connections, it was interesting to see the link between Jesus, Barabbas, Pilate, and Joseph of Arimathea, who are all directly connected to Jesus. Note that Joseph is only identified by his first name in the Croatian and Polish versions of the graph.

In the King James version of the English translation, which is used in this analysis, the name James Zebedee is used, while in the Croatian and Polish translations, the name is Jakob or Jakov. In graphs, it is evident that the same person is in question, since there is a connection between him, his brother John, and his father (signed just as Zebedee). For some other figures, a 'new' or a bit untraditional method was used. For the research, the surnames that are not present in the gospel were added to distinguish similar names that refer to different people.

TABLE I. SELECTED SNA METRICS (D – DEGREE, C – CLOSENESS, B - BETWEENNESS) ORDERED BY THE BETWEENNESS CENTRALITY IN THE ENGLISH TRANSLATION

	ENTITY	ENGLISH			CROATIAN			POLISH		
		D	C	B	D	C	B	D	C	B
0	Jesus Christ	0.29	0.38	0.3	0.43	0.46	0.25	0.28	0.33	0.15
1	Andrew - brother of Peter	0.32	0.34	0.24	0.09	0.22	0	0.08	0.17	0
2	John the Baptist	0.15	0.28	0.12	0.14	0.31	0.09	0.14	0.24	0.07
3	Simon Peter	0.18	0.38	0.11	0.17	0.31	0.02	0.17	0.24	0.01
4	Mary Magdalene	0.09	0.27	0.08	0.09	0.27	0	0.08	0.22	0.05
5	James Zebedee	0.15	0.37	0.07	0.14	0.3	0.02	0.14	0.24	0.02
6	John Zebedee	0.15	0.37	0.07	0.14	0.3	0.02	0.14	0.24	0.02
7	Elias	0.12	0.32	0.02	0.11	0.31	0.01	0.11	0.24	0.01
8	Zebedee	0.06	0.25	0	0.06	0.2	0	0.06	0.16	0
9	Alphaeus	0.24	0.26	0	0	0	0	0	0	0
10	Bartholomew	0.24	0.26	0	0.17	0.17	0	0.22	0.22	0
11	James Alphaeus	0.24	0.26	0	0.17	0.17	0	0.22	0.22	0
12	Matthew	0.24	0.26	0	0.17	0.17	0	0.22	0.22	0
13	Philip	0.24	0.26	0	0.17	0.17	0	0.22	0.22	0
14	Simon the Canaanite	0.24	0.26	0	0.17	0.17	0	0.22	0.22	0
15	Thaddaeus	0.24	0.26	0	0.17	0.17	0	0.22	0.22	0
16	Thomas	0.24	0.26	0	0	0	0	0.22	0.22	0
17	Joses Nazarean	0.09	0.09	0	0.11	0.11	0	0.11	0.11	0
18	Juda Nazarean	0.09	0.09	0	0.11	0.11	0	0.11	0.11	0
19	Mary mother of Jesus	0.09	0.09	0	0.11	0.11	0	0.11	0.11	0
20	Simon Nazarean	0.09	0.09	0	0.11	0.11	0	0.11	0.11	0
21	Herod the King	0.09	0.21	0	0.09	0.21	0	0.08	0.17	0
22	Herodias	0.09	0.21	0	0.09	0.21	0	0.08	0.17	0
23	Philip, brother of Herod	0.09	0.21	0	0.09	0.21	0	0.08	0.17	0
24	Moses	0.09	0.3	0	0.09	0.28	0	0.08	0.22	0
25	Bartimaeus	0.03	0.03	0	0.06	0.27	0	0.03	0.03	0
26	Timaeus	0.03	0.03	0	0.06	0.27	0	0.03	0.03	0
27	Pontius Pilate	0.09	0.26	0	0.09	0.27	0	0.08	0.21	0
28	Barabbas	0.06	0.26	0	0.06	0.27	0	0.06	0.21	0
29	Alexander of Simon C.	0.06	0.06	0	0.06	0.06	0	0.06	0.06	0
30	Rufus of Simon C.	0.06	0.06	0	0.06	0.06	0	0.06	0.06	0
31	Simon Cyrenian	0.06	0.06	0	0.06	0.06	0	0.06	0.06	0
32	Mary, mother of James and Joses	0.06	0.2	0	0.09	0.27	0	0.06	0.16	0
33	Salome	0.06	0.2	0	0.09	0.27	0	0.06	0.16	0
34	Joseph of Arimathea	0.06	0.26	0	0.06	0.27	0	0.06	0.21	0
35	David the King	0	0	0	0.03	0.26	0	0	0	0
36	Andrew	0	0	0	0.17	0.17	0	0.22	0.22	0
37	James Nazarean	0	0	0	0.11	0.11	0	0.11	0.11	0

It is common for biblical names to be translated differently in different languages, depending on the linguistic and cultural traditions of each language. James has become a common English name and is often used to refer to the biblical figure. Many English-speaking readers may not recognize that the name Jacob refers to the same person. However, some English translations use the name Jacob instead of James, particularly in recent translations that seek to be more faithful to the original Hebrew and Greek texts.

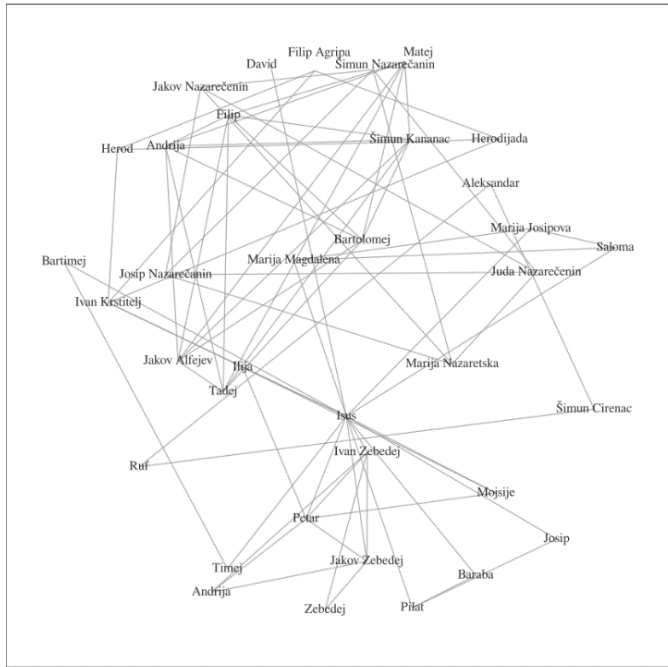


Fig. 1. Entity graph for English language.

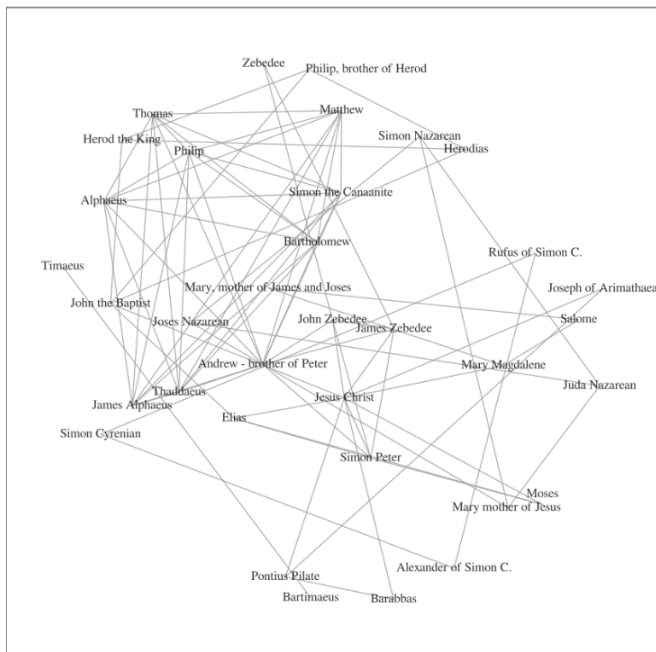


Fig. 2. Entity graph for Croatian language.

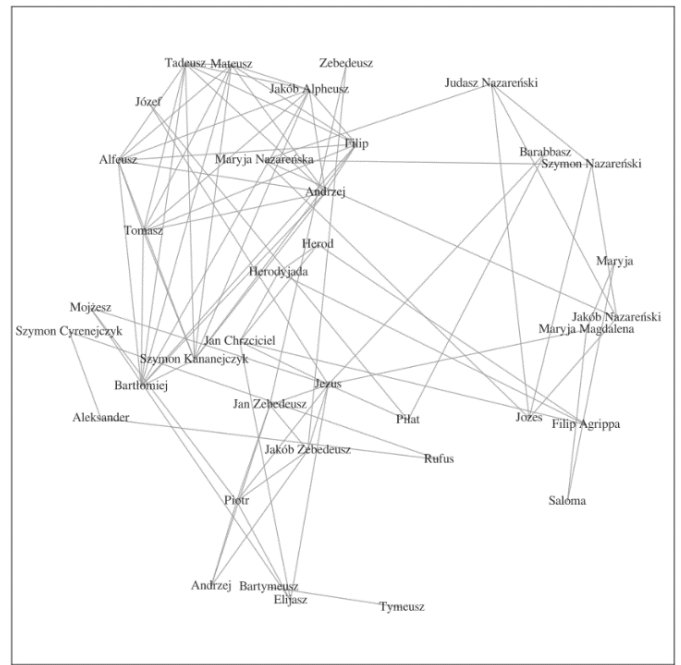


Fig. 3. Entity graph for Polish language.

V. DISCUSSION

A. Theological Perspective

During the manual evaluation and correction process, the greatest diversity came from the differences in the languages and the use of names in specific languages. Some Bible translations do not follow the phonetic development of an original name but instead use more common names that evolved over time and even morphed into different names. However, some names needed to be differentiated even more in the text itself. The first task was to distinguish between the names of places, towns, regions, and individuals. It was also necessary to differentiate between functions such as Caesar, Pharisee, etc., and individuals since the program recognized them as specific names of people rather than functions. In that sense, a boat (pol. *łódź*) was recognized as a proper name in the Polish version of the text, probably because there is a town with the same name in Poland.

Furthermore, the main challenge in identifying individuals in the Gospel is that the text provides only a first name, occasionally giving a last name or what we would call, in today's terms, a last name [12]. In biblical terms, they almost always represent a connection with a particular person in parent-child relationships. The name Bartimaeus, the son of Timaeus, is the most notable and precise connection between father and son (Mk 10:46-52) [4]. It is a Greek name that means "son of Timaeus." "There are no similar connections in other places of the gospel. Instead, one can see "son of Alphaeus," "son of Zebedej," or even "Son of David," which is sometimes applied to "Jesus". The problem emerged because those connections are not evident in the text unless manually marked. As a result, last names were required to make clearer distinctions. Moreover, the apostles James and John were given the surname Zebedej, but that did not resolve the issue, especially since there were other people named

James. One is known as “James Alphaeus” and is sometimes referred to as “the lesser”. Additionally, he is supposed to be Levi’s (Mathew) brother. The information on Levi as Mathew is reconstructed from other synoptic gospels and textual connections that are not obvious from a single reading of the story.

There is also a problem with the name Simon, which is given to Peter as he becomes a disciple, but it is also the same name of individuals who are difficult to distinguish. In some ways, one is unsure if those are two different people or if they are referred to differently, as in the case of Levi.

In our analysis of the text, it was important to differentiate Christ as a name from Christ as a reference to Jesus, as Ehrman concludes [13]. In other words, a problem emerged during the name analysis because it was necessary to distinguish between Christ as a name and Christ as a function. That is, in terms of theology, Ehrman’s work made sense to this. The concept of the Messianic secret in the Gospel of Mark has important implications for understanding the relationship between Jesus and his disciples, the apostles.

B. Philosophical Perspective

Except for the beginning of the Gospel of Mark, where the syntagm ‘Jesus Christ’ is used, it cannot be sure if Christ refers to Jesus in other places. However, it is clear in some places that it does not or can be questioned. This is a challenge for linguistic analysis because the question is: when Christ refers to Jesus, is it only when there is a syntagm Jesus Christ, and is it possible that Christ does not refer to any specific person in some places?

The famous location is (Mk 8: 29), where Jesus asks his disciples: “...But whom say ye that I am? And Peter answered and saith unto him, Thou art the Christ.” [4] As a result, the syntagm Jesus Christ is not the only place where Christ refers to Jesus.

For linguistic analysis, “Christ” can be interpreted as a function in Frege’s terms [14], with Jesus serving as the object that fulfills the function of being “Christ.” In the case of “Christ” in the Gospel of Mark, one could interpret it as a function that assigns the value of being the Messiah, the Savior, or some other similar role to a particular object. Depending on the context, the object that fulfills this function could be Jesus or another individual. However, the function would remain the same regardless of the specific object that performed it.

If Jesus is the only object that fulfills this function, then the function would have a unique value, and any use of the term “Christ” would necessarily refer to Jesus. However, Frege’s theory of functions allows for the possibility of multiple values that fulfil a function. The specific number and nature of these values would be determined by the context and interpretation of the term “Christ.”

If one wants to limit “being a Christ” to one person, then it is better to understand “Christ” as a name and not a function. According to Kripke [15], names are rigid designators; therefore, the meaning of a name is fixed and independent of any specific description or property of the individual. With

this definition, it can be said that Christ, as a rigid designator, has the same fixed meaning as the name Jesus when it is known that these two names refer to the same person (compare [16, 17]). This meaning, or the object of reference, could be Jesus or another figure, such as the Messiah or the Savior. Still, the name “Christ” would always refer to the same individual, regardless of the specific properties or descriptions used to describe them. Kripke’s theory also allows the possibility of reference failure, in which a name does not refer to any individual in the world. This could happen if the name is used in a context where no individual matches the description associated with it (for example, before Jesus was born) or if it is unknown that Christ has the same reference as Jesus. In the case of “Christ” in the Gospel of Mark, it is also possible that the name Christ does not refer to any specific individual in some contexts, either because it is used in a general sense or because there is ambiguity or uncertainty about its referent.

Therefore, “being a Christ” can be understood as a function in Frege’s terms. Kripke’s theory of reference provides a more useful framework for understanding how names refer to individuals in the world when this individual is unknown or uncertain.

VI. CONCLUSION

From a theological perspective, this study demonstrates that in some cases it is difficult to distinguish an individual from the mission or even from another individual. Almost everyone in the world knows who Jesus is and perhaps who Peter or Judas is. Still, this research showed that some interesting connections between individuals who seemed less important could be discovered through deep learning. Even with this small extract of the text and only a few languages, it will become even more prominent on a larger scale.

Given the importance and influence of Bible, the relative scarcity of available language resources is surprising. Therefore, our future research will initially focus on creating more relevant resources, starting with improving parallel multilingual corpora and annotated data sets for all the major languages of the Bible. These resources are necessary to train more accurate multilingual language models for named entity recognition and entity resolution. Based on these data sets, the intention is to develop an accurate end-to-end deep neural model for joint entity recognition and resolution, which would facilitate numerous other research projects in digital humanities, theology, and related fields.

Another unexpected early finding was that the newer NER models performed worse than older models on biblical texts, contrary to the standard benchmarks, which should be thoroughly researched and explained.

Further research using other metrics and techniques from SNA in studying the interaction of entities in the Bible and validating those findings against traditional research should yield interesting results.

Finally, given the richness and importance of human and other proper names in biblical texts, several interesting studies in onomastic sciences and sociolinguistics could be conducted, ranging from analyses of the similarity and differences of

biblical names across different languages to the creation of alternative similarity matrices among languages.

REFERENCES

- [1] C. Christodouloupoulos and M. Steedman, "A massively parallel corpus: the Bible in 100 languages," *Language Resources and Evaluation*, Language Resources and Evaluation, 2015.
- [2] F. Moretti, *Distant reading*, Verso Books, 2013.
- [3] L. W. Hurtado, *The Earliest Christian Artifacts: Manuscripts and Christian Origins*, Wm. B. Eerdmans Publishing Co., 2008.
- [4] *The Holy Bible: King James Version*, Dallas, TX: Brown Books Publishing, 2004.
- [5] J. Rubalcaba, *Who's Who in the Bible: Everyone You Need to Know from the Old and New Testaments*, National Geographic, 2018.
- [6] M. Büchler and L. Mellerin, "JDMDH Special Issue on Computer-Aided Processing of Intertextuality in Ancient Languages," *Journal of Data Mining and Digital Humanities*, 2017, pp. 1-10.
- [7] D. E. Mills, *Authorship attribution applied to the Bible*, Texas Tech University, 2003.
- [8] I. Goodfellow, Y. Bengio and A. Courville, *Deep learning*, MIT press, 2016.
- [9] L. M. E. Goodlad, "A Study in Distant Reading: Genre and the Longue Durée in the Age of AI," *Modern Language Quarterly*, vol. 81, no. 4, 2020, pp. 491-525.
- [10] L. Xue, A. Barua, N. Constant, R. Al-Rfou, S. Narang, M. Kale, A. Robert and C. Raffel, "Byt5: Towards a token-free future with pre-trained byte-to-byte models," *Transactions of the Association for Computational Linguistics*, vol. 10, 2022, pp. 291-306.
- [11] S. P. Borgatti and D. J. Brass, *Social Networks at Work*, Routledge, 2019.
- [12] D. Hey, *The Oxford Guide to Family History*, Oxford University Press, 2002.
- [13] B. D. Ehrman, *The Orthodox Corruption of Scripture: The Effect of Early Christological Controversies on the Text of the New Testament*. Updated and with a new afterword ed., Oxford University Press, 2011.
- [14] G. Frege, "Function and Concept," in *Translations from the philosophical writings of Gottlob Frege*, Basil Blackwell, 1960, pp. 21-35.
- [15] S. Kripke, *Naming and Necessity*, Harvard University Press, 2001.
- [16] J.T. Turner. Hylemorphism, rigid designators, and the disembodied 'Jesus': a call for clarification. *Religious Studies*, vol. 57, no. 2, 2021. pp. 193-208.
- [17] W. Wood. *Analytic theology and the academic study of religion*. Oxford University Press, USA. 2021.