

Generating Textual Summaries from the Bibliographies Contained in Scientific Literature

Master's Thesis

Otto Mättas

Supervisor:

Prof. Kees van Deemter

24. February 2024

Master Artificial Intelligence
Faculty of Science
Utrecht University
The Netherlands

Abstract

This thesis explores the development and evaluation of a method for generating computer-generated textual summaries of bibliographies in scientific papers and conference proceedings. This method aims to enhance the utility of bibliographies by providing informative overviews that encapsulate the essence of cited sources. By integrating content analysis with metadata considerations, the method distills key information such as temporal trends, authorship patterns, research themes, and source characteristics into coherent summaries. This approach aids researchers in comprehending the contextual and thematic relevance of bibliographies and streamlines the research process by mitigating the need for exhaustive perusal of sources.

The research investigates features and elements that bolster the efficacy of bibliography summaries, guided by the hypothesis that such summaries can aid scientists in their daily work. Through a mixed-methods evaluation involving qualitative interviews and a quantitative questionnaire, the study elucidates the preferences and requirements of academic professionals, revealing an inclination towards summaries that balance conciseness with comprehensive insight.

Future avenues for this research encompass advanced topic analysis to unearth deeper thematic connections, incorporation of citation networks to elucidate scholarly discourse dynamics, and the development of customisable, interactive summary tools to cater to diverse user needs. Additionally, broadening user evaluations and integrating the summarisation method into existing research platforms are identified as critical steps towards enhancing the accessibility and impact of this method. Ultimately, this research posits that computer-generated bibliographic summaries hold the potential to transform literature review practices and foster more efficient scientific inquiry.

Acknowledgements

I would like to express my deepest gratitude to all those who have supported and contributed to the completion of this research project.

First and foremost, I would like to thank my advisor Prof. Kees van Deemter for their invaluable guidance, expertise, and unwavering support throughout this journey. Their insightful feedback, constructive criticism, and encouragement have been instrumental in shaping this research.

I am also grateful to my colleagues who have provided valuable insights, stimulating discussions, and assistance during the various stages of this project. Their contributions have greatly enriched the quality of this research.

I extend my appreciation to the participants who generously dedicated their time and shared their insights during the interviews and evaluations conducted for this research. Their valuable input has significantly contributed to the development and evaluation of the proposed method. In alphabetical order, the researchers who participated were Prof. Albert Gatt, Prof. Alex C. Telea, Prof. Alvo Aabloo, Prof. Argo Rosin, Prof. Fabiano Dalpiaz, Prof. Floris J. Bex, Dr. Juhan-Peep Ernits, Prof. Mehdi M. Dastani, Dr. Natasha A. Alechina, Prof. Pinar Yolum Birbil, Prof. Rens van de Schoot, Dr. Robbert Jan Beun, Prof. Tanel Tammet, Dr. Till Miltzow.

I would like to acknowledge the academic and research institutions, namely Utrecht University, Tallinn University of Technology, and Tartu University, whose work and resources have been instrumental in the implementation and evaluation of this research.

Furthermore, I would like to express my gratitude to my friends and family for their continuous support, understanding, and encouragement throughout this research endeavour. Their belief in my abilities and their presence in my life have been a constant source of motivation.

Lastly, I would like to express my heartfelt appreciation to all the researchers, scholars, and authors whose contributions form the foundation of scientific literature. Their dedication to advancing knowledge and their commitment to sharing their findings have made this research possible. Special thanks to Kittipitch Kuptavanich for his valuable input and support.

This work would not have been possible without the collective effort and support of all those mentioned above, as well as countless others who have contributed in various ways. I am sincerely thankful to each and every one of them.

1. Introduction	7
2. Natural Language Generation	8
2.1.Evaluation in Natural Language Generation	8
3. Gathering Requirements for Summary Generation from Interviews with Scientists	10
3.1.Interviewing Methods	10
3.2.Participants	10
3.3.Findings	10
3.4.Insights	12
3.5.Implications for Generating Textual Summaries	12
3.2.Conclusion about Requirements	14
4. Towards an Improved Algorithm for Generating Textual Summaries from Bibliographies	15
4.1.Algorithm Evaluation	15
5. Design and Implementation of the Model	17
5.1.Simulated Perceptual Grouping: Enhancing Human-Computer Interaction	17
5.2.Pre-processing Stage	17
5.3.Information Extraction	18
5.4.Summary Generation	20
6. Method Evaluation	23
6.1.Evaluating Method through Interviews with Scientists	23
6.2.Evaluating Method through Web Questionnaire	23
6.3.Evaluating Method through Interviews with Developers	24
6.4.Integrating Evaluation Insights	26
6.5.Ethical Considerations and Data Handling	26
6.6.Pilot Testing	26
7. Results and Findings	27
7.1.Qualitative Findings	27
7.2.Quantitative Findings	28
7.3.Method Implementation Feasibility Findings	28
7.4.Integrated Findings	29

7.5.How this Thesis Extends Kuptavanich’s Work	29
8. Future Research	31
8.1.General Themes	31
8.2.The Proposed Method	31
9. Conclusion	33
10.Bibliography	34
11. Appendices	35
Appendix A. Interview Questions - Reviewing a Paper	36
Appendix B. Interview Questions - Finding Research Focus	37
Appendix C. Interview Questions - Summary Generation Method Understandability for Implementation	38
Appendix D. Experimentation Plan	39
Appendix E. Experimentation Plan Part 1: Semi-Structured Interviews	40
Appendix F. Experimentation Plan Part 2: Structured Web Questionnaire	41
Appendix G. Consent Procedure for Interviews	43
Appendix H. Consent Procedure for Web Questionnaire	44
Appendix I. Pseudo-code for Summarisation Algorithm for Consumer Products	45
Appendix J. Pseudo-code for Summarisation Algorithm for Bibliographies (Kuptavanich Method)	47
Appendix K. Pseudo-code for Summarisation Algorithm for Bibliographies (Mättas Method)	48
Appendix L. Generated Summary (Kuptavanich Method)	49
Appendix M. Generated Summary (Mättas Method)	50

1. Introduction

This thesis delves into the potential of enhancing the utility of bibliographies, typically appended to scientific papers and conference proceedings, by providing readers with computer-generated textual summaries.

The proliferation of peer-reviewed research undoubtedly presents advantages, such as a robust body of work that can be referenced and built upon. This facilitates the branching of topics into subtopics, enabling individual researchers or teams to contribute to the main topic while focusing on specific areas. Moreover, the availability of expertly processed data on a subject provides numerous reference points for making informed contributions towards further describing the subject.

However, the vast quantities of data also pose challenges. Among the potential drawbacks are the dilution of truth due to the difficulty of detecting low-quality research among published papers, and the potential for topic branching to create divides among scientists, complicating contributions towards a main topic. Instead of defending their research against opposing camps in the field, researchers could better utilise their time through effective collaboration.

As scientific fields progress, they often become increasingly interdisciplinary. Artificial intelligence, for instance, merges computer science, psychology, and philosophy, among other disciplines. Developments in one area can unlock new avenues in others, fostering a symbiotic relationship between fields.

While there are problematic aspects of contemporary scientific practice, such as publication bias and citation bias, this research project will not delve into these issues. Instead, after highlighting some of the challenges in the field of science, I propose a supportive approach that helps build bridges and develop understanding in science. The focus of this work is the development of a method to summarise bibliographies of scientific papers and conference proceedings, supporting scientists in making informed research decisions.

In the following sections, I will explore relevant work in Natural Language Generation, including a previous attempt to produce computer-generated summaries of scientific bibliographies (Section 2), describe a qualitative investigation conducted to inform this project (Section 3), draw conclusions from this investigation and inform the implementation of the project (Section 4), detail the implementation and technical aspects of the project (Section 5), evaluate the method and explain the results (Sections 6 and 7), and discuss possible future avenues for this research topic (Section 8).

The research question posed is: **What are the specific features and elements that enhance a summary of a bibliography?**

Furthermore, I propose a hypothesis: **The availability of summarised bibliographies will aid scientists in their daily work.**

2. Natural Language Generation

Reiter & Dale (2002) describe Natural Language Generation (NLG) as a specialised field within artificial intelligence and computational linguistics. It focuses on developing computer systems capable of generating coherent texts in English or other human languages, using non-linguistic information representations such as databases or structured data sources like spreadsheets as a basis. Porter et al. (2009) emphasise that this definition primarily applies to the transformation of data into text, a process known as data-to-text. They suggest that such applications could be utilised to distill complex information, such as medical data, into a format easily digestible by humans.

Furthermore, NLG systems can facilitate text-to-text generation, which includes summarising existing texts or generating new texts from textual information, a concept explored by Haim & Graefe (2017). Gatt and Krahrmer (2018) have also noted a rise in systems that employ images as a foundation for text generation.

NLG systems find applications in various domains where clear and efficient communication of information is crucial, such as news reporting, weather forecasting, and business intelligence. These systems typically convert structured input data into a textual output that encapsulates the data. The input data can take multiple forms, including tables, graphs, and lists. While most examples can be categorised as data-to-text, the key distinction lies in whether the data is internally generated or pre-exists, as in the case of medical data.

NLG systems often generate data summaries and can be customised to produce text in a specific style or tone. For instance, a weather NLG system might generate succinct, clear weather descriptions, while a business NLG system might produce more detailed and formal reports. These systems can generate text in any language and are often integrated with machine translation systems to produce multilingual text.

The challenges of NLG are primarily linguistic. The process of converting non-linguistic data into a textual representation involves tasks such as determining the content to include, forming cohesive text from multiple sentences, and selecting appropriate words, phrases, or expressions. As van Deemter (2016) highlights, even the generation of proper names requires careful consideration.

Gatt & Krahrmer (2018) identified three primary approaches to creating NLG systems: modular, planning, and integrated (or global). Modular approaches divide the various tasks into different systems, which can be designed and implemented in diverse ways. Planning approaches use more integrated systems that follow a series of planned actions. Integrated approaches aim to incorporate all tasks into a single system. These approaches can further be divided into rule-based and stochastic or statistical approaches. In recent years, statistical approaches have gained popularity, particularly with the advancement of machine learning, neural networks, and deep learning techniques, as discussed by Le Glaz et al. (2021).

2.1. Evaluation in Natural Language Generation

Evaluation in Natural Language Generation remains a critical frontier in the field, poised between advancing technical sophistication and achieving a deeper resonance with end-user needs. As underscored by van der Lee et al. (2021), the contemporary pivot in NLG evaluation leans towards embracing user-centric metrics, challenging the traditional confines of linguistic attributes such as fluency and coherence. This transition not only denotes a shift towards practical applicability and satisfaction from the user's perspective but also hints at the need for a holistic approach in evaluating system outputs.

Building upon this foundation, the methodologies proposed by Kuptavanich et al. (2024) introduce a forward-thinking perspective by extending their summarisation algorithm from consumer products (Kuptavanich, 2018) to academic bibliographies (Kuptavanich et al., 2024). This expansion not only showcases the algorithm's versatility but also responds adeptly to the changing landscape of NLG evaluation criteria. This thesis not only applies Kuptavanich et al.'s (2024) principle of "Overview first, zoom and filter, then details-on-demand" (p.1) to the summarisation of scholarly texts but also endeavours to enhance this method. It seeks to refine and augment the algorithm to better suit the intricacies of academic literature, aligning with the evolving direction towards evaluation frameworks that incorporate both qualitative and quantitative analyses. This dual approach enriches the evaluation process, integrating structured, rule-based techniques that have demonstrated effectiveness in commercial applications.

Moreover, this thesis synergises the summarisation principles of Kuptavanich et al. (2024) with the evaluation frameworks suggested by van der Lee et al. (2021), creating a methodological blend tailored to meet the complex demands of summarising both commercial and academic texts. This synthesis underscores a commitment to developing dynamic, user-focused evaluation paradigms. By highlighting the summarisation techniques' ability to identify and prioritise key information, this research echoes critical evaluation criteria in NLG, particularly the need to define and convey content that maximises value for end-users in terms of clarity, relevance, and usability.

The incorporation of Kuptavanich et al.'s (2024) algorithm into this study represents an effort to integrate effective summarisation techniques with the sophisticated evaluation criteria demanded by the modern NLG field.

This integration aims to advance NLG evaluation, ensuring that developed systems not only demonstrate technical excellence but also resonate with the diverse needs and preferences of users. The merging of these methodologies illustrates the potential of adaptable summarisation strategies to fulfil the requirements of real-world applications, thereby broadening the scope and enhancing the impact of NLG technologies across multiple sectors.

Ultimately, this thesis marks a deliberate shift in NLG research, embracing Kuptavanich et al.'s (2024) summarisation algorithm as a cornerstone to pave the way towards a comprehensive evaluation paradigm. This paradigm accommodates the advanced evaluation criteria outlined by van der Lee et al. and aligns with the overarching goals of contemporary NLG research. As the field of NLG progresses, the fusion of such methodologies will be ensuring that generated texts not only meet but also surpass the complex expectations of various application areas, expanding the capabilities of NLG.

3. Gathering Requirements for Summary Generation from Interviews with Scientists

Natural Language Generation (NLG) systems are typically constructed utilising machine learning algorithms (Gatt & Krahmer, 2018), or through the emulation of target texts collaboratively crafted by NLG designers and potential users (Reiter & Dale, 2000). However, the method of conducting focused interviews with potential readers is less frequently employed in NLG research. This work aims to explore its application in discerning user preferences for generated texts and to ascertain the value of such interviews in the NLG system design process. By meticulously selecting interview participants, I can glean more precise and insightful information, thereby enhancing the efficacy of the generated texts. To gain a more profound understanding of scientific workflows, I conducted a qualitative interview with a select group of seasoned researchers from Utrecht University. This approach allowed me to gather valuable insights on the topic and foster a deeper comprehension of the needs and preferences of those who engage with scientific information.

3.1. Interviewing Methods

Interviews serve as a powerful tool for delving into the narrative behind a participant's experiences, allowing the interviewer to extract in-depth information surrounding the topic. Given that these interviews were intended to serve as a foundation for formulating a hypothesis, a degree of flexibility was incorporated into the interview setup.

The structure of research interviews is a distinct field of study in itself (Qu, 2011). Broadly, there are three types of interviews - structured, semi-structured, and narrative. Each type brings its unique advantages to qualitative research. After careful consideration, I opted for the open-ended narrative interview (Turner, 2010). This format aligns with the objectives of this research project as it allows me to guide the conversation and delve into specifics as needed. The open-ended nature of the interview also offers the interviewee ample freedom to articulate their personal perspective without the burden of meeting any preconceived expectations.

While there was no rigid format, I developed a standard set of questions to pose to each participant. The investigation focused on two activities influenced by bibliographies - reviewing a research paper and exploring a research field based on conference proceedings. Each activity was accompanied by its own set of scenario-specific questions. These questions can be found in Appendix A and Appendix B.

3.2. Participants

The sample size for this study was carefully selected to provide sufficient information relevant to the scope of the research, with a slight surplus for added depth. Without specifying the exact distribution per scenario, half of the participants responded to questions about reviewing a research paper, while the other half addressed inquiries about exploring a research field based on conference proceedings. Both scenarios were represented across all seniority levels. The interviewees comprised of two full professors, five associate professors, and one assistant professor, all from the Department of Information and Computing Sciences at Utrecht University.

3.3. Findings

This section presents the insights gained through the interview process conducted as part of this thesis work. The interviews served as a crucial avenue for gathering firsthand information and perspectives from the participants, allowing for a comprehensive exploration of the research topic.

Firstly, interviewees typically investigate a paper's bibliography when prompted by certain elements within the scientific literature. These include:

- A related work section that seems lacking or missing key research.
- An interesting related work or methods section.
- References in the text that spark their interest, such as a surprising fact.
- References in the text from familiar authors.
- References in the text that cite the reader.
- Claims by the researcher of completely novel work without citing previous work in the field.

Secondly, when seeking out a reference in the bibliography, interviewees are interested in the following aspects:

- Publication date.
- Number of authors.
- Location of the original author(s).
- Affiliation of the original author(s).
- Publication venue of the referenced scientific work.

- Citation count.
- For previously unknown work to the reader, finding it to add to their collection.

Thirdly, interviewees use the bibliography when examining conference proceedings (or books) to gain a general understanding of the topic. They seek answers to questions such as:

- Is there a survey on the topic available?
- What general topics are investigated in the field? This is gathered and assumed from the title of the proceedings (or books).
- What methods might be used in the field? This is gathered and assumed from the title and author(s) of the proceedings (or books).
- Which scientific works are representative of the scientific field to add to their reading list?

Fourthly, there are aspects related to the bibliography that are currently inaccessible:

- Research area(s) of the referenced paper.
- Keywords and topics of the referenced work.
- Scientific work related to the referenced scientific work (e.g., references in the same bibliography).
- Publication venue of the referenced scientific work.
- Integrity and rating of the publication venue.
- General timeline of the research field.
- Noteworthy developments in the research field over time.
- License under which the scientific work is published (e.g., Open Access).
- Research techniques used in a technical research work.
- Reproducibility in a technical research work.
- Pointer to the exact sentence in the referenced work that is used.
- Novelty of the referenced work.

Fifthly, interviewees rarely focus on a scientific work's bibliography unless required (e.g., during a reviewing process) due to potential negative connotations that might bias their view of the scientific work:

- A large number of references in the bibliography can be overwhelming or confusing.
- A small number of references may indicate that the author has overlooked important work in the field.
- Citing pre-prints can raise questions about the scientific work's integrity.
- It is challenging to determine the relevance of every reference from the bibliography alone.

Sixthly, certain aspects of the bibliography can overshadow the scientific work's content:

- A small number of references in the bibliography can devalue the scientific work.
- A large number of self-references in the bibliography can devalue the scientific work also.

As the interview stage of the thesis work is considered preparatory, it will not be formally evaluated. The insights gathered are used to determine the requirements for designing a method for generating textual summaries from the bibliographies.

Before listing specific findings, it is important to note that bibliographies have minimal impact on investigating a research field based on conference proceedings when deciding on a new research focus. During interviews, it was commonly mentioned that the process of investigating a research field often extends beyond just examining conference proceedings. Research group heads typically decide on a new research focus using a complex workflow, often connected to their current or previous scientific work. If there is a need to focus on an unfamiliar area, they seek relevant expertise by hiring someone knowledgeable in the new field of interest.

All participants agreed that they often feel overwhelmed by the sheer number of scientific works present in the conference proceedings. Due to the complex workflows and high volume of scientific works, conference proceedings are often not considered as a source of inspiration for entirely new research. However, there might be untapped potential in these conference proceedings and their bibliographies if presented in a more digestible format.

3.4. Insights

3.4.1. Key Insights from Interviews

The interview participants consistently expressed that they are well-versed with the ongoing advancements in their respective research fields. Interestingly, several participants observed a trend of their research fields becoming increasingly diverse and expansive as they gain popularity. This phenomenon presents a double-edged sword. On one side, the growing community makes it more challenging to stay abreast of all developments. Conversely, the increased number of contributors propels the field forward at a faster pace.

The bibliography of a scientific work serves as a valuable tool for researchers to gain a deeper understanding of the community. Every interviewed researcher emphasised the importance of being familiar with the work of other authors in their workflows. Often, recognising the author's name also serves as a form of personal validation.

An additional insight that emerged was the potential interest in the geographical location of authors. This information could provide researchers with a broader context of the global distribution and diversity within their field.

3.4.2. Assessing the Credibility of Published Work

Each participant has developed a unique, subjective internal metric for assessing the quality of published work over time. Upon further exploration, this metric appeared to be more of an intuitive 'gut feeling' rather than a quantifiable or directly measurable standard. Certain authors and publication venues (such as conferences, journals, newspapers) are assigned varying degrees of credibility based on the participant's personal experience and familiarity with them. This subjective evaluation process underscores the nuanced and individualised nature of academic assessment.

3.4.3. Interpreting the Content of Published Work

Occasionally, the content of scientific work may be obscured by an enigmatic title or an abstract that is difficult to decipher. This is where experience and community connections come into play. Researchers who are well-integrated into their academic community often have a sense of their peers' ongoing work and research interests. This familiarity allows them to anticipate the content of a colleague's paper, regardless of how cryptic the title may appear. This highlights the importance of networking and maintaining active engagement within one's research community.

3.4.4. Perceived Value of Published Work

All interview participants were familiar with the citation count as a metric for evaluating the impact of a research paper. However, it was disliked without exception. Participants described it as misleading, unjust, and often a distraction from the true essence of advancing science and making progress in their respective fields. They also mentioned derived metrics, such as the h-index, which attempt to provide a more nuanced comparison of authors and scientific work. However, the prevailing sentiment was one of disappointment, as there seems to be no satisfactory method for measuring the popularity or influence of scientific work that accurately reflects its content and the contributions of the author.

3.5. Implications for Generating Textual Summaries

The interviews revealed six major themes, as outlined in section 3.3. To encapsulate and pragmatically interpret these overarching themes:

1. Researchers seldom concentrate specifically on bibliographies in their workflows.
2. Researchers utilise bibliographies subjectively.
3. Researchers perceive that bibliographies contain valuable information that can be unlocked through summarisation and metadata analysis.

There are several aspects of scientific workflows where summarising bibliographies can provide benefits to the reader. However, to maintain focus and scope on the individual researcher and how to directly assist them in their work, I had to make a selection for the thesis project. I chose to further investigate certain features for summarisation. Creating a summary based on the following list of features can provide value to the reader without requiring any additional effort on their part:

1. Timeline of the references in the entire bibliography.
2. Author information of the referenced work from the bibliography.

There are also additional features that could be beneficial to the researcher and can be gathered by investigating metadata:

1. Frequently studied topics.
2. Publishing venue information of the referenced work from the bibliography.
3. Familiarity of the referenced work.

It is important to note that while citation count is a specific metric that can be easily collected and presented to the reader, it comes with several inherent issues. Most notably, it is a metric that can be disconnected from the relevance of the work's content in the research field. For instance, a classically important work may have a high citation count, while more recent work that improves upon the classical work and provides a definitive answer to a research question might not have been cited as much. To better comprehend these implications, let's examine each one in detail.

1.1. Timeline

A recurring theme among interviewees was the significance of the publication date of referenced works. While the ideal representation of this feature within a summary remains unclear, most researchers hinted at its potential benefits. This information was particularly intriguing in the context of summarising references related to conference proceedings.

1.2. Author Information

The interviewees also expressed a common interest in accessing information about the authors of the referenced works. Specifically, they were interested in details that could be quickly gleaned, such as the number of authors, their affiliations, and potentially their general locations. Highlighting self-references was also deemed valuable by the readers.

1.3. Topic Information

The ability to summarise the topics of referenced work could provide a valuable snapshot of current trends in a field. This information could help researchers, both junior and senior, to quickly familiarise themselves with a new field. It could also provide insights into the subject matters that have influenced the work at hand. It is crucial to examine each reference individually and consider any available metadata, such as often-overlooked keywords. Additional information about the subject matters, such as a timeline and activity in the field, could also be valuable.

1.4. Publishing Venue Information

Given the multitude of venues catering to similar audiences, researchers might not be aware of all the journals and conferences active in their field. This sparked a discussion about the potential value of providing information about these venues to researchers. The consensus was that the prestige of a venue should be investigated and included in the generated summary. Additional details could include whether the article is published, a pre-print, or perhaps even a blog post or newspaper article, among others. It was noted that research published solely on a website might be unjustly undervalued, highlighting the importance of considering diverse publication platforms.

1.5. Familiarity of the References

In addition, researchers often have familiarity with certain aspects of the referenced work. This could include recognising familiar authors or conferences based on their own previous research or involvement in the field. Including this information in the generated summary would help researchers quickly identify and connect with work that is familiar to them, enhancing their understanding and engagement with the scientific literature.

3.2. Conclusion about Requirements

In conclusion, the task of bibliography summarisation presents an opportunity to provide valuable insights to researchers with minimal additional effort on their part. While bibliographies are currently viewed as mundane and often overlooked in scientific workflows, this research project aims to help change that perception by extracting meaningful information from bibliographies.

It is evident that bibliographies are subjectively used by researchers in specific cases. They are often relied upon to verify information, overcome personal doubts, or gain a general understanding of the content, particularly in cases where there is a large amount of material, such as conference proceedings. However, there is much more to be explored in the references themselves, beyond the content of the scientific work itself.

Researchers recognise the potential value that can be derived from bibliographies if further investigation is undertaken. By collecting metadata and drawing conclusions from it, the reader can better contextualise the work and gain a deeper understanding of its significance. Summarising bibliographies has the potential to unlock this untapped potential and provide researchers with valuable insights.

Based on the insights gathered from the interviews, the following larger topics will be focused on in this thesis project:

- Timeline of the references in the whole bibliography.
- Author information.
- Topic information.
- Publishing venue information.
- Familiarity of the references.

By addressing these, the generated summaries can provide researchers with key information that will enhance their understanding and engagement with the literature, ultimately facilitating their research process and contributing to scientific progress.

4. Towards an Improved Algorithm for Generating Textual Summaries from Bibliographies

In previous work by Kuptavanich (2018), efforts were made to generate human-like and descriptive textual summaries, particularly focusing on consumer products. The concept of highlighting features with the highest impact on price was explored, and a summarisation algorithm for consumer products was proposed. See Appendix I for pseudo-code of the algorithm.

Kuptavanich's (2018) findings indicated that human-like and descriptive summaries were more useful to readers compared to shorter versions with the same factual information. However, the investigation into text comprehension between the two summary versions yielded inconclusive results.

When it comes to generating summaries from bibliographies, the relevance of highlighting features related to primary metric such as price diminishes, as each researcher's workflow and preferences differ. The quality and effectiveness of the generated summary become more subjective and personalised for the reader. Rather than focusing on identifying important features among all the available ones, the emphasis will be on achieving an optimal balance and granularity of different features within the summary. The goal is to generate summaries that stand out and cater to the unique needs of each reader. In this regard, Kuptavanich et al.'s (2024) experience with summarising bibliographies of scientific work is taken into consideration. See Appendix J for pseudo-code of the algorithm.

Similar to the previous findings, Kuptavanich et al. (2024) discovered that readers preferred human-like and descriptive summaries over purely factual details or no summary at all. The investigation into reader preferences regarding the level of human-like expressivity in generated summaries also yielded inconclusive results.

Kuptavanich et al.'s (2024) evaluation suggests that his approach to bibliography summarisation could benefit academic readers, but it lacked a deep understanding of their specific needs. To address this limitation, the decision was made to conduct interviews with researchers as a starting point for this thesis. The goal is to gain insights into what makes bibliography summaries more useful to them and apply these lessons to generate improved summaries.

In contrast to Kuptavanich et al.'s (2024) approach, the use of citation count as an implicit metric in the summary will be completely excluded in the proposed algorithm. The interviews highlighted clear issues and limitations with relying solely on citation count as a measure of importance or relevance.

Furthermore, the proposed algorithm will expand on the time dimension of references by summarising the timeline of a particular research area based on keywords. These same keywords will be used to provide a general summary of the entire bibliography in terms of the research areas covered. Author information will be investigated in depth, including details about their location and affiliation. Venue information will also be explored to provide additional context. To personalise the summary and make it more relevant to the individual researcher, the topic information derived from keywords, will be examined. Additionally, familiarity of the references, such as any prior collaborative work between the referenced author and the reader, will be considered.

By incorporating these aspects into the proposed algorithm, the generated summaries aim to be more informative, personalised, and useful to researchers, ultimately enhancing their experience with bibliographies and facilitating their research process.

4.1. Algorithm Evaluation

In the pursuit of evaluating NLG algorithms, especially those aimed at generating summaries that cater to the sophisticated requirements of discerning readers, a multi-faceted approach to evaluation is paramount. Recognising the limitations of metric-based evaluations like BLEU, which necessitate large training and test datasets not readily available for this project, the focus shifts towards a more nuanced, human-centred evaluation methodology.

To rigorously assess the efficacy of the proposed algorithm for summarising scientific bibliographies, I have adopted a two-pronged approach.

4.1.1. Expert Evaluation

Expert researchers have been interviewed to conduct a qualitative assessment of two distinct types of summaries. The first summary type adheres closely to Kuptavanich et al.'s (2024) method, with minor modifications to enhance its comparability with the second summary type proposed in this thesis. The second type incorporates a more elaborate structure, offering additional information and descriptive elements.

The interviewees were tasked with evaluating these summaries through a semi-structured interview process. This format is chosen for its flexibility, allowing experts to express their judgments on the clarity and utility of each summary and to state their preferences based on their personal or professional needs. The semi-structured nature

of the interviews also opens the door for open-ended feedback, providing rich, qualitative insights that extend beyond predefined assessment criteria.

4.1.2. Focus Group Evaluation

Complementing the interviews, a focus group evaluation involves a broader cohort of academic readers engaging with both summary types. Unlike the expert researchers, the focus group participants will employ a standardised questionnaire for their assessment, incorporating quantitative measures such as Likert scale ratings to evaluate various summary features.

This dual-layered evaluation framework—combining the depth of expert insights with the breadth of focus group data—aims to furnish a holistic understanding of the algorithms' performance. By integrating structured and open-ended feedback mechanisms, this approach not only captures quantitative metrics of effectiveness but also qualitative nuances that might inform future enhancements.

The synthesis of findings from both the expert and focus group evaluations will be instrumental in refining the summarisation algorithms. This iterative process, informed by diverse perspectives and detailed feedback, ensures that the developed algorithms are not only technically sound but also resonate with the practical and nuanced needs of the research community.

5. Design and Implementation of the Model

In this research, I present a method that focuses on generating textual summaries from bibliographies found in scientific literature. My method takes into account the content and metadata of the bibliography to create a comprehensive and useful overview for readers. The generated summary is organised into distinct sections, each providing valuable insights into the bibliography.

The sections included in my summary are as follows:

- **Introduction:** This section sets the context and provides a brief overview of the bibliography. It highlights the main topic or research area covered by the sources.
- **Timeline:** The timeline section provides information about the publication dates of the referenced works. It gives readers a sense of the chronological development of the research field and allows them to understand the historical context.
- **Authors:** In the authors section, I provide details about the authors of the referenced works. This includes information such as the number of authors, their affiliations, and potentially their geographical locations. By highlighting the authors, readers can gain insights into the expertise and credibility of the sources.
- **Topics:** The topics section focuses on summarising the main subjects or themes explored in the referenced works. It helps readers quickly grasp the areas of research covered by the bibliography and identify relevant topics of interest.
- **Character:** This section highlights the diverse range of publishing venues, from peer-reviewed journals to blogs and pre-prints, acknowledging the importance of understanding where research is disseminated.
- **Familiarity:** The familiarity section considers the familiarity between the reader and the authors referenced in the bibliography. It may highlight any prior collaborations or interactions, providing a personal connection and context for the reader.

By structuring the summary into these distinct sections, my method aims to help readers quickly grasp the context and significance of the sources within a bibliography. It eliminates the need for readers to extensively read each source and provides a concise overview to inform their further exploration. This approach empowers readers to make informed decisions about which sources to prioritise and delve deeper into based on their individual interests and research goals.

5.1. Simulated Perceptual Grouping: Enhancing Human-Computer

Interaction

Simulated Perceptual Grouping (SPG), a method proposed by Thórisson (2019), has gained attention in the realm of human-computer interaction (HCI) for its potential to create more intuitive interfaces. SPG is designed to mimic the human perceptual grouping process, a key element in how humans process and understand visual data.

The fundamental principles of SPG are derived from the Gestalt laws of perceptual organisation, such as proximity, similarity, continuity, and closure. These principles guide human interpretation of visual stimuli.

The core of the SPG system is the Grouping Engine, designed to mimic the perceptual grouping process in real time. This engine uses algorithms to calculate the strength of various perceptual groupings based on the input data. By assessing the relationships among visual elements, the Grouping Engine pinpoints the most likely perceptual groupings, equipping computers with the ability to comprehend visual information much like humans. Experiments with the Grouping Engine have shown its capacity to accurately predict and form meaningful perceptual groupings, leading to more intuitive and user-friendly interfaces.

Simulated Perceptual Grouping marks a significant milestone in advancing more natural and efficient human-computer interaction. By emulating human perceptual processes, SPG boosts the computer's understanding and responsiveness to visual data, ultimately enhancing the overall HCI experience. In my work, I integrate some of the SPG algorithms and concepts to further enhance my system's capabilities.

5.2. Pre-processing Stage

In the pre-processing stage, my primary focus is on handling data directly from each source in the bibliography, without any interaction with APIs or metadata. This stage involves a thorough cleaning and preparation of data. By implementing methods such as cleaning, instance selection, normalisation, one-hot encoding, transformation, and feature extraction, I can ensure the data's quality and suitability for further analysis. This rigorous data preprocessing is crucial for ensuring the quality of the results in subsequent stages of my project. Please note that at this stage, I do not delve into the structure of the bibliography.

5.3. Information Extraction

My method for generating textual summaries from bibliographies involves a critical step: information extraction. This phase, following the initial pre-processing stage, revolves around obtaining supplementary data about each source from APIs. Information extraction is the process of automatically extracting structured information from unstructured and semi-structured machine-readable documents. In my unique application, I utilise APIs to access databases and gather information that isn't directly available in the bibliography. This could encompass additional details such as the impact factor of the journal where the source was published or extra information about the authors.

The process of information extraction unfolds through several sub-steps:

- **API Selection:** Based on the type of information I aim to extract, I choose the most appropriate APIs. For instance, I might opt for the ORCID API to gather more extensive details about the authors.
- **Query Construction:** Using the metadata procured in the pre-processing stage, I construct queries for the selected APIs. These queries are designed to fetch the most relevant and beneficial information for my summary.
- **Data Retrieval:** The constructed queries are sent to the APIs, leading to the retrieval of the required data. This retrieved data is subsequently processed and prepared for integration into my summary.
- **Data Preparation:** The retrieved data is often inconsistent, from API to API and from record to record. Following Simulated Perceptual Grouping (SPG) principles as described before, the retrieved data is processed and grouped to provide the most useful summary based on the available data.
- **Data Integration:** The final step involves the integration of the retrieved data with the metadata extracted in the pre-processing stage. This combined dataset forms the backbone of my summary.

The extraction of additional information from public databases allows me to present a more comprehensive and informative summary. This not only enhances the value of the summary for the reader, but also provides a more holistic picture of the bibliography. In turn, this empowers readers to swiftly understand the context and significance of the sources, guiding their deeper exploration of the scientific literature.

To sum up, my method for generating textual summaries from bibliographies is a blend of pre-processing, information extraction, and the application of HCI principles. This structured approach is designed to deliver a comprehensive and beneficial overview of the bibliography, thereby facilitating a more informed exploration of the literature.

5.3.1. List of commonly used APIs

In the rapidly evolving landscape of digital research, Application Programming Interfaces (APIs) have emerged as invaluable tools for accessing and manipulating a vast array of scholarly data. The following section delineates a curated list of APIs that have been instrumental in this thesis, each offering unique functionalities that extend the horizons of academic research. From facilitating the exploration of academic literature to enabling the analysis of complex citation networks, these APIs provide the foundational data and tools necessary for advanced scholarly analysis. Below is an overview of each API, along with a brief description of its capabilities and direct links for further exploration.

- **Lens.org API:** Provides access to academic information on research papers, patents, and scientific documents, enabling researchers to explore academic literature and gain insights into various research domains. (<https://docs.api.lens.org>)
- **Microsoft Academic Graph (MAG) API:** Offers access to a large dataset containing information about academic papers, authors, venues, and more, facilitating metadata retrieval, citation network analysis, and exploration of academic entities. (<https://www.microsoft.com/en-us/research/project/academic-knowledge>)
- **Semantic Scholar API:** Developed by the Allen Institute for AI, this search engine for academic publications allows querying their database for metadata, citations, and other scholarly paper information. (<https://www.semanticscholar.org/product/api>)
- **CrossRef API:** Provides access to a searchable database of scholarly works, enabling researchers to retrieve metadata about research papers, including DOIs, publication dates, authors, and references. (<https://www.crossref.org/documentation/retrieve-metadata/rest-api/a-non-technical-introduction-to-our-api>)
- **Europe PMC API:** Part of the worldwide network of PubMed Central repositories, this API offers access to metadata, abstracts, and full-text articles in life sciences, providing a valuable resource for data and information extraction. (<https://europepmc.org/RestfulWebService>)
- **arXiv API:** A preprint server API for research papers in various scientific disciplines, allowing search and access to metadata for articles, including titles, authors, abstracts, and submission dates. (<https://arxiv.org/help/api>)

- OpenAIRE API: Promoting open science by providing open access to research publications and data, this API enables access to metadata, full-text articles, and scholarly work information. (<https://graph.openaire.eu/docs/apis/home>)
- CORE API: Aggregates open access research papers from repositories and journals worldwide, allowing access to metadata, full-text articles, and scholarly work information. (<https://core.ac.uk/services/api>)
- PubChem API: A database of chemical molecules and their activities, this API provides access to chemical information, bioassay data, and other scientific data related to chemical compounds. (<https://pubchem.ncbi.nlm.nih.gov/docs/pug-rest>)
- ORCID API: Provides unique researcher identifiers and a method of linking research activities, allowing access to data about researchers, their affiliations, publications, and more. (<https://info.orcid.org/documentation/api-tutorials/api-tutorial-add-and-update-data-on-an-orcid-record>)
- Unpaywall API: Offers access to a database of millions of free scholarly articles, helping researchers find open access versions of paywalled research papers and access full-text versions when available. (<https://unpaywall.org/products/api>)

These APIs offer a wide range of resources and functionalities for accessing academic information, retrieving metadata, and analysing research papers. Researchers can leverage these APIs to explore scholarly literature, gain insights, and conduct various research activities.

5.4. Summary Generation

The summary generation process involves organising the extracted information and derived statistics into distinct sections, each of which is described in detail below.

5.4.1. Section: Introduction

There are [#] sources in the bibliography.

- Report the total number of sources included in the bibliography, giving readers an idea of the scope of the literature.
- Pseudo-code:

```
1. Read bibliography data
2. Count the total number of sources
```

By describing the Introduction section, readers are provided with a more comprehensive understanding of the scope of the bibliography.

5.4.2. Section: Timeline

The weighted average publishing year among all the sources is [#]. [#] of the sources were published before that time, with the earliest source dating back to [#]. The latest source was published in [#].

- Calculate the frequency-weighted average publishing year among all sources to represent the central tendency of the publication years, which can help readers understand the general timeframe of the literature. Multiply publication year for each source by the number of times it appears in the bibliography, summing these products, and then dividing by the total number of sources. The result is a frequency-weighted average that gives more importance to years with a higher number of publications.
- Identify the range of publication years by determining the earliest and latest sources, providing insights into the overall timespan of the research and highlighting any possible gaps or bursts of research activity.
- Report the number of sources published before and after the frequency-weighted average publishing year, offering a balanced view of the distribution of sources over time.
- Pseudo-code:

```
1. Read bibliography data
2. Extract publication years from sources
3. Calculate the frequency-weighted average publication year
4. Count sources published before the frequency-weighted average year
5. Identify the earliest and latest publication years
```

By describing the Timeline section, readers are provided with a more comprehensive understanding of the temporal distribution of sources in the bibliography. This includes a detailed analysis of publication years, trends, and the evolution of research topics, offering insights into the development of the field over time.

5.4.3. Section: Authors

There are [#] authors listed for the sources in the bibliography. [#]% of authors have published [#]% of the work. The three most referenced authors are [A] (with [#] publications), [B] (with [#] publications), and [C] (with [#] publications).

Most authors cited in this paper come from [D] ([#] authors), closely followed by [E] ([#] authors) and [F] ([#] authors).

Based on the authors' affiliations with universities, we can analyse the global distribution of sources. [#]% of the authors are/were active in [G] (primarily in the [H] and [I]), [#]% of the authors in [J] (mostly in the [K]), [#]% of the authors in [L] (notably in [M] and [N]), with [#]

% of the authors representing the rest of the world. Among all the sources, [#] ([#]%) were self-references.

- Count the total number of authors listed for the sources in the bibliography, giving readers an idea of the number of researchers contributing to the field.
- Identify the 3 most referenced authors and report their respective number of sources, highlighting key contributors and thought leaders in the field.
- Analyse author affiliations to determine the 3 most common universities or organisations and their corresponding number of authors, providing insights into the main institutions driving research in the field. The information about affiliations can be gathered using APIs as previously described.
- Estimate the general region of authors' activity based on their affiliations and calculate the percentage of authors representing each region (e.g., Europe, Asia, North America, etc.), offering a geographical perspective on the research landscape.
- Determine the number of self-references within the bibliography, giving readers an understanding of the extent to which authors build upon their own previous work.
- Pseudo-code:

```
1. Read bibliography data
2. Extract author information and affiliations from sources
3. Count authors and their contributions
4. Identify top 3 most referenced authors
5. Identify top 3 organisations based on author affiliations
6. Determine primary research regions based on affiliations
7. Count self-references
```

By describing the Authors section, readers are provided with a more comprehensive understanding of the researchers contributing to the field, their affiliations, and the geographical distribution of research activity. In addition, this extension explores collaboration patterns, self-references, and diversity among authors, offering a more detailed perspective on the research community and its dynamics.

5.4.4. Section: Topics

Most frequently described fields of study are [O] ([#]% of sources), [P]([#]% of sources), and [Q] ([#]% of sources). As the three most frequently described topics, researchers have used "[R]" ([#]% of sources), "[S]" ([#]% of sources), "[T]" ([#]% of sources) as keywords to illustrate the content of their work.

- Identify the 3 most commonly listed fields of study, giving readers a clear understanding of the dominant research areas within the bibliography. The information about fields of study can be gathered using APIs as previously described.
- Determine the 3 most frequently used topics within the sources, highlighting their prevalence in the literature, as well as any potential overlaps or connections between topics. The information about topics can be gathered using APIs as previously described.
- Pseudo-code:

```
1. Read bibliography data
2. Extract topic information for sources via programmatic means (APIs)
3. Identify top 3 most prevalent fields of study and topics
```

By describing the Topics section, readers are provided with a more comprehensive understanding of the main research areas, keywords, and trends in the bibliography, offering a more nuanced perspective on the research landscape.

5.4.5. Section: Character

The three most frequently used types of sources are [U]s ([#]% of sources), [V]s ([#]% of sources), and [W]s ([#]% of sources).

- Analyse the character of sources (e.g., journal articles, conference proceedings articles, and book chapters, etc.) and identify three most represented types of publishing venues, providing readers with an understanding of the distribution of different source types within the bibliography. The information about character of sources can be gathered using APIs as previously described.

- Report character of sources section only if at least 50% of the sources have this information available about them.

- Pseudo-code:

```
1. Read bibliography data
2. Extract source types (e.g., journal, conference proceedings)
3. Identify top 3 most represented types of research
```

By describing the Character section, readers are provided with a more comprehensive understanding of the types and nature of sources in the bibliography.

5.4.6. Section: Familiarity

As you have uploaded an example of your own published research, we have found that you have previously collaborated with the following authors from the bibliography: [X], [Y], and [Z].

- If the user has provided an example of their own published research, identify any previous collaborations or connections (e.g., affiliation) with authors listed in the bibliography, and report the names of the connected authors.

- Assess the user's familiarity with the sources by determining the number and percentage of sources they have cited in their own work, which can help readers understand the user's level of engagement with the literature. The information about familiarity of sources can be gathered using APIs as previously described.

- Report familiarity of sources only if co-authors are present in the bibliography.

- Pseudo-code:

```
1. Read bibliography data
2. Extract author information from sources
3. Compare authors with the reader's previous collaborators
4. Identify any authors that the reader has collaborated with
```

By describing the Familiarity section, readers are provided with a more comprehensive understanding of the influence and impact of the sources in the bibliography using alternative metrics (without citation count). This extension also explores the user's level of engagement with the literature, offering a more detailed perspective on the research landscape.

6. Method Evaluation

In this evaluation segment, I analysed the bibliographic summaries generated through my methodology, setting them against the backdrop of the framework established by Kuptavanich et al. (2024). The primary goal was to shed light on the advancements or distinctive variances my approach brought to the forefront in the articulation of bibliographic data. Special emphasis was placed on pivotal aspects such as clarity and utility, specifically tailored to resonate with an erudite audience comprising scientists and academicians with higher degrees. Moreover, the exploration extended to the pragmatic feasibility of integrating my proposed methodology within real-world applications.

Parallel to this analytical endeavour, I delved into an investigation aimed at unraveling how individuals interact with these summaries devoid of any preconceived context or designated tasks. This strategy was deliberately chosen to capture a spectrum of personal preferences and methodologies, thereby illuminating the potential of my bibliographic summaries to augment their research processes. This inquiry not only enriches the evaluation but also aligns with the evolving paradigms of Natural Language Generation (NLG) evaluation, spotlighting the need for more nuanced and comprehensive assessment techniques that cater to both end-user experiences and the technical viability of implementing such NLG systems.

6.1. Evaluating Method through Interviews with Scientists

In this section, I delve into the qualitative evaluation of my bibliographic summaries, aimed at understanding their reception within the academic community regarding effectiveness, clarity, and utility. A more detailed account of this evaluative approach is delineated in the final section of my thesis, refer to Appendix E. The summaries used for the evaluation are described in Appendix L and Appendix M.

6.1.1. Approach

To gather these insights, I conducted semi-structured interviews with a select group of researchers hailing from a wide array of scientific disciplines. The intent was to tap into their personal experiences and perceptions related to the bibliographic summaries. The semi-structured nature of these interviews facilitated a dynamic exchange, allowing the conversation to organically pivot to explore intricate details of the bibliographic summaries, as and when required by the flow of dialogue.

6.1.2. Participants

The participant pool for this study was composed of eight distinguished academics, including full professors and leading researchers from renowned institutions such as Utrecht University, Tallinn University of Technology, and Tartu University. This carefully curated group provided a broad spectrum of perspectives, ensuring a rich and varied collection of insights into how bibliographic summaries are perceived and utilised in the realm of academic research.

6.1.3. Analysis

To distill the essence of the interviews, I employed thematic analysis, a method that enabled me to identify and interpret patterns within the data. This analytical process was instrumental in uncovering the underlying themes and preferences related to the bibliographic summaries, thereby offering a layered and nuanced understanding of their value as perceived by the academic community. Through this analysis, I aimed to capture not just the surface-level reactions but also the deeper, more subtle nuances that contribute to the overall perception and utility of the bibliographic summaries in scholarly research.

6.2. Evaluating Method through Web Questionnaire

In this segment, I delve into the quantitative evaluation of my bibliographic summaries in comparison to those developed by Kuptavanich et al. My objective was to methodically measure the academic community's reception of these summaries, focusing on key aspects such as clarity, information presentation, and overall usefulness. A detailed exposition of this evaluative approach can be found in Appendix F of my thesis. The summaries used for the evaluation are described in Appendix L and Appendix M.

6.2.1. Approach

To achieve this, I rolled out a structured web-based questionnaire to an extensive group of academics and graduate students. This questionnaire featured anonymised summaries generated by both Kuptavanich's method and mine, allowing for a direct comparison based on a set of predefined criteria. This strategy was designed to elicit

clear, comparative feedback from participants, shedding light on the relative strengths and weaknesses of each approach.

6.2.2.Participants

The questionnaire was structured to collect data on various parameters, including reading frequency, the number of published papers, and preferences for summary types. Initially, 138 participants engaged with the questionnaire, from which 114 responses were deemed suitable for analysis. These 114 respondents were all graduate degree holders from the United States, selected for their active engagement with scientific literature. The demographic selection, facilitated by SurveyMonkey, was balanced in terms of age and gender, ensuring a representative sample of the academic readers' community.

- **Reading Frequency:** The valid participants were almost evenly split in their engagement with scientific literature, with approximately one-third reading weekly or frequently, another third reading monthly or occasionally, and the final third reading yearly or rarely.
- **Published Papers:** Half of the valid participants have not published any papers, indicating a significant portion of early-career researchers or those from non-research-intensive disciplines. A smaller segment has published between 1 to more than 6 papers, showcasing a range of experience in publishing within the academic community. A third of all the valid participants have published more than 6 papers.

6.2.3.Analysis

The analysis of the collected data was conducted using non-parametric statistical test, specifically the Wilcoxon Signed-Rank Test. This test was chosen for its suitability for the non-parametric nature of my data, stemming from Likert scale responses, and the paired design of my study, which compared two sets of summaries evaluated by the same participants. The Wilcoxon Signed-Rank Test is particularly adept at identifying significant shifts in the median differences between the pairs, offering a robust analysis of central tendency differences without the assumption of data normality.

The application of the Wilcoxon Signed-Rank Test was aimed at determining whether statistically significant differences exist in the clarity, usefulness, and overall preference ratings between the summaries generated by my method and those by Kuptavanich et al. (2024). A significant finding in any of these aspects would underscore a meaningful discrepancy in the perceived quality and favourability of the summaries, thus validating or challenging my hypotheses about the comparative efficacy and user preference between the two summary methods.

This rigorous statistical approach not only bolsters the robustness of my findings but also complements the qualitative insights obtained from the interviews, providing a holistic understanding of the academic community's preferences and perceptions regarding the bibliographic summaries.

6.3.Evaluating Method through Interviews with Developers

The objective of this segment was to evaluate the technical feasibility and identify any potential challenges associated with the implementation of my proposed method for generating bibliographic summaries, all from the perspective of software developers. A complete interview setup is described in Appendix C.

6.3.1.Approach

To gain a comprehensive understanding, I conducted in-depth, semi-structured interviews with a diverse group of programmers and software developers. These interviews were designed to probe the developers' comprehension of the proposed method, their confidence in its implementability, anticipated technical hurdles, and any suggestions they might have for enhancing the method. The participants included a mix of professionals ranging from a junior software developer and a senior software engineer to a data scientist, a self-taught full-stack developer, and a cybersecurity analyst. This diversity ensured a well-rounded evaluation of the method's technical viability.

6.3.2.Participants

The selected group of five developers was intentionally diverse, encompassing a wide range of expertise and backgrounds (junior software developer, senior software engineer, data scientist, full-stack developer, and cybersecurity analyst). This selection criteria were aimed at ensuring a comprehensive technical evaluation of the proposed method. Each developer brought a unique perspective to the table, facilitating a rich discussion on various aspects of the method's technical implementation.

6.3.3. Analysis

The interviews were analysed using thematic analysis, allowing for the extraction of significant insights and feedback related to the method's clarity, technical complexity, and feasibility of implementation. This analysis revealed a general consensus among the developers regarding their understanding of the method and their confidence in its potential implementation. However, it also brought to light certain challenges, particularly concerning specific technical details and data handling aspects. This evaluation was crucial in assessing the practicality of bringing the proposed method to life from a software development standpoint.

6.4. Integrating Evaluation Insights

The synthesis of findings from the evaluations offers a clear perspective on the proposed method's practicality and viability. This integration is crucial for refining the bibliographic summary generation technique to ensure it aligns with the academic community's requirements and is technically feasible.

The evaluation framework aimed to assess the method comprehensively, gathering essential insights for its refinement. These insights are pivotal for enhancing the method and informing discussions within the Natural Language Generation (NLG) field about effective bibliographic summary generation.

By merging feedback from both the academic and developer communities, this approach underscores the importance of a multifaceted evaluation in NLG. The gathered insights are instrumental in improving the method, ensuring its relevance for academic research and its feasibility for technical implementation, thereby contributing to the advancement of NLG methodologies.

6.5. Ethical Considerations and Data Handling

In conducting this research, stringent ethical considerations and data handling protocols were adhered to, ensuring the integrity and confidentiality of the information gathered. Ethical approval was obtained prior to the commencement of the study, and informed consent was secured from all participants, guaranteeing their awareness and agreement to the terms of their involvement. A commitment to secure data handling, analysis, and storage was maintained throughout the research process, in strict compliance with prevailing privacy and research ethics guidelines at Utrecht University. The complete details of the ethical considerations and data handling procedures are documented in Appendix G and Appendix H of the thesis.

6.6. Pilot Testing

To enhance the reliability and effectiveness of the method validation, pilot testing was conducted on the interview questions and web questionnaire. A select group of participants was involved in this preliminary phase, providing initial feedback on the clarity and effectiveness of the questions. This feedback was instrumental in refining the questions, ensuring they were well-crafted to elicit meaningful and relevant responses from the larger participant pool in the subsequent phases of the main evaluation.

7. Results and Findings

My evaluation framework, underpinned by narrative interviews and a quantitative analysis, was designed to assess the efficacy and utility of my bibliographic summary technique. The insights harvested from this comprehensive approach not only validated the effectiveness of my method but also illuminated potential avenues for future enhancements. This endeavour ensured that the summaries I generated were in tune with the evolving demands of the academic sphere.

7.1. Qualitative Findings

The qualitative analysis of semi-structured interviews with eight professors across diverse academic disciplines provided critical insights into the design and effectiveness of bibliographic summaries. These insights are organised into thematic categories, each highlighting key preferences and needs that can inform the development of more effective bibliographic tools.

7.1.1. Balancing Conciseness and Detail

Researchers expressed a nuanced preference for balancing conciseness with detail in bibliographic summaries. While brevity aids quick comprehension, detailed information is essential for in-depth understanding. This balance is crucial for bibliographic summaries to be adaptable to different research phases, from initial scanning to deep analysis.

7.1.2. Structural Clarity and Organization

The importance of clear structure and logical organisation in bibliographic summaries was emphasised. Well-defined section headings and an intuitive arrangement facilitate the quick location of relevant information, making summaries more user-friendly and effective in supporting research workflows.

7.1.3. Emphasis on Valuable Features

Quantitative details such as author contributions, geographic distribution, and temporal trends were highlighted as particularly valuable. These features offer deeper insights into the bibliography's scope and impact, suggesting a need for summaries to include specific metrics like self-references that can provide context and relevance to the researcher's specific area of study.

7.1.4. Context-Dependent Utility

The utility of bibliographic summaries was noted to vary significantly with the researcher's context and specific needs. Summaries need to provide relevant, actionable insights tailored to the researcher's current task, whether it be for an overview of the field or for in-depth analysis in paper reviewing or editing.

7.1.5. Need for Content Insights

There is a strong desire for summaries to go beyond bibliometrics and provide insights into the content and quality of referenced works. This suggests that effective summaries should encapsulate a deeper understanding of the bibliography's contributions to the field, aiding researchers in assessing the relevance and quality of the cited works.

7.1.6. Importance of Context in Summary Use

The findings underscore the context-dependent nature of bibliographic summary utility. Researchers' needs for summaries can range widely, from requiring quick overviews to needing detailed analyses for tasks like paper reviewing. This variability necessitates the development of adaptive summaries that can be customised to suit a range of academic tasks.

7.1.7. Length Preferences and Usefulness Hypothesis

Preferences for the length of summaries varied, aligning with the usefulness hypothesis. The need for customisable summaries emerged as a significant theme, indicating that tools should allow users to select the level of detail that meets their current requirements, enhancing the summaries' practicality in diverse research scenarios.

7.1.8. Shifting Relevance of Traditional Metrics

A notable shift away from traditional reliance on citation counts for evaluating publication validity was observed. Researchers are seeking more nuanced and context-specific metrics, such as conference rankings, suggesting a move towards more qualitative indicators of research quality and relevance.

7.1.9. Demand for Interactivity and Flexibility

The desire for direct links to publications and adjustable detail levels in summaries highlights a need for more interactive bibliographic tools. Such features would not only enhance accessibility to source materials but also allow researchers to tailor summaries to their specific needs, thereby increasing their utility in personal workflows.

7.1.10. Refining the "Familiarity" Section

Feedback on the "familiarity" section suggests a need for clearer naming and definition to better reflect its content and purpose. This adjustment could improve the section's value by helping researchers connect the bibliography with their existing knowledge and networks, thus enriching their understanding of the research landscape.

7.2. Quantitative Findings

The summary generated based on the method developed in this thesis will be called a "Mättas" summary and the summary generated based on the method developed by Kuptavanich et al. (2024) will be called a "Kuptavanich" summary for simplification and comparative purposes going forward.

The questionnaire distributed among the wider academic community offered valuable insights into the perceptions of clarity, usefulness, and overall preference for the Mättas and Kuptavanich bibliographic summaries, revealing nuanced preferences and the importance of content presentation in meeting research needs.

7.2.1. Clarity Ratings

The Wilcoxon Signed-Rank Test revealed no statistically significant difference in the perceived clarity between the Mättas and Kuptavanich summaries (p-value: 0.3234). This suggests that both summaries are comparably clear to the participants, indicating that the format or presentation style of each does not significantly impact clarity perception.

7.2.2. Usefulness Ratings

A significant difference was observed in the usefulness ratings, with the Mättas summary being perceived as more useful than the Kuptavanich summary (p-value: 0.00055). This finding indicates that the content or the way information is presented in the Mättas summary aligns better with the participants' expectations or needs in terms of research utility.

7.2.3. Preference

To indicate the more useful version, 71 chose Mättas summary while 43 chose Kuptavanich summary. When asked which summary participants would prefer to use regularly, 66 chose Mättas summary while 48 chose Kuptavanich summary. 17 people who chose Mättas summary to be more helpful would rather use Kuptavanich summary more regularly. 12 who chose Kuptavanich summary to be more helpful would rather use Mättas summary more regularly.

The preference data further corroborated the usefulness ratings, with a majority favouring the Mättas summary over the Kuptavanich summary for both immediate utility and regular use. However, a notable subset of participants displayed a mismatch between the summary they found more helpful and the one they would prefer to use regularly, suggesting nuanced preferences that may depend on specific use cases or personal workflow needs.

7.3. Method Implementation Feasibility Findings

The interviews conducted with software developers provided insights into the technical feasibility of implementing the proposed bibliographic summary generation method.

7.3.1. Developer Understanding and Technical Aspects

Developers demonstrated a solid understanding of the method's objectives and the processes involved, such as extracting key information from bibliographies and organising it into structured sections. Some developers

pointed out difficulties in grasping certain technical elements, particularly the computation of the weighted average publication year, highlighting areas where the method's description could be enhanced for clarity.

7.3.2. Implementation Confidence and Challenges

Developers generally felt confident in their ability to implement the method, albeit some anticipated the need for further training or support to address specific technical challenges. Key challenges identified included ensuring the accuracy and relevance of the generated summaries, managing the extensive datasets involved, and, notably from the cybersecurity analyst, upholding data security and privacy.

7.3.3. Suggestions for Improvement

There was a call for more explicit explanations of the calculations and algorithms employed in the method, to demystify any complex components and make the method more accessible. Additionally, developers suggested incorporating a quality control or evaluation mechanism to guarantee the accuracy and relevance of the summaries, ensuring they meet the intended standards. Enhancing data security protocols was also recommended to ensure the method's suitability for handling sensitive or private bibliographic data, emphasising the importance of privacy and security in the implementation.

7.4. Integrated Findings

The comprehensive evaluation of the proposed method through qualitative interviews and a quantitative questionnaire has unveiled a rich array of insights into the effectiveness of the Mättas and Kuptavanich summary methods. This integrated analysis sheds light on the nuanced preferences within the academic community, emphasising the crucial balance between summary conciseness, detail, and structure to cater to varied research demands.

The findings distinctly highlight a preference for the Mättas method, praised for its clarity, coherence, and overall utility in academic contexts. Participants valued its ability to present comprehensive insights succinctly, striking an optimal balance between brevity and depth of information. This preference aligns with the quantitative analysis results, where Mättas summaries outperformed in usefulness ratings, notwithstanding the lack of significant disparity in clarity compared to the Kuptavanich summaries.

The insights also bring to the fore the subjective nature of summary preferences, shaped by individual research contexts and requirements. While the Mättas summary generally garnered favour, a segment of participants exhibited a penchant for the Kuptavanich summaries, especially in situations necessitating quick overviews or under time constraints.

Qualitative interviews underscored the value of structural clarity and logical organisation in enhancing summaries' readability and utility. The expressed need for adaptive summaries, capable of being customised to specific academic endeavours, indicates a burgeoning demand for flexible bibliographic tools.

The quantitative data intriguingly revealed a significant divergence between the summaries' perceived helpfulness and the preference for their regular use. This suggests that factors beyond informational value, such as ease of use, time efficiency, and personal preference for a particular summary style, significantly influence user preferences.

Participant feedback has highlighted several avenues for enhancing summary methods, including the creation of more adaptive, user-centric summary generation tools and the exploration of alternative evaluation metrics. Moreover, incorporating interactive elements could further augment the engagement and utility of bibliographic summaries. Finally, describing citation count does not seem to affect the evaluation of summary clarity, usefulness or preference at all.

7.5. How this Thesis Extends Kuptavanich's Work

In recognising the foundational work of Kuptavanich et al. (2024) in establishing the initial framework for bibliographic summary generation, this research builds upon their methodology by introducing an approach that incorporates both qualitative and quantitative analyses. The qualitative research conducted aimed to identify potential enhancements to Kuptavanich's method, focusing on the intricacies of bibliographic data and its utility to end-users. Subsequent quantitative research evaluated these improvements, demonstrating their efficacy and utility in real-world applications.

An outcome of this study is the deeper understanding it provides of how different user groups, specifically scientists and graduate degree holders, engage with bibliographic summaries. This insight not only extends Kuptavanich et al's (2024) work but also sheds light on the potential of Natural Language Generation (NLG) to tailor

bibliographic information to diverse academic needs, enhancing the accessibility and applicability of bibliographic summaries in scholarly work.

7.5.1. Proposed Methodology Enhancements

By conducting qualitative research, I delved into the nuanced preferences and needs of bibliographic summary users, particularly scientists and graduate degree holders. This approach allowed for the identification of potential enhancements to Kuptavanich et al's (2024) method, focusing on the detailed utility of bibliographic data.

Following qualitative insights, I employed quantitative research to assess the real-world applicability and effectiveness of the proposed improvements. This dual approach ensures that the enhancements are not only theoretically sound but also practically beneficial.

A significant outcome of this study is the deepened understanding of how different user groups interact with bibliographic summaries. These insights extend Kuptavanich et al's (2024) work by highlighting the potential of Natural Language Generation (NLG) to customise bibliographic information for diverse academic requirements, thereby improving the summaries' relevance and utility in scholarly contexts.

7.5.2. Proposed Summary Content and Structure Enhancements

Unlike Kuptavanich et al's (2024) method, my approach provides extensive details about authors, including unique contributors, notable authors, their contributions, and affiliations, offering a more comprehensive understanding of the bibliography's authorship landscape.

My summaries include prevalent subject matters and subdomains, providing users with a clearer picture of the bibliography's thematic focus and scope.

By incorporating information on venue type and domain, my summaries offer additional context that can be critical for understanding the relevance and prestige of the cited works.

Similarly to Kuptavanich et al's (2024) method, my approach includes self-citation counts, providing insights into the bibliography's self-referential nature.

7.5.3. Addressing Challenges

To tackle incomplete or inconsistent metadata, I developed heuristics and fallback mechanisms to infer missing information, enhancing metadata completeness and reliability.

Given the challenge of managing a large number of sources, clustering and dimensionality reduction techniques were employed to distill the most pertinent information, enabling efficient extraction of key insights.

To ensure the generated summaries are coherent and readable, natural language processing techniques were leveraged. These techniques aided in structuring the summaries and maintaining a logical flow, making them more accessible to researchers.

By overcoming these challenges, my method not only refines the process of generating bibliographic summaries but also contributes to the broader goal of providing researchers with insightful, coherent, and useful textual summaries. These advancements enhance researchers' understanding of the literature and support their scholarly endeavours, marking a step forward in the application of NLG in academic research.

8. Future Research

In the evolving landscape of academic research, the development and optimisation of bibliographic summaries stand as a pivotal frontier for innovation. The drive toward enhancing the efficiency and effectiveness of academic workflows has spotlighted the potential that lies in technical innovations, user experience improvements, and a methodical approach to evaluation and adaptation. These facets, when weaved into a cohesive research strategy, promise not only to revolutionise the utility of bibliographic tools but also to tailor them to the dynamic and diverse needs of the academic community. As we venture into future research endeavours, prioritising these areas will be crucial in fostering a more integrated, user-centred, and intuitive framework for academic research and collaboration.

8.1. General Themes

Future research in the development and optimisation of bibliographic summaries presents opportunities for enhancing the efficiency and effectiveness of academic work. These opportunities span across technical innovations, user experience improvements, and methodical evaluation and adaptation based on feedback and quantitative analysis as also indicated by van der Lee et al. (2021)

8.1.1. Technical Innovations and User Experience Enhancements

Addressing the needs for technical advancements, future research could focus on creating algorithms and interfaces that support adaptive summary generation, allowing users to adjust the granularity of summaries according to their particular tasks and preferences. This includes incorporating interactive elements, such as direct links to full-text articles and visualisation of author networks, which will make summaries not only more engaging but also more informative and easier to navigate. The exploration of alternative evaluation metrics, like conference rankings or qualitative indicators, can offer nuanced insights into the importance and relevance of cited works. Moreover, the design of these tools must pivot around the user, taking into account the diverse roles within academia (researchers, reviewers, editors) and developing features that resonate with their specific requirements. Research should also delve into context-aware summarisation, where content is dynamically modified based on the user's current task, history, and preferences, employing machine learning techniques for personalisation and enhanced utility.

8.1.2. Evaluating and Adapting Based on User Interactions and Preferences

Beyond the development of these tools, it's crucial to continually assess their clarity and usefulness through quantitative measures and qualitative user feedback. Insights indicate that although summaries might be comparable in terms of clarity, their practical value varies, highlighting the importance of identifying which aspects make bibliographic summaries more useful to different user segments. A thematic analysis of user responses can uncover specific features or areas that necessitate improvement, offering a deeper understanding of user needs. Additionally, addressing the discrepancy in preferences for summary use suggests exploring the contexts in which summaries are deemed more helpful versus those where simplicity, ease of reading, or specific needs drive regular use. This exploration could unveil subtle personal preferences or biases towards presentation style, structure, and writing style, guiding refinements in summary designs.

To comprehensively address these aspects, future studies should not only iterate on the development of more sophisticated, user-centred bibliographic tools but also embody a robust evaluation framework that captures diverse user feedback and preferences. Follow-up studies and direct engagement with the user base can provide the necessary clarity and direction for adjusting features or adding new functionalities. Such an approach ensures that bibliographic summaries evolve in alignment with academic demands, fostering a research ecosystem that is more efficient, informed, and collaborative.

8.2. The Proposed Method

The proposed method, while offering a comprehensive framework for analysing bibliographic data, lays the groundwork for future enhancements aimed at refining and expanding its capabilities. Future work could focus on integrating advanced computational techniques, expanding the scope of data analysis, and improving user interaction with the generated summaries. By addressing these areas, the method will evolve to meet the ever-changing needs of the academic community, providing more nuanced insights into the vast landscape of scientific research and fostering a deeper understanding of the intricate web of knowledge that carries academic disciplines forward.

8.2.1.Section: Introduction

In analysing the bibliography, I observed several notable trends and patterns that provide insights into the distribution of sources. These include the presence of a few highly influential papers that have significantly impacted the field. Additionally, I identified a significant number of interdisciplinary studies that bridge multiple fields, indicating the growing collaboration and convergence of knowledge across disciplines.

8.2.2.Section: Timeline

In examining the temporal trends, I identified periods of increased or decreased research activity within the literature. I also explored potential correlations with external factors such as historical events or technological advancements, which may have influenced the direction and focus of research over time. By observing the evolution of research topics, I gained insights into how the field has developed and responded to emerging challenges and opportunities. The distribution of publication years using Gestalt principles from SPG, such as a histogram or line chart, would allow readers to quickly grasp the temporal trends in the literature.

8.2.3.Section: Authors

Analysing the patterns of collaboration among authors revealed recurring co-authorship teams and research clusters, shedding light on the structure of the research community and the relationships between researchers. I also investigated the diversity of authors in terms of gender, ethnicity, and career stage, providing insights into the inclusivity and demographic representation within the field. Using Gestalt principles from SPG, visual representations could be created, such as a histogram or line chart with a map view, to showcase the distribution of authors, enabling readers to easily grasp the trends in the literature.

8.2.4.Section: Topics

Examining the distribution of topics across sources, I identified clusters of highly related papers and observed interdisciplinary studies that bridge multiple fields. This provided insights into the thematic structure of the bibliography and the evolving focus of research areas over time. I also identified gaps or underrepresented areas in the literature, suggesting potential avenues for future research or unexplored aspects of the field. By using Gestalt principles from SPG, visual representations such as word clouds or network diagrams to visually convey the most frequent topics and their relationships could be created, facilitating a quick understanding of the thematic trends in the literature.

8.2.5.Section: Character

Analysing the publishing venues across the sources allowed me to identify the most common venues in the field, highlighting the prevailing venues used by researchers. Examining the distribution of venue types, I observed patterns and trends, such as the predominance of certain types of venues. These findings shed light on the character of the sources in the bibliography and the diversity of approaches employed by researchers. Visual representations, such as pie charts or bar graphs, based on Gestalt principles from SPG, could provide a clear overview of the distribution of source types, facilitating readers' comprehension of the character of the sources. In addition, discreet venues could be pointed out to give the reader more information on the sources used in the bibliography.

8.2.6.Section: Familiarity

In refining the approach to assessing the influence of sources within the bibliography, a significant emphasis was placed on the familiarity of authors to the reader. Recognising the uniqueness of each reader's academic journey, future enhancements to this method could specifically target identifying authors with whom the reader has previously collaborated or whose work they have extensively cited. This adjustment could underline the connections that are most personally relevant to the reader, transforming the bibliography into a map of familiar scholarly terrain.

9. Conclusion

This thesis explored the potential of computer-generated textual summaries to enhance the utility of bibliographies within scientific literature, positing that such summaries could aid not only scientists but all academic readers in efficiently navigating the vast landscape of scientific research. Anchored by the hypothesis that such summaries would not only facilitate scientists but also extend their benefits to the broader academic readership, this thesis sought to streamline the navigation through the expansive terrain of scientific research. The evaluation of this extended hypothesis, corroborated by insights from a web questionnaire, underscores the universal applicability of the summarisation method developed herein.

Central to this thesis is the integration of key components within the summarisation method - metadata extraction, temporal trend analysis, authorship patterns, and thematic exploration - each playing a pivotal role in the method's efficacy. For instance, metadata extraction underpins the accuracy of bibliographic details, while temporal trend analysis illuminates the evolution of research topics, collectively making scientific research more accessible and navigable for all readers.

The findings support the hypothesis, illustrating that succinct, coherent, and informative summaries markedly aid readers in navigating the extensive corpus of literature more effectively. By offering a multifaceted overview of bibliographies, the method expedites the research process. It addresses challenges such as metadata inconsistency and summary coherence through the application of heuristics, clustering, dimensionality reduction, and natural language processing, thereby maintaining the summaries' quality, readability, and relevance.

A contribution of this thesis to the field is the development of evaluation frameworks tailored to the changing landscape of Natural Language Generation (NLG). Recognising the dynamic nature of NLG and its applications, these frameworks are designed to assess the effectiveness of generated summaries in real-world academic contexts. The importance of these frameworks lies in their ability to adapt to the evolving demands of NLG, ensuring that the generated summaries remain relevant, user-friendly, and conducive to the advancement of scientific research. This focus on evaluation underscores a commitment to continuous improvement and responsiveness to the needs of the academic community.

Looking to the future, this research opens several avenues: advancing topic analysis for deeper thematic insights, integrating citation networks to offer richer narratives of academic impact, developing customisable and interactive tools for summary generation, broadening user evaluations to encompass diverse disciplines, and embedding the summarisation method within existing research platforms for wider accessibility.

In conclusion, this thesis fortifies the premise that computer-generated textual summaries can enhance the utility of bibliographies, aligning with the initial hypothesis and addressing the research question. By pursuing the future research directions outlined, coupled with the ongoing development of evaluation frameworks, the field is poised to meet the dynamic needs of the scientific community. This fosters more efficient, informed, and collaborative research practices, heralding a future where the accessibility and comprehension of scientific literature are streamlined, catalysing the advancement of knowledge across disciplines.

10. Bibliography

- Dash, A., Shandilya, A., Biswas, A., Ghosh, K., Ghosh, S., & Chakraborty, A. (2019). Summarizing User-generated Textual Content: Motivation and Methods for Fairness in Algorithmic Summaries. *Proceedings of the ACM on Human-Computer Interaction*, 3, 1–28. <https://doi.org/10.1145/3359274>
- Deemter, K. (2016). Designing Algorithms for Referring with Proper Names. *Proceedings*, 31–35. <https://doi.org/10.18653/v1/W16-6605>
- El-Kassas, W. S., Salama, C. R., Rafea, A. A., & Mohamed, H. K. (2021). Automatic text summarization: A comprehensive survey. *Expert Systems with Applications*, 165, 113679. <https://doi.org/10.1016/j.eswa.2020.113679>
- Gatt, A., & Krahmer, E. (2018). Survey of the state of the art in natural language generation: Core tasks, applications, and evaluation. *Journal of Artificial Intelligence Research*, 61, 65–170. <https://doi.org/10.1613/jair.5477>
- Haim, M., & Graefe, A. (2017). Automated News. *Digital Journalism*, 5 (8), 1044–1059. <https://doi.org/10.1080/21670811.2017.1345643>
- Kuptavanich, K. (2018). Using textual summaries to describe a set of products. *Proceedings*, 563–567. <https://doi.org/10.1145/3240323.3240328>
- Kuptavanich, K., Reiter, E., Van Deemter, K., & Siddharthan, A. (2024). Textual Summarisation of Large Sets: Towards a General Approach. <https://doi.org/10.48550/arXiv.2401.09041>
- Le Glaz, A., Haralambous, Y., Kim-Dufor, D. H., Lenca, P., Billot, R., Ryan, T. C., Marsh, J., DeVlyder, J., Walter, M., Berrouguet, S., & Lemey, C. (2021). Machine Learning and Natural Language Processing in Mental Health: Systematic Review. *Journal of Medical Internet Research*, 23 (5), e15708. <https://doi.org/10.2196/15708>
- Portet, F., Reiter, E., Gatt, A., Hunter, J., Sripada, S., Freer, Y., & Sykes, C. (2009). Automatic generation of textual summaries from neonatal intensive care data. *Artificial Intelligence*, 173 (7–8), 789–816. <https://doi.org/10.1016/j.artint.2008.12.002>
- Reiter, E., & Dale, R. (2002). Building Applied Natural Language Generation Systems. *Natural Language Engineering*, 3. <https://doi.org/10.1017/S1351324997001502>
- Thórisson, K. R. (2019). Simulated Perceptual Grouping: An Application to Human-Computer Interaction. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society*. <https://api.semanticscholar.org/CorpusID:6265048>
- Turner, D. W. (2010). Qualitative Interview Design: A Practical Guide for Novice Investigators. *The Qualitative Report*, 15(3), 754–760. Retrieved from <http://nsuworks.nova.edu/tqr/vol15/iss3/19>
- Qu, S. Q., & Dumay, J. (2011). The Qualitative Research Interview. *Qualitative Research in Accounting and Management*, 8 (3), 238–264. Available at SSRN: <https://ssrn.com/abstract=2058515>
- van der Lee, C., Gatt, A., van Miltenburg, E., & Krahmer, E. (2021). Human evaluation of automatically generated text: Current trends and best practice guidelines. *Computer Speech & Language*, 67, 101151. <https://doi.org/10.1016/j.csl.2020.101151>
- Vougiouklis, P., Elshahar, H., Kaffee, L. A., Gravier, C., Laforest, F., Hare, J., & Simperl, E. (2017). Neural Wikipedian: Generating Textual Summaries from Knowledge Base Triples. *Journal of Web Semantics*, 52–53. <https://doi.org/10.1016/j.websem.2018.07.002>
- Vougiouklis, P. (2020). Neural Generation of Textual Summaries from Knowledge Base Triples. In *Studies on the Semantic Web*. <https://api.semanticscholar.org/CorpusID:191521779>
- Yu, J., Reiter, E., Hunter, J., & Mellish, C. (2007). Choosing the content of textual summaries of large time-series data sets. *Natural Language Engineering*, 13, 25–49. <https://doi.org/10.1017/S1351324905004031>

11. Appendices

The appendices included at the conclusion of this thesis encompass a comprehensive array of supporting materials pivotal to the elucidation and evaluation of the research methodology and findings. Appendix A and Appendix B present the interview protocols and configurations that underpin the development of the proposed methodology, offering insight into the initial investigative phase. Appendix C elucidates the interview queries that explore the practical feasibility of implementing the proposed method from a technical standpoint. Appendix D, Appendix E, Appendix F detail the experimental setup and the inquiries posited to substantiate the efficacy and applicability of the proposed method. Appendix G and Appendix H delineates the consent procedures meticulously followed during the experimental phase to ensure compliance with the ethical standards mandated by Utrecht University. Appendix I furnishes the pseudo-code for the summary generation techniques as articulated by Kuptavanich (2018), Appendix J does the same for Kuptavanich et al. (2024) and Appendix K does the same for the method proposed in the thesis. Lastly, Appendix L and Appendix M showcase exemplar of summaries generated via the method developed by Kuptavanich et al. (2024) and the method proposed in the thesis, serving as empirical evidence of the comparative performance and potential advancements introduced by the proposed methodology.

Appendix A. Interview Questions - Reviewing a Paper

Topic - Generating Insights from Bibliographies

Premise - Bibliographies hold information that is beneficial to the researcher in their work.

Outcome -

- Firstly, an interview will be held to gain insight into what researchers deem necessary in their work.
- Secondly, an open sourced technical prototype will be created. It will be able to summarise bibliographies.
- Thirdly, the prototype will be thoroughly evaluated.

Interview Details - Individual answers will stay private between the interviewee and the researcher and will not be published recognisably.

Open-ended warm-up questions

- How connected to the scientific community do you consider yourself to be? In your respective field? In science in general?

Scenario-specific questions

[Stage 1: Personal workflow details]

Suppose a scientific journal asks you to review a paper that was submitted to said scientific journal. Papers typically contain a bibliography, listing the papers cited in the paper. Looking at the bibliography of the submitted paper might help you in your assessment. My conversation today focuses on such bibliographies and what makes them useful for you as a reviewer.

- When you look at a bibliography, what do you tend to do? Please describe as you see fit.
- How closely do you look at the references?
- Do you read the bibliography from start to finish?
- Is there anything important about bibliographies that I have not discussed yet?

[Stage 2: Example case]

A research paper is presented. You are given five minutes to browse the paper before answering the questions.

- Which of the references in the bibliography did you find most important? Why?
- Are there any that could or should have been left out?
- Are any important literature references missing?

[Stage 3: Summary]

Suppose the journal asks you to write a short summary of the bibliography. It can consist of anything from content to formatting to choice of authors or topics etc.

Please take your time to write down the summary with a maximum of 100 words.

Open-ended cool-down questions

We are building a tool for generating insights from bibliographies. The tool will take the bibliography as input and give back a summary as output.

- What insights would you want to see in that summary so that it helps you as a researcher in the review process?
- If you could change one thing about bibliographies or how I approach them as a scientific community, what would it be?

Appendix B. Interview Questions - Finding Research Focus

Topic - Generating Insights from Bibliographies

Premise - Bibliographies hold information that is beneficial to the researcher in their work.

Outcome -

- Firstly, an interview will be held to gain insight into what researchers deem necessary in their work.
- Secondly, an open sourced technical prototype will be created. It will be able to summarise bibliographies.
- Thirdly, the prototype will be thoroughly evaluated.

Interview details - Individual answers will stay private between the interviewee and the researcher and will not be published recognisably.

Open-ended warm-up questions

- How connected to the scientific community do you consider yourself to be? In your respective field? In science in general?

Scenario-specific questions

[Stage 1: Personal workflow details]

Suppose a university gives you an opportunity to establish a new research group. You might not have an idea immediately, so you wonder what area should the group be focusing on. For inspiration, you browse conference proceedings of different research areas. Let me assume you focus on a specific conference proceedings — a large volume containing hundreds of conference papers. It is full of work that you are not sure how to place yet. Looking at the bibliography of said conference proceedings might help you in your assessment. My conversation today focuses on such bibliographies and what makes them useful for you as a research group head.

- When you look at a conference proceedings in terms of bibliography, what do you tend to do?
- How closely do you look at the references at the end of each paper?
- Do you read the bibliographies from start to finish?
- Is there anything important about bibliographies that I have not discussed yet?

[Stage 2: Example case]

Proceedings from a recent and familiar conference are presented. You are given five minutes to browse the proceedings before answering the questions.

- Having formed a quick first impression, is there anything that struck you about the works that the authors chose to cite?
 - Which of the references in the bibliography did you find most important? Why?
 - Are there any that could or should have been left out?
 - Are any important literature references missing?
 - From the bibliographies, is it possible to form an impression of the way in which the topic of the proceedings has been studied over time?
 - For how long has this topic been studied? Is it currently more or less popular than 10 years ago?

[Stage 3: Summary]

Suppose summarising the bibliography would help you set the focus for your research group. You would like to highlight some things from the bibliographies that struck you. The summary can consist of anything from content to formatting to choice of authors or topics etc.

Please take your time to write down the summary with a maximum of 100 words.

Open-ended cool-down questions

We are building a tool for generating insights from bibliographies. The tool will take the collection of papers, extract their bibliographies for input and give back a summary as output.

- What insights would you want to see in that summary so that it helps you choose a new research focus for the new research group you establish?
 - If you could change one thing about bibliographies or how I approach them as a scientific community, what would it be?

Appendix C. Interview Questions - Summary Generation Method

Understandability for Implementation

The purpose of this interview is to validate a proposed method for generating textual summaries from bibliographies in scientific literature. The focus will be on understanding the level of understandability of the method and its technical implementability.

The interview will last between 30 minutes to 60 minutes and will cover various aspects of the method.

Interview Questions

1. How would you rate your understanding of the proposed method on a scale of 1-10?
2. Could you explain the method in your own words?
3. What parts of the method, if any, are unclear or confusing to you?
4. How technically capable do you find yourself to implement this method?
5. What challenges, if any, do you foresee in implementing this method?
6. What resources would be needed to implement this method?
7. How could this method be improved to increase its understandability and ease of implementation?
8. What are your initial thoughts on the proposed method for generating textual summaries from bibliographies in scientific literature?

Thank you for your time and valuable insights. Your feedback will be instrumental in refining and improving the proposed method for generating textual summaries from bibliographies in scientific literature.

Appendix D. Experimentation Plan

General research question: What are the specific features and elements that enhance a summary of a bibliography?

General Hypothesis: The availability of summarised bibliographies will aid scientists in their daily work.

To validate my thesis on generating textual summaries from bibliographies, I will execute an extensive experimentation plan that involves qualitative and quantitative methods. This plan will directly assess the utility, clarity, and preference of my method (Mättas summaries) compared to an alternative method (Kuptavanich summaries).

Evaluation Hypotheses

1. Mättas summaries are preferred over Kuptavanich summaries.
2. Mättas summaries are judged as clearer than Kuptavanich summaries.
3. Mättas summaries are judged as more useful than Kuptavanich summaries.

Appendix E. Experimentation Plan Part 1: Semi-Structured Interviews

Objective

I will evaluate the preference, perceived clarity, and usefulness of Mättas summaries over Kuptavanich summaries among academic professionals.

Participants

I will select a diverse group of dozen professors across various disciplines to ensure a broad range of perspectives.

Materials

I will prepare two summaries of the same bibliography using both the Mättas and Kuptavanich methods, anonymising them to prevent bias.

Procedure

1. I will present both summaries to the participants without revealing which method was used for each.
2. I will conduct semi-structured interviews, incorporating the specified comparative questions to elicit detailed feedback on each summary's informative value, structure, specific useful features, assistance in daily work, suggested improvements, and overall challenge addressing capability.
3. I will ensure to probe deeper based on the responses to gather rich, qualitative data.

Questions

- Upon reviewing these two summaries, which one do you find more informative for understanding the bibliography, and why?
- How does the structure of each summary influence your ability to quickly grasp the key information from the bibliography?
- Are there specific features or elements in one summary that you find particularly useful, which are absent in the other?
- Considering your research needs, how would each summary assist you in your daily work?
- What improvements would you suggest for each summary to better meet the needs of researchers like yourself?
- Which summary do you think better addresses the challenges of dealing with extensive bibliographies?
- How do you evaluate the comprehensiveness of the information provided in each summary?
- Do you have any concerns about the accuracy or clarity of the information presented in the summaries?
- Considering both summaries, which one would you prefer to use regularly? Choose one even if you value both summaries at exactly the same level.
 - Why did you choose this one over the other one? (Open-ended)
- Is there anything that you would like to add that has not been said so far?
- Do you agree to being mentioned in the acknowledgment section of the thesis by name?

Analysis

I will perform a thematic analysis on the interview transcripts to extract common themes related to the summaries' clarity, usefulness, and overall preference. This process will include the following steps:

1. Initial Coding: Transcripts will be coded to extract significant insights concerning the clarity and usefulness of the summaries, along with explicit preferences between the Mättas and Kuptavanich methods. Reasons behind these preferences will be noted.
2. Theme Identification: Codes will be grouped into themes that represent the summaries' clarity, usefulness, and reasons for preference. This will include aspects such as comprehensiveness, ease of understanding, and relevance to research needs.
3. Review and Refinement of Themes: Themes will be reviewed to ensure they accurately reflect interview responses, with adjustments made as necessary for coherence and completeness.
4. Definition of Themes: Each theme will be clearly defined, detailing its relation to the summaries' evaluation and preference. For example, a theme might highlight the balance between brevity and comprehensiveness as a factor in preference.
5. Integration and Cross-Theme Analysis: Preference-related findings will be integrated with clarity and usefulness themes to provide a comprehensive understanding of the summaries' perceived value. A cross-theme analysis will explore how various factors interrelate to influence preferences.

Appendix F. Experimentation Plan Part 2: Structured Web Questionnaire

Objective

I will quantitatively evaluate the broader academic community's preference, perceived clarity, and usefulness of Mättas summaries versus Kuptavanich summaries.

Participants

I will target individuals with at least a graduate degree, utilising academic networks and platforms like SurveyMonkey, Prolific for recruitment. I will gather no less than a 100 responses.

Materials

I will use the same anonymised summaries from Part 1. Additionally, two more anonymised summaries are used for each method. For direct comparison, full summaries are split into sections.

Procedure

1. I will randomise the presentation order of the summaries to each participant to control for order effects.
2. I will ask participants to rate each summary based on clarity and usefulness using a Likert scale.
3. I will ask participants to compare two summaries for direct comparison.
4. I will provide space for open-ended feedback to capture qualitative insights that complement the quantitative data.

Questions

- How often do you read scientific literature? Choose one option:
 - Weekly / Frequently
 - Monthly / Occasionally
 - Yearly / Rarely
 - Never
- How many research papers have you published? Choose one option:
 - 0
 - 1
 - 2-5
 - 5+
- Rate the clarity of Summary 1 on a scale from 1 (Very Unclear) to 5 (Very Clear).
- Rate the clarity of Summary 2 on a scale from 1 (Very Unclear) to 5 (Very Clear).
- On a scale from 1 (Not Useful) to 5 (Extremely Useful), how would you rate the usefulness of Summary 1 for your research work?
 - On a scale from 1 (Not Useful) to 5 (Extremely Useful), how would you rate the usefulness of Summary 2 for your research work?
- Which summary (1 or 2) do you find more helpful in providing a quick overview of the bibliography? Choose one.
- Considering both summaries, which one would you prefer to use regularly? Choose one even if you value both summaries at exactly the same level.
 - Why did you choose this one over the other one? (Open-ended)
 - What information do you think is missing or could be improved in Summary 1? (Open-ended)
 - What information do you think is missing or could be improved in Summary 2? (Open-ended)

Analysis

I will employ a non-parametric test suitable for paired data. I will compile the ratings for clarity, usefulness, and preference given by each participant for both the Mättas and Kuptavanich summaries. I will use the Wilcoxon Signed-Rank Test to compare the ratings for each summary pair provided by the same participant. This test does not assume normal distribution of the data and is suitable for Likert scale data, making it ideal for this analysis.

Analysis Steps

1. Compile Ratings: I will gather clarity, usefulness, and preference ratings for both the Mättas and Kuptavanich summaries from each participant.
2. Determine Differences: I will calculate the difference in scores for clarity, usefulness, and preference between the Mättas and Kuptavanich summaries for each participant.
3. Rank Differences: I will rank these differences, ignoring signs, from smallest to largest, to prepare for the Wilcoxon Signed-Rank Test.
4. Assign Signed Ranks: I will give ranks their appropriate signs based on whether the difference favoured Mättas or Kuptavanich summaries.

5. Summarise Ranks: I will aggregate the positive and negative ranks separately to form the basis of the statistical test.

6. Statistical Testing: I will apply the Wilcoxon Signed-Rank Test to the ranked differences to assess whether there are statistically significant differences in clarity, usefulness, and preference ratings between the two types of summaries.

Interpretation

A significant outcome from the Wilcoxon Signed-Rank Test across any of the aspects (clarity, usefulness, preference) will indicate a meaningful difference in the perceived quality and favourability of the Mättas summaries in comparison to the Kuptavanich summaries. This result will help validate or refute the hypothesis that Mättas summaries are preferred and perceived as clearer and more useful than Kuptavanich summaries, providing a quantitative basis to support the qualitative findings.

Appendix G. Consent Procedure for Interviews

Welcome to the research interview. This interview is part of a study for a Masters Thesis at Utrecht University. The core of the thesis is a method which aims to offer a comprehensive view of bibliographies, enhancing literature comprehension.

We are conducting this study to better understand how I can help people who read scientific papers (in their work). I am interested in your views on how summaries of scientific bibliographies can be made and used. I will show you some examples of these summaries, generated from real scientific articles, and ask what you think about them.

We highly value your contribution to this research. Completing this interview is expected to take about 30 minutes of your time. Your thoughtful responses will play a pivotal role in enhancing the effectiveness of bibliographic summarisation in scientific literature.

Anonymous Participation

Your answers will not be disclosed explicitly but rather used for thematic analysis by me.

Voluntary Participation

I want to assure you that your participation in this interview is entirely voluntary. You are under no obligation to take part, and if you choose not to participate, it will not have any negative consequences for you.

Informed Consent

Before I begin, I need to obtain your consent to participate in this interview. By continuing with the interview, you are indicating your understanding of the research's purpose and your willingness to participate.

Withdrawal

Please remember that you have the right to withdraw from this interview at any time and for any reason. If at any point you feel uncomfortable or wish to stop, please let me know, and I can conclude the interview immediately.

Opportunity to Ask Questions

Before I proceed, do you have any questions about the research or the interview process? If you have any concerns or need clarification on anything, please feel free to ask. Additionally, I will provide you with my contact information so you can reach out later if needed.

Time to Consider

Please take a moment to consider whether you wish to proceed with the interview. There is no rush, and I want to ensure that you are comfortable before I begin.

Contact Information

If you have any further questions or concerns about the research or your participation, you can reach me at my Utrecht University email address o.mättas@students.uu.nl, the address I used to ask you to participate in the interview.

Appendix H. Consent Procedure for Web Questionnaire

Welcome to the research questionnaire. This questionnaire is part of a study for a Masters Thesis at Utrecht University. The core of the thesis is a method which aims to offer a comprehensive view of bibliographies, enhancing literature comprehension.

We are conducting this study to better understand how I can help people who read scientific papers (in their work). I am interested in your views on how summaries of scientific bibliographies can be made and used. I will show you some examples of these summaries, generated from real scientific articles, and ask what you think about them.

We highly value your contribution to this research. Completing this questionnaire is expected to take about 20 minutes of your time. Your thoughtful responses will play a pivotal role in enhancing the effectiveness of bibliographic summarisation in scientific literature.

Voluntary Participation

Your participation in this questionnaire is completely voluntary. You are not obligated to take part, and if you choose not to participate, it will not have any negative consequences for you.

Informed Consent

Before you proceed with the questionnaire, I need to obtain your consent to participate. By clicking the "Agree" button below, you are indicating your understanding of the research's purpose and your willingness to participate.

Withdrawal

Please remember that you can exit the questionnaire at any time and for any reason. If you decide to stop participating, simply close the browser window, and your responses will not be saved.

Opportunity to Ask Questions

If you have any questions about the research or the questionnaire, please don't hesitate to contact me at o.mattas@students.uu.nl. I will be happy to provide clarification or address any concerns you may have.

Time to Consider

Please take a moment to review the information provided and consider whether you wish to proceed with the questionnaire. There is no rush, and I want to ensure that you are comfortable before continuing.

Parental/Legal Guardian Consent

If you are under the age of 18, please make sure you have obtained consent from your parent or legal guardian before proceeding with the questionnaire.

Appendix I. Pseudo-code for Summarisation Algorithm for Consumer Products

Introductory Paragraph Pseudo-Code:

```
input: data of prices and products (including product size and attributes)
output: range of prices and median price
```

```
def price_range_and_median(data):
    # exclude outliers and round prices to the nearest 5
    prices = data["price"].round(-1)
    prices = prices[(prices - prices.median()).abs() <= (3 * prices.std())]
```

```
# calculate the range of prices and median price
min_price = prices.min()
max_price = prices.max()
median_price = prices.median()
```

```
# report the range of prices and median price
print("For", data["product_size"].iloc[0], "inch TVs, the price of most products (",
len(prices), "out of", len(data), "models) falls in the range of", min_price, "to",
max_price, "pounds with a median price of about", median_price, "pounds.")
```

Collective Description of Products Pseudo-Code:

```
input: data of products and their attributes
output: common features of products
```

```
def common_features(data):
    # find the 7 most common features of the products
    features = data.columns.tolist()
    features_count = {}
    for feature in features:
        features_count[feature] = data[feature].value_counts().iloc[0]
    common_features = [k for k, v in sorted(features_count.items(), key=lambda item:
item[1], reverse=True)][:7]
```

```
# report the common features of the products
print("Most", data["product_size"].iloc[0], "inch TVs have the following features:", ",
".join(common_features))
```

Highlighting Important Features Pseudo-Code:

```
input: data of prices and products (including product size and attributes)
output: features with a strong impact on price
```

```
def important_features(data):
    # calculate the average price for each subset of a feature
    features = data.columns.tolist()
    features.remove("price")
    features_price = {}
    for feature in features:
        feature_subsets = data.groupby(feature)["price"]
        feature_subsets_avg = feature_subsets.mean()
```

```
features_price[feature] = feature_subsets_avg

# calculate the standard deviation of the average prices for each subset of a feature
features_price_sd = {}
for feature, feature_price in features_price.items():
    features_price_sd[feature] = feature_price.std()

# find the features with the highest standard deviation
important_features = [k for k, v in sorted(features_price_sd.items(), key=lambda item:
item[1], reverse=True)][:7]

# report the features with a strong impact on price
print("The features that have a strong impact on the price of",
data["product_size"].iloc[0], "inch TVs are:", ", ".join(important_features))
```

Appendix J. Pseudo-code for Summarisation Algorithm for Bibliographies (Kuptavanich Method)

Introductory Paragraph Pseudo-Code:

```
input: data of references (including venue type and number of citations)
output: total number of references and most common venue type
```

```
def reference_summary(data):
    # Calculate the total number of references
    total_references = data["references"].count()

    # Calculate the most common venue type
    venue_type = data["venue_type"].value_counts().idxmax()
    venue_type_percentage = data["venue_type"].value_counts(normalize=True).max() * 100
```

Quantifiers Pseudo-Code:

```
input: data of references (including publication year and self-citation)
output: range and center of publication year and self-citation percentage
```

```
def publication_year_and_self_citation(data):
    # Calculate the range and center of publication year
    year_range = data["publication_year"].agg(["min", "max"])
    year_center = data["publication_year"].mean()

    # Calculate the percentage of self-citations
    self_citation_count = data[data["self_citation"] == True]["self_citation"].count()
    self_citation_percentage = (self_citation_count / data["references"].count()) * 100
```

Venue Subdomains' Top Publications Pseudo-Code:

```
input: data of references (including venue type and number of citations)
output: detailed subdomains of venues and reference with highest citation count for each subdomain
```

```
def venue_subdomains_top_publications(data):
    # Group references by venue type and calculate the reference with highest citation count for each subdomain
    grouped_data = data.groupby("venue_type").agg({"citations": "idxmax"})
```

List of Authors with Highest Citation Count Pseudo-Code:

```
input: data of references (including authors and number of citations)
output: top 7 authors with highest citation count
```

```
def authors_with_highest_citation_count(data):
    # Calculate the top 7 authors with highest citation count
    top_authors = data.groupby("author").agg({"citations": "sum"}).nlargest(7, "citations")
```

Appendix K. Pseudo-code for Summarisation Algorithm for Bibliographies (Mättas Method)

Introductory Section Pseudo-Code:

1. Read bibliography data
2. Count the total number of sources

Timeline Section Pseudo-Code:

1. Read bibliography data
2. Extract publication years from sources
3. Calculate the frequency-weighted average publication year
4. Count sources published before the frequency-weighted average year
5. Identify the earliest and latest publication years

Authors Section Pseudo-Code:

1. Read bibliography data
2. Extract author information and affiliations from sources
3. Count authors and their contributions
4. Identify top 3 most referenced authors
5. Identify top 3 organisations based on author affiliations
6. Determine primary research regions based on affiliations
7. Count self-references

Topics Section Pseudo-Code:

1. Read bibliography data
2. Extract topic information for sources via programmatic means (APIs)
3. Identify top 3 most prevalent fields of study and topics

Character Section Pseudo-Code:

1. Read bibliography data
2. Extract source types (e.g., journal, conference proceedings)
3. Identify top 3 most represented types of research

Familiarity Section Pseudo-Code:

1. Read bibliography data
2. Extract author information from sources
3. Compare authors with the reader's previous collaborators
4. Identify any authors that the reader has collaborated with

Appendix L. Generated Summary (Kuptavanich Method)

Introduction

There are 150 publications in the bibliography.

Timeline

Comparing to the paper publication year of 2015, most of the references are quite recent, dating from 1990 to 2021 with a centring in 2010. Among these, 15 (10%) are self-references.

Authors

Authors with highest citation counts appeared in the references, in descending order, are Dr. Alice Smith, Dr. Bob Johnson.

Topics

The three most popular fields of study are Artificial Intelligence (35% of papers), Quantum Computing (25% of papers), and Molecular Biology (20% of papers). As the three most popular topics, researchers have used "Machine Learning" (30% of papers), "Quantum Entanglement" (22% of papers), "DNA Sequencing" (18% of papers) as keywords to illustrate the content of their work.

Character

Most references (40%) are from Journal Articles.

Appendix M. Generated Summary (Mättas Method)

Introduction

There are 150 sources in the bibliography.

Timeline

The weighted average publishing year among all the sources is 2015. 70 of the sources were published before that time, with the earliest source dating back to 1990. The latest source was published in 2021.

Authors

There are 85 authors listed for the sources in the bibliography. 30% of authors have published 60% of the work. The three most referenced authors are Dr. Alice Smith (with 15 publications), Dr. Bob Johnson (with 12 publications), and Dr. Carol Lee (with 10 publications).

Most authors cited in this paper come from MIT (18 authors), closely followed by Stanford University (16 authors) and Harvard University (12 authors).

Based on the authors' affiliations with universities, I can analyse the global distribution of sources. 40% of the authors are/were active in Europe (primarily in the UK and Germany), 30% of the authors in North America (mostly in the USA), 20% of the authors in Asia (notably in China and Japan), with 10% of the authors representing the rest of the world. Among all the sources, 15 (10%) were self-references.

Topics

Most frequently described fields of study are Artificial Intelligence (35% of sources), Quantum Computing (25% of sources), and Molecular Biology (20% of sources). As the three most frequently described topics, researchers have used "Machine Learning" (30% of sources), "Quantum Entanglement" (22% of sources), "DNA Sequencing" (18% of sources) as keywords to illustrate the content of their work.

Character

The three most frequently used types of sources are Journal Articles (40% of sources), Conference Proceedings Articles (30% of sources), and Book Chapters (20% of sources).

Familiarity

As you have uploaded an example of your own published research, I have found that you have previously collaborated with the following authors from the bibliography: Dr. Alice Smith, Dr. Bob Johnson, and Dr. Carol Lee.