



LUNDS UNIVERSITET

Ekonomihögskolan

Institutionen för informatik

Artificial Intelligence in Data Analytics

A Qualitative Study on how Data Analysts perceive and experience the usage of AI in Data Analytics

Kandidatuppsats 15 hp, kurs SYSK16 i Systemvetenskap.

Författare: Tony Do
Adam Jacobson

Handledare: Niki Chatzipanagiotou, PhD-Senior Lecturer

Rättande lärare: Blerim Emruli
Nam Aghaee

ENGELSK TITEL: Artificial Intelligence in Data Analytics: A Qualitative Study on how Data Analysts perceive and experience the usage of AI in Data Analytics – SVENSK TITEL: Artificiell Intelligens inom Data Analys: En kvalitativ studie om hur Data Analytiker uppfattar och upplever användandet av AI inom dataanalys

FÖRFATTARE: Tony Do, Adam Jacobson

UTGIVARE: Institutionen för informatik, Ekonomihögskolan, Lunds universitet

EXAMINATOR: Osama Mansour, Docent

FRAMLAGD: Maj, 2024

DOKUMENTTYP: Kandidatuppsats

ANTAL SIDOR: 99

NYCKELORD: Data Analyst, Data Analytics, Artificial Intelligence, AI-Based Tools, Sensemaking Theory, Qualitative Research.

SAMMANFATTNING:

The bachelor's thesis explores the utilization of Artificial Intelligence (AI) in data analytics, aiming to understand data analysts' perceptions and experiences with AI-based tools and technologies. Through qualitative research involving semi-structured interviews with six data analysts from Sweden and the United States, thematic analysis identified five key themes. These findings were discussed in light of a theoretical framework encompassing data analytics concepts, AI in data analytics, accountability in decision-making AI tools, and the sensemaking theory. The findings show that data analysts recognize AI's benefits, particularly in error reduction, time efficiency and bias mitigation, yet acknowledge its inability to fully replace human analysts. They predominantly rely on specialized software to mitigate risks associated with multi-purpose AI tools and utilize self-developed AI solutions to enhance machine learning benefits. Concerns persist regarding limitations, data integrity, and the need for ongoing education, emphasizing the importance of accountability and effective communication among stakeholders. While AI adoption in data analytics is growing, manual workflows remain prevalent, reflecting cautious optimism and ongoing evaluation of AI's capabilities and challenges. Theoretical contributions include insights into the dynamics surrounding AI integration, while practical implications underscore the effective utilization and integration of AI-based technologies in data analytics workflows.

Kandidatuppsatsen undersöker användningen av artificiell intelligens (AI) inom dataanalys med syftet att förstå dataanalytikernas uppfattningar och erfarenheter av AI-baserade verktyg och teknologier. Genom kvalitativ forskning med intervjuer av sex dataanalytiker från både Sverige och USA identifierades fem huvudteman genom tematisk analys. Dessa resultat diskuterades med utgångspunkt från ett teoretiskt ramverk som omfattar begrepp inom dataanalys, AI inom dataanalys, ansvarsutkrävande vid beslutsfattande med AI-verktyg och "sensemaking"-teorin. Resultaten visar att dataanalytiker erkänner fördelarna med AI, särskilt vad gäller minskning av misstag och fördomar samt tidsbesparingar, men de erkänner också dess begränsningar i att helt ersätta mänskliga analytiker. De förlitar sig huvudsakligen på specialiserad programvara för att hantera specifika arbetsflöden, och använder multipurpose-verktyg för assistans i deras övriga analytiska arbete. De använder dessutom egna AI-lösningar för att ta del av fördelarna med maskininlärning. Oro kvarstår angående begränsningar, dataintegritet, vilket betonar vikten av ansvarsutkrävande och effektiv kommunikation mellan intressenter. Trots att AI-användningen inom dataanalys växer är manuell analys fortfarande en stor del av det analytiska arbetet överlag, vilket återspeglar en försiktig optimism och en pågående utvärdering av AI:s förmågor och utmaningar. "Sensemaking"-teorins bidrag inkluderar insikter i dynamiken kring AI-integration, medan praktiska konsekvenser understryker vikten av att använda och integrera AI-baserade teknologier på ett effektivt sätt i arbetsflöden för dataanalys.

Acknowledgments

We would like to thank our supervisor Niki Chatzipanagiotou for guiding us through the bachelor's thesis. We could not have asked for a better supervisor. Thank you for your major contribution, knowledge and support of our work.

We would also like to express our gratitude to the respondents who without hesitation gave their time and willingly participated in this study. Your valuable insights, high competence and willingness to contribute have been essential for our research. Thank you for being an integral part of this bachelor's thesis.

A special thank you to Nenad Brdarski, who facilitated and supported our research prior to its initiation.

May, 2024

Adam Jacobson & Tony Do

Table of Contents

1. Introduction	1
1.1 Background	1
1.2 Problem Identification and Previous Research	1
1.3 Research Purpose and Research Question	2
1.4 Delimitations	3
2. Literature Review	4
2.1 Search procedure	4
2.2 Data Analytics	7
2.2.1 Definition	7
2.2.2 Steps in Data Analytics	7
2.2.3 Numerical Data Analytics	8
2.2.4 Non-Numerical Data Analytics	8
2.2.5 Types of Data Analytics techniques	9
2.3 Artificial Intelligence	9
2.3.1 Definition	9
2.3.2 Machine Learning	10
2.4 AI in Data Analytics	10
2.4.1 The Evolution of AI in Data Analytics	10
2.4.2 The benefits of AI usage in Data Analytics	11
2.4.3 The Concerns of AI Usage in Data Analytics	11
2.5 AI hallucination	12
2.6 AI Based Tools and Technologies	12
2.6.1 Natural Language Processing	12
2.6.2 Generative Pre-Trained Transformers	12
2.6.3 AI in Business Intelligence	13
2.6.4 Natural Language Processing in Business Intelligence	13
2.6.5 No Code AI tools	13
2.6.6 AI assistants	13
2.7 Accountability in decision making AI tools	14
2.7.1 Perception of Moral Responsibility of AI	14
2.7.2 Designing Responsible AI based systems	14
2.8 Summary of the Literature Review Findings	15
2.9 Sensemaking Theory	15
3. Methodology	17
3.1 Research approach	17

3.2 Method of Data Collection	17
3.2.1 Semi-structured interviews	17
3.2.2 Selection Criteria, Sample Technique, and Respondents	17
3.3 Method of Data Analysis	18
3.4 Validity and Reliability	19
3.5 Ethical considerations	20
4. Empirical Findings	21
4.1 Theme 1: Data Analytics	21
4.1.1 Subtheme: Numerical data	21
4.1.2 Subtheme: Non-numerical data	22
4.2. Theme 2: AI-based tools in Data Analytics	23
4.2.1 Subtheme: AI tools with specific purposes	23
4.2.2 Subtheme: The use of GPT	24
4.2.3 Subtheme: AI in Business Intelligence	25
4.3 Theme 3: AI benefits	26
4.4 Theme 4: AI Concerns	27
4.4.1 Subtheme: Data challenges & Risks	27
4.4.2 Subtheme: Limitations	28
4.4.3 Subtheme: AI hallucination	29
4.5 Theme 5: Accountability	29
5. Discussion	31
5.1 Data Analytics	32
5.1.1 Numerical Data Analytics	33
5.1.2 Non-numerical Data Analytics	33
5.2 AI-based tools in Data Analytics	34
5.2.1 Tailoring easy to use AI to fit specific analytical needs	34
5.2.2 Using GPT in Data Analytics	34
5.2.3 Business Intelligence (BI) can benefit from AI advancements	35
5.3 AI Benefits	36
5.3.1 Enhanced Efficiency and Time Savings	36
5.3.2 Enhanced Precision and Reliability	36
5.3.3 Complementary Role of AI in Elevating Human Analysts	37
5.4 AI Concerns	37
5.4.1 Data challenges & Risks	37
5.4.2 Limitations	37
5.4.3 AI Hallucinations	38
5.6 Discussion of the Findings with Sensemaking Theory	39
6. Conclusion	41
6.1 Conclusions	41
6.2 Contributions	41
6.3 Future Research	42

References	43
Appendices	46
Appendix A. Interview Guide	46
Appendix C. Interview 1	48
Appendix D. Interview 2	60
Appendix E. Interview 3	67
Appendix F. Interview 4	72
Appendix G. Interview 5	80
Appendix H. Interview 6	88
Appendix I. AI statement	99

List of Tables

Table 1: Literature Review Articles	5
Table 2: Respondent Overview	18
Table 3: Theme and Subthemes Overview	19
Table 4: Findings Addressing the Research Question	31

1. Introduction

1.1 Background

The growth of data has increased significantly in recent years due to technological advancements, digitalization, social trends and strategic business goals. This development has transformed how decision-making within organizations is conducted (Harvard Business Review et al., 2019). Clake, George & Halawi (2022) explain data analytics as the process of utilizing machine learning and statistical analysis to extract pertinent data from datasets, to identify patterns and insights for informed decision making.

Artificial Intelligence (AI) describes a machine's ability to perform tasks that typically require human intelligence. Tasks like learning, reasoning, problem solving, and making decisions (Ertel, 2017). Implementing AI has a wide range of potential advantages, including increasing productivity, boosting creativity, and tackling complex problem solving in various fields and industries (Russel & Norvig, 2022).

The field of data analytics, which has been primarily a manual and time-consuming task, has gone through significant advancements with the advent of AI technologies. Leveraging AI within data analytics has revolutionized traditional processes by automating tasks, identifying complex patterns, and providing valuable insights at exceptional rates (Russel & Norvig, 2022). To fully leverage the advantages of AI in data analytics, organizations rely on the expertise of data analysts to ensure the effective implementation and utilization of the technology. Data analysts help bridge the knowledge gap between raw data and useful insights by applying their subject-matter expertise and analytical abilities. Thus, the exponential growth of data and artificial intelligence technology has made data analysts more crucial to enterprises.

1.2 Problem Identification and Previous Research

By implementing AI within data analytics processes, it presents enhanced business opportunities for organizations by offering the chance of improving decision-making and gaining a competitive edge in today's data-driven business environment. Nevertheless, despite the potential benefits associated with the new technology, the implementation of AI is not without its challenges as explained below.

A key challenge is comprehending the perception and experience of data analysts when they adopt AI technologies into their work practices. Data analysts have an in-depth knowledge of both the technical intricacies and practical implications of data analytics, enabling them to effectively connect AI technologies with organizational business requirements. Their viewpoint is therefore essential for successfully implementing and maximizing the value derived from AI technologies in data analysis processes in enterprises (Porter et al., 2018). The perspectives and experience of data analysts can have a significant impact on successful

AI initiatives in businesses, by affecting factors such as user acceptance, adoption rates and the utilization of AI-generated insights in decision-making processes (Iansiti & Lakhani, 2020). Organizations can adjust their strategies to encourage innovation, knowledge sharing, and encourage collaboration between data analysts and AI tools by learning more about how data analysts view and experience AI implementation in data analytics (Iansiti & Lakhani, 2020).

The impact of utilizing various AI tools in data analytics has been studied from a number of different industries in previous research. Saivasan and Lokhande (2023) examine the potential of Generative AI (GAI) tools to transform financial analytics. Through the application of GAI in financial analytics, businesses can improve their capacity for assessing risk, derive meaningful insights from large datasets, and make more precise judgments.

In healthcare, the use of GAI tools in analyzing datasets can improve the predictive capabilities and address numerous crucial challenges within the industry. According to Gupta (2023), GAI tools can be the disruptive factor in solving problems like data shortages, privacy issues and the intrinsic unpredictability of medical problems in healthcare. Researchers and practitioners in the healthcare sector can collaborate together more efficiently by using AI to create synthetic data. This allows for the design of algorithms that are applicable worldwide while maintaining patient privacy (Gupta, 2023).

However, besides the benefits, the use of AI in data analytics raises some ethical and privacy concerns, as identified by Aldboush and Ferdous (2023). Their study provides useful advice for organizations on how to efficiently manage and mitigate those risks.

These studies show the possible benefits and risks of utilizing AI in the field of data analytics. However, there remains a gap in understanding the overall influence of AI tools, and the experiences and perspectives of those who work directly with the technology. While studies have examined the impact of AI tools in organizations and risks associated with the implementation, there have been few who have delved into the nuanced perceptions and experiences of the individuals directly involved in the data analytics processes. Addressing this gap will help us acquire a better understanding of the factors that either promote or inhibit the successful integration of AI in data analysis procedures. To enhance organizational objectives, streamline workflow processes, and maximize the benefits of AI technologies, it is necessary for businesses to comprehend the perceptions and experiences of data analysts during the adoption of these technologies.

1.3 Research Purpose and Research Question

The research purpose of this bachelor's thesis is to investigate how data analysts utilize various AI tools in their data analytics work. As well as exploring how they perceive and experience the implementation and impact of AI in data analytics. By exploring their perspective this study aims to provide valuable insights on how the use of AI affects data analytics. Hence, the following is our research question:

How do data analysts perceive and experience the usage and impacts of AI in data analytics?

1.4 Delimitations

In this bachelor's thesis we focus on the impact of AI in data analytics from the perspective of data analysts. This means that the study adopts individuals' point of view, and not an organization's point of view. The study was conducted with data analysts situated in companies based in Sweden and the United States, which may have limited the findings' applicability to other areas or cultural contexts outside of these two nations.

2. Literature Review

2.1 Search procedure

The literature review was conducted via the website of Lund University Library (LUBsearch) to access a number of academic databases related to our bachelor's thesis. The main subject area of the selected databases was information systems, in order to find the most relevant data. The databases selected were mainly Emerald, ACM, and Springer, due to their reputation including reputable publications. Within these databases, we searched for keywords related to our main themes of our bachelor's thesis which was "Data Analytics" and "Artificial Intelligence", and also combined these keywords with the words associated with usage of these themes, such as "foundation", "implementing", "benefits" and "challenges".

We restricted our search to peer-reviewed research publications published in academic journals, conferences, and books. The main sources used were published within 10 years to assure the relevance of the information. Sources older than 10 years were only used when describing different theories. The literature review search concluded with the following scholarly material presented in Table 1.

Table 1. Literature Review Overview

Article	Source/Database	Motivation
Annasingh, F. & Sesay, J.B. (2022). Data Analytics for Business Foundations and Industry Applications. Routledge, Taylor & Francis Group.	Emerald	Article about Data Analytics
Zheng, Y. (2019). A comparison of tools for teaching and learning data analytics. In Proceedings of the 20th Annual Conference on Information Technology Education (SIGITE 2019), pp. 160. Association for Computing Machinery, Inc.	ACM	Conference paper about tools in Data Analytics
Kotronoulas, G., Miguel, S., Dowling, M., Fernández-Ortega, P., et al., & Papadopoulou, C. (2023). An Overview of the Fundamentals of Data Management, Analysis, and Interpretation in Quantitative Research. Seminars in Oncology Nursing.	LUBSearch	Article about defining Quantitative / Numerical data analysis
Dolan, H.R., Alvarez, A.A., Freylersythe, S.G., & Crane, T.E. (2023). Methodology for Analyzing Qualitative Data in Multiple Languages. Nursing Research, 72(5), pp. 398-403	LUBSearch	Article defining Qualitative / Non numerical data analytics.
Martinez, R. (2019). Artificial Intelligence: Distinguishing between Types & Definitions. Nevada Law Journal, Vol. 19, Iss. 3, pp. 1015-1042.	LUBSearch	Defining AI less based on advanced technical concepts, but rather defining how AI relates to human intelligence. Can AI replace human analysts?
Huawei Technologies Co., Ltd. (2023). Artificial Intelligence Technology. 1st ed. Springer Nature Singapore: Imprint: Springer.	Springer	Defining machine learning in an e-book
Feeney, O. Murphy, B. (2023). AI, Data Analytics and the Professions. In The Future of Work: Challenges and	Springer	E-book about the future of IT related fields. The chapter delves into the history of Data Analytics and how AI has

Prospects for Organisations, Jobs and Workers (pp. 35-51). Cham: Springer International Publishing.		changed the profession, and might change it.
Beutel, G., Geerits, E., & Kielstein, J.T. (2023). Artificial hallucination: GPT on LSD? <i>Critical Care</i> , 27(1).	Springer	Study made on AI hallucination, defining and explaining the term.
Sen, M., Sen, S.N., & Sahin, T.G. (2023). A New Era for Data Analysis in Qualitative Research: ChatGPT! <i>Shanlax International Journal of Education</i> , 11(1), pp. 1-15.	LUBSearch	Many of our participants used chatGPT specifically, and therefore we felt the need to include chatGPT.
Edge, D., Larson, J., & White, C. (2018). Bringing AI to BI: Enabling Visual Analytics of Unstructured Data in a Modern Business Intelligence Platform. In <i>Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems</i> .	ACM	Delved into how AI can change BI for many years to come, which was an occurring theme.
Liu, C. & Downing, C. (2024). Using Text Analytics AI Insights in Microsoft Power BI Desktop to Score Sentiments, Extract Key Phrases, and Discover Unstructured Data Patterns. <i>Journal of Information Systems Education</i> , Winter 2024, Vol. 35, Issue 1, pp. 48-55.	LUBSearch	Although the example in the article is in the work process of BI, the overall gist is about no-code, and extracting keywords for non-numerical analysis.
John, R.J.L., Bacon, D., Chen, J., Ramesh, U., Li, J., Das, D., Claus, R., Kendall, A., & Patel, J.M. (2023). DataChat: An Intuitive and Collaborative Data Analytics Platform. In <i>Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2023</i> . Association for Computing Machinery.	ACM	Delved into an easy-to-use AI based data analytics platform. One of many AI tools and technologies

Lima, G., Grgić-Hlača, N., & Cha, M. (2021). Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making.	ACM	Studies public perception in AI decision making systems, which aligns with a theme in our thesis
Sattlegger, A., Van den Hoven, J., Bharosa, N., Hagen, L., Solvak, M., & Hwang, S. (2022). Designing for Responsibility. In Proceedings of the 23rd Annual International Conference on Digital Government Research (dg.o 2022), Virtual Event, South Korea, pp. 499-499.	ACM	Similar to the reference above, but delves more into the developers' perspective.

2.2 Data Analytics

2.2.1 Definition

Annansingh and Seay (2022) define data analytics as the process of evaluating unprocessed data to identify patterns, trends and insights that can help support decision-making. The term "data analytics" is broad and covers a wide range of data analysis techniques. Data analytics techniques may be applied to any kind of information to get insights that can be leveraged to make improvements (Annansingh and Seay, 2022).

Through the use of various software tools, data analysts are able to collect, store, process and report data, which would otherwise be lost in the volume of information (Annansingh and Seay, 2022). Spreadsheets, programming languages for manipulating databases, and visualization software for compiling and analyzing data via dashboards and reports are all examples of data analytics tools (Zheng, 2019).

A large number of data analytics methods and procedures have been mechanized into mechanical procedures and algorithms that operate on unprocessed data for human use. Many organizations, including corporations, may utilize data analytics to maximize profit, streamline operations, improve performance, or help them make more strategically sound decisions (Annansingh and Seay, 2022).

2.2.2 Steps in Data Analytics

Annansingh and Seay (2022) describes the data analytics process in the following steps:

1. *Discovery*: The first step entails gathering information and data from a wide range of sources. Finding out if there is adequate information is an essential aspect of the discovery process. If data is collected from different sources, different types of data are reconstructed into the same format so that they may ultimately be examined.

2. *Data storage and preparation:* To store, manipulate, and grant access to the collected data, a database is necessary. Hence, building a database of this kind to handle the data is the next stage in the analytics process.
3. *Data modeling:* The third step is the analysis stage. It involves interpreting the collected and stored data, to create models, which will highlight trends and insights that can be used to interpret future data. Python and R, or other open-source programming languages help achieve this.
4. *Evaluation:* A crucial component of machine learning is assessing a data mining technique's success. Model to model variations might exist in evaluation techniques. However, regression, clustering, and classification are the most widely used data processing methods.
5. *Communication:* The final stage of the process is to arrange the data so stakeholders and decision-makers can access and comprehend it. Adequate data presentation is crucial to assess the accuracy of the results and make sure the project meets all user criteria.

2.2.3 Numerical Data Analytics

According to Kotronoulas, Papathanassoglou, and Burns-Cunningham (2023), the capacity to analyze numerical data to identify trends, connection seen using statistics like mean, median, and mode (Kotronoulas et al., 2023). Inferential statistics, on the other hand, go deeper, examining correlations between variables and drawing conclusions about the larger population from sample data.

It is important to acknowledge that statistics function as a means of measuring correlations to some extent, since it's only possible to factor in certain variables such as relationships, differences, or consequences to a certain degree (Kotronoulas et al., 2023).

2.2.4 Non-Numerical Data Analytics

In contrast to quantitative data analytics, qualitative data analytics revolves around non-numerical data that are systematically examined and interpreted to find themes, patterns, and interpretations. According to Dolan et al. (2023), it is the process of comprehending and making meaning of the individualized experiences, viewpoints, and actions of people or groups. With this strategy, researchers can investigate different patterns in more detail while capturing subtleties and contextual elements that may be missed by quantitative approaches.

Working with non-numerical data consists of analyzing qualitative data, which includes data transcription, independent and collaborative coding, analysis and interpretation, and distribution of results (Dolan et al. 2023). The process of analyzing qualitative data is experimental and iterative, requiring close consideration to the context provided, and viewpoints of different types of data.

2.2.5 Types of Data Analytics techniques

According to Annansingh and Seay (2022) there are four main techniques to analyze data:

1. *Descriptive analytics*: This category involves analyzing events that have transpired over a certain period of time. The main objective is to give insights into trends, patterns, and linkages by gathering and merging data. It benefits a wide range of corporate audiences by presenting data in a readily readable style. It seeks to find issues and possibilities inside current processes and functions by utilizing methods like drill down and online analytical processing. The technique seeks to answer, for example, “What has transpired” and “Are sales stronger this quarter compared to the last?”.
2. *Predictive analytics*: In order to predict future patterns or outcomes, predictive analytics use statistical algorithms and machine learning techniques to analyze previous data. Its goal is to forecast future events by utilizing patterns and connections found in the data. Predictive analytics aims to answer, for instance, “What could happen, and why will it happen?” and “How will the summer's weather affect ice cream sales?”
3. *Prescriptive analytics*: Prescriptive analytics suggests a more advanced way to optimize results than predictive analytics, going beyond prediction suggesting a course of action. It utilizes complex algorithms, modeling techniques, and optimization models to promote actionable insights and solutions. It seeks to answer, for example, “Why should we do it?” and “What should we do?”
4. *Diagnostic analytics*: Aims to find patterns or correlations in the data to determine why specific events happened. This technique is mainly used when assessing performance. It seeks to discover, for instance, “Did the weather affect beer sales?” and “How did that latest marketing campaign impact sales?”

2.3 Artificial Intelligence

2.3.1 Definition

Martinez, R. (2019) research tries to distinguish the different terms and contexts that relate to Artificial Intelligence. AI is the study and development of computer systems and programs that display behavior or carry out tasks that ordinarily need human intelligence. This includes a broad range of tasks, such as complex communication, pattern identification, problem-solving and decision-making.

One key aspect in defining AI is the intelligence part itself, as it is often associated with human characteristics. In the past, artificial intelligence research has concentrated on developing computers that can simulate human behavior and intelligence, first through chess games and problem-solving exercises. It's important to note the distinction between systems that mimic human intelligence, and those that truly exercise it, the human race.

The development of increasingly complex systems has resulted from advances in AI, including machine learning, complicated communication, and pattern recognition. Advances in this field strive to duplicate and, in certain situations, exceed human capabilities.

Different definitions and understandings within the field have resulted in difficulty in regard to defining intelligence itself as well as the variety of approaches to AI development. Where it's easy to define the artificial part of AI, the intelligence part is much more difficult (Martinez, 2019).

2.3.2 Machine Learning

Within the field of artificial intelligence, machine learning focuses on creating algorithms that, through practice, become more proficient at a particular activity (Huawei, 2023). Machine learning techniques enable a computer to learn from examples or data rather than having to be explicitly programmed to carry out a task.

For example, in the task of identifying junk email, traditional programming methods may struggle to find a set of rules that accurately distinguish between junk and regular emails. Instead, machine learning algorithms can automatically learn from examples of both types of emails and identify patterns that differentiate them. This statistical-based approach enables machines to learn complex rules and handle more complicated tasks.

Machine learning is adaptable and can be applied to various problems in the field of AI. It is particularly suitable for tasks that require complex solutions or involve large amounts of data where the probability distribution of the data is unknown (Huawei, 2023).

2.4 AI in Data Analytics

2.4.1 The Evolution of AI in Data Analytics

The development of artificial intelligence (AI) in data analytics over the last few decades has been revolutionary (Murphy & Feeney, 2023). At first, the use of AI in data analytics was restricted to automating repetitive operations and supporting simple decision-making procedures. But as processing power has grown and technology has progressed, AI in data analytics has expanded to include a wide range of functions.

One of the key milestones in the evolution of AI in data analytics is the development of machine learning algorithms. These algorithms enable computers to analyze vast amounts of data, identify patterns, and make predictions or decisions without explicit programming.

Integrating AI-powered tools and platforms into current data analytics workflows is a significant step in the advancement of AI in data analytics. Professionals working with data analysis can choose from a variety of AI-based software programs that offer features like anomaly detection, generative AI, visualizations, and natural language processing.

2.4.2 The benefits of AI usage in Data Analytics

According to Murphy & Feeney (2023), AI in data analytics can elevate the work of data analysts in terms of efficiency, decision making, cost effectiveness and continuous improvement.

1. *Efficiency and Automation:* AI and machine learning algorithms automate repetitive tasks such as data entry, categorization, and analysis (Murphy & Feeney, 2023). Professionals are free to concentrate on higher-value tasks that call for human judgment and interpretation because this automation minimizes the amount of manual labor. AI-powered software, for example, can handle repetitive bookkeeping operations in accounting, freeing up accountants to give clients strategic financial advice instead.
2. *Better Decision Making:* Professionals can swiftly and precisely extract insights from massive amounts of data-thanks-to-data analytics. AI algorithms are accurate in forecasting future events by examining patterns and trends in historical data. AI systems in the legal sector, for instance, can examine statutes, case precedents, and regulations to give attorneys important information for formulating legal strategies and forecasting case outcomes.
3. *Cost reduction:* AI in data analytics can help cost reduction for companies as well as customers by automating repetitive jobs and optimizing workflows. Automation in accounting lowers the time and costs associated with conducting financial analyses and audits.
4. *Continuous Machine Learning and Improvement:* As AI algorithms process more data and get feedback, they learn and get better over time which is known as unsupervised machine learning. This iterative learning guarantees that professionals have access to the most recent insights and analysis while improving the accuracy and efficiency of AI-powered products (Murphy & Feeney, 2023).

2.4.3 The Concerns of AI Usage in Data Analytics

While AI in data analytics holds potential benefits, it also presents challenges and concerns (Murphy & Feeney, 2023).

1. *Data Security and Privacy:* Data security and privacy are two of the main issues with AI use in data analytics (Murphy & Feeney, 2023). Large-scale sensitive data analysis by AI algorithms raises the possibility of data breaches, unauthorized access, and misuse of personal data. Reducing these risks and preserving stakeholder and client trust requires implementing strong data protection protocols and adhering to data privacy laws.
2. *Fairness and Bias:* AI systems are prone to biases present in the training data, which can result in unfair treatment and skewed results. Particularly in delicate sectors like recruiting, lending, and criminal justice, biased algorithms in data analytics have the potential to reinforce already-existing inequities and discrimination. Careful data selection is necessary to address prejudice and promote fairness in AI algorithms.

3. *Overreliance on Technology*: Lastly, there is a worry regarding the possible undervaluation of human experience and judgment due to an excessive reliance on technology. While data analytics and artificial intelligence (AI) provide strong tools for problem-solving and decision-making, professionals need to find a balance between utilizing technology and maintaining the human-centric components of their work (Murphy & Feeney, 2023).

2.5 AI hallucination

AI hallucination refers to instances when results produced by AI considerably depart from reality or the intended context (Beutel et al. 2023) This phenomena presents substantial issues in the field of data analytics. Large language models (LLMs) are one type of AI model that may generate outputs that are not accurate, coherent, or relevant to the query or input data. A number of things, such as restrictions in the training data, biases embedded in the model, or mistakes in the interpretation of certain patterns in the data, might cause these hallucinations.

AI hallucination compromises the validity and dependability of analytical insights in data analytics. Data-driven suggestions are used by decision makers to guide risk management, resource allocation, and strategic decisions. AI-generated insights, however, have the potential to produce incorrect conclusions, or even destructive ones, if they are founded on false or hallucinogenic data (Beutel et al. 2023).

2.6 AI Based Tools and Technologies

2.6.1 Natural Language Processing

Natural Language Processing (NLP) means an artificial software's ability to converse, through mainly chatbots and voiceover systems, based on different inputs. According to Gowrabhathini et al. (2024), it is an AI-based tool with immense potential in data analytics. It enables operations like sentiment analysis, text summarization, and language translation.

2.6.2 Generative Pre-Trained Transformers

Generative pre-trained transformers (GPT), a more advanced type of NLP, can be effective in data analytics (Sen et al., 2023). GPT demonstrated the ability to quickly and accurately evaluate data, categorizing it into different segments with adaptability and functionality. Despite occasional errors, GPT proves itself as a useful tool for qualitative research analysis.

ChatGPT, developed by OpenAI, is a GPT available to the public (Sen et al., 2023). Capable of producing codes and patterns from data, ChatGPT utilizes its natural language processing skills to provide insightful analyses. It enhances the analysis process' dependability and transparency by quoting directly from qualitative text.

2.6.3 AI in Business Intelligence

Business Intelligence (BI) involves interpreting data to facilitate internal organizational decision-making (Edge et al. 2018). While historically focused on organized numerical data in spreadsheets and databases, the evolving data landscape now includes unstructured sources like social media and news, posing challenges for traditional BI solutions. Efforts to address the challenges posed by unstructured data streams involve integrating AI with BI systems (Edge et al. 2018).

This AI-based tool comprises two main components:

Back-end pipelines: These utilize AI to extract meaningful patterns from unstructured text, media, and metadata streams, recognizing key phrases, named entities, and different subjects.

Front-end representations: After the back-end step, AI generates visual representations based on visual analytics literature to visualize the extracted data. Dashboards and other visualizations aim to provide an easy-to-use interface for users to explore and understand different data structures, catering to a wide range of users, including those without specialist knowledge in data science (Edge et al. 2018).

2.6.4 Natural Language Processing in Business Intelligence

Integrating NLP with business intelligence (BI) systems offers several advantages, including improved decision-making, faster processes, and easier implementation of preventative actions (Gowrabathini et al., 2024). Businesses can extract insights from unstructured text data sources such as customer reviews and social media posts, with the help of chatbots, enhancing their understanding of customer attitudes and preferences. NLP also aids in automating data processing and analysis operations, maximizing the effectiveness of BI systems. For instance, sentiment analysis algorithms can automatically classify consumer comments as neutral, positive, or negative, providing valuable information about client satisfaction levels.

2.6.5 No Code AI tools

The use of text analytics, a type of no-code AI technology, to analyze unstructured data and enhance company operations and educational initiatives is examined by Liu and Downing (2024). Given that forecasts indicate a large increase in unstructured data by 2025, it emphasizes the growing significance of unstructured data analysis for enterprises. Businesses may effectively extract important insights from enormous amounts of unstructured data by using text analytics tools driven by AI (Liu and Downing, 2024). These technologies allow enterprises to fully utilize their data resources in an efficient and effective manner by enabling individuals to evaluate unstructured data without requiring considerable programming knowledge.

2.6.6 AI assistants

Rogers et al. 2023 discusses how AI assistants can assist both technical and non-technical people in businesses in an easy-to-use, collaborative environment for data analytics. It tackles the difficulty of drawing significant conclusions from data without requiring consumers to

become experts in complicated data science programming. They provide a collection of advanced data science functions that address typical data analysis tasks like data visualization, the creation of machine learning-based insights, data wrangling (data cleaning), and collaborative features. Normally, the process of executing a skill involves users creating a directed acyclic graph (DAG) of skill requests. The AI is used to complete these activities, and the user receives results based on how well they were executed. Based on user-defined inputs, the platform dynamically improves the execution of skills. Artifacts created through an AI assistant come with a summary, outlining the procedures needed to create it, which emphasizes the transparency of the platform. Based on this, users can edit and update their workflow as inputs change.

2.7 Accountability in decision making AI tools

2.7.1 Perception of Moral Responsibility of AI

The question of assigning accountability for the actions of self-governing artificial intelligence (AI) systems can be a difficult topic. The results show that there is an inconsistency in people's perceptions of the moral responsibility of AI agents (Lima et al., 2021).

In comparison to AI agents, human agents are assigned greater degrees of responsibility in terms of task performance and oversight. On the other hand, degrees of responsibility in terms of liability and blame do not significantly differ. Although people believe that humans are more accountable for tasks, both AI and human agents share the blame for any negative outcomes (Lima et al., 2021).

The study emphasizes how important it is to comprehend public opinion on blame attribution because, when deploying AI systems, stakeholder perceptions are very important. In high-stakes situations, it implies that both AI and human decision-makers are required to provide justification for their choices, highlighting the necessity for explainable AI. The findings implies that in addition to holding AI systems accountable alongside other stakeholders, users and designers should also assume accountability for the activities of AI systems (Lima et al., 2021).

2.7.2 Designing Responsible AI based systems

Sattlegger et al. (2022) explore the use of advanced self-learning algorithms to automate and standardize decision-making in public administration. They point out that while predictive data might lead to more effective decision-making, there are issues associated with the adoption of artificial intelligence (AI). They suggest that it is difficult to hold people accountable for the harm that these new socio-technical decision systems may produce. To address this challenge, the authors propose viewing it as a design challenge to design for circumstances of accountability. They create a conceptual framework to investigate the task duties of respondents in the development and use of algorithmic decision-making, based on ideas of responsibility.

The article also discusses the need to establish necessary responsibility conditions - moral agency, freedom, control, and knowledge - for individuals to take and be held accountable when developing AI. They list distinct task-related obligations and stress how crucial it is to build responsibility into designs in order to stop future harm. The authors also emphasize how critical it is to involve citizens and system users in the responsible design of algorithmic decision-making in public administration.

2.8 Summary of the Literature Review Findings

In summary, our literature review highlights the significance of data analytics and AI, but most importantly how AI affects data analytics.

The process of examining raw data to identify patterns, trends, and insights is known as data analytics, and it includes a broad range of methods and resources. The field offers a variety of techniques to extract value from data, ranging from non-numerical data analytics, which focuses on qualitative data interpretation, to numerical data analytics, which uses statistical analysis to find trends and connections. Additionally, descriptive, predictive, prescriptive, and diagnostic analytics techniques cater to different analytical needs, ranging from historical trend analysis to future prediction and decision support.

AI, on the other hand, includes the research and development of computer systems capable of simulating human intelligence. Machine learning, a subset of AI, enables computers to learn from data and improve performance over time. The evolution of AI in data analytics has been transformative, enabling automation, efficiency, and improved decision-making.

Significant advantages come from integrating AI into data analytics procedures, including increased productivity, improved decision-making, cost reduction, and continuous improvement. AI-powered tools and technologies, such as natural language processing (NLP) and generative pre-trained transformers (GPT), have improved advanced data analysis and made it easier. Furthermore, AI assistants, and other AI tools are used in business intelligence (BI) systems to glean insights from unstructured data sources, which provides important information for strategic decision-making.

Despite AI's potential advantages in data analytics, there are a number of issues and concerns that must be addressed. These include concerns related to data security and privacy, fairness and bias in AI algorithms, and the risk of AI hallucinations, where AI-generated insights deviate from reality. Furthermore, the significance of ethical issues and stakeholder involvement in AI development and deployment is highlighted. The perception of moral responsibility needs to be considered, as well as the need for designing responsible AI-based systems.

2.9 Sensemaking Theory

The selected theory for our bachelor's thesis is the sensemaking theory. Kaur et al. (2022) explains the theory as the process of identifying factors that impact individuals' capacity of understanding and interpreting complex information. First introduced in the 1960s by Karl E. Weick, Weick breaks down the theory in 7 steps (Kaur et al., 2022).

1. *Interpretation of ambiguity*: Making sense of their surroundings enables people to navigate uncertainty.
2. *Social*: To make sense of confusing circumstances, people rely on social interactions and group interpretations.
3. *Enactment of environment*: Sensemaking is an interactive process in which people use interpretation and action to actively change their environment.
4. *Retrospective*: People can better comprehend and adjust their tactics for future acts by thinking back on prior occurrences.
5. *Ongoing process*: People constantly interpret new information and improve their knowledge over time as part of the ongoing sensemaking process.
6. *Surprise and Mindfulness*: Unexpected occurrences compel individuals to seek out sensemaking in order to restore coherence, which emphasizes the significance of awareness in this process.
7. *Driven by plausibility rather than accuracy*: Sensemaking gives more weight to explanations that align with one's values and views than it does to the actual, objective reality.

Since the aim for our bachelor's thesis is to explore how data analysts perceive and experience the usefulness and ease of use of AI tools in data analytics, the sensemaking theory is best suited. The sensemaking theory enhances the understanding of how AI-driven data analytics activities are interpreted and understood by data analysts. It studies the cognitive processes required to work with discoveries produced by AI.

Thus, the literature review formed the theoretical basis of this work by presenting the main concepts of data analytics, artificial intelligence (AI) and AI in data analytics, AI-based tools and technologies, accountability in decision-making AI tools. The literature review along with the theory of sensemaking formed the theoretical framework of this bachelor's thesis, which is used to understand, analyze and discuss the research findings.

3. Methodology

3.1 Research approach

For our bachelor's thesis research, we have selected the qualitative approach, by conducting interviews with data analysts who are actively working with AI. On the contrary to quantitative research, which focuses on numerical data and statistical analysis, qualitative research facilitates a comprehensive examination of human experiences, enabling a comprehension and interpretation of intricate areas (Griffiths, et al. 2022).

This research approach aligns best with our aspiration of the bachelor's thesis. Our main objective is to answer our research question, by gaining a deeper understanding of the topic by engaging in conversations with data analysts, and learning more about their experiences, perspectives and challenges they encounter while utilizing AI in data analytics.

3.2 Method of Data Collection

3.2.1 *Semi-structured interviews*

In this bachelor's thesis study, the empirical data will be collected through qualitative methods, using semi-structured interviews. Semi-structured interviews are flexible, enabling in-depth discussion with respondents while offering a structure to ensure continuity throughout interviews with the help of an interview guide (Griffiths, et al. 2022).

In our interview guide, we included open-ended questions in order to get as comprehensive and in-depth responses as we could. If case follow up questions were needed, the respondents were asked in order to obtain more thorough responses and a deeper understanding. The interview guide with the questions can be found at Appendix A.

3.2.2 *Selection Criteria, Sample Technique, and Respondents*

Data analysts with prior experience working with AI in data analytics environments were interviewed. The respondents' experience working with AI differs. By capturing a wide range of experiences and viewpoints, this method seeks to offer deep and complex insights on the use and effects of AI in data analysis.

To conduct our bachelor's thesis research, respondents were selected working within data analytics in different companies. To be able to answer our research question, it was essential for us to ensure that the respondents had prior experience within data analytics processes and AI. To get a broad insight and perception, 6 respondents with both junior and higher positions were selected.

The *snowball effect* sampling technique *was* utilized to choose the respondents for the research. The snowball effect is a method which involves identifying respondents who meet the specific criteria, and then asking them to refer to other potential respondents (Griffiths, et al. 2022). Firstly, we reached out to people we knew that were currently working with data analytics, or they had previously worked in data analytics companies. We then asked those respondents to suggest to us other potential respondents that meet the aforementioned criteria, and who might be willing to participate in our research.

We initiated contact with possible respondents who met our selection criteria, via LinkedIn and email. We provided a brief introduction to our bachelor's thesis research and explained our interest in their specific knowledge. We then asked if they were willing to partake in our research and share their expertise and experience of data analytics and AI with us. We concluded the following 6 respondents as shown in Table 2. All of them had different roles but all worked with data analytics together with AI in their everyday work.

Table 2. Respondents Overview

Company	Role	Interview length	Date and Place	Appendix
Qlik	Technical Program Manager in AI	40 Minutes	2024-04-11 Online	Appendix C
Occtoo	Solution Architecture	35 Minutes	2024-04-12 In person	Appendix D
Company x	Junior Data Engineer	40 Minutes	2024-04-15 In person	Appendix E
Skatteverket	Chief Data Scientist	45 minutes	2024-04-16 Online	Appendix F
Qlik	Senior Data Scientist	40 minutes	2024-04-17 Online	Appendix G
BXASE	Managing Director in Data Analytics	40 minutes	2024-04-18 In person	Appendix H

3.3 Method of Data Analysis

All interviews were recorded and transcribed by the use of the transcription tool in Microsoft Teams and Tactic. The data collected through these interviews were analyzed using thematic analysis. Griffiths et al. (2022) refers to thematic analysis as the process of identifying

recurring themes or patterns in the collected data. By identifying trends and managing the large amount of data, it sets the base to develop themes (Griffiths, et al. 2022). These themes categorize specific parts of the data, and are in direct relation to our research question, with every theme presenting a specific aspect of the research question.

Our data analysis process started with filtering through the transcripts and selecting the relevant information to answer our research question. Example of this was to separate the descriptive information to explain the overall context, with the information which is correlated with the research question. Examples of keywords that were highlighted were 'data', 'analytics' and 'artificial intelligence'. These keywords together with other sentences that were meaningful to the research question consisted of our initial codes, which were then grouped into different themes. During the early stages of the categorization the themes were more general. But with iterative work we managed to put together a collection of themes that were highly correlated with our research question. In table 3 below you find the finalized themes and subthemes. The themes and their respective subthemes represent our findings and are presented and explained in the coming chapter 4 supported with the respondents' quotations.

Table 3. Finalized Themes and Subthemes

Theme	Subthemes
Data Analytics	Numerical & Non numerical
AI Tools	AI tools with specific purposes, The use of GPT & AI in Business Intelligence
AI Concerns	Challenges, Limitations & Risks
AI Benefits	
AI Accountability	

3.4 Validity and Reliability

Collecting empirical data that is both valid and reliable is essential for a successful research project (Griffiths, et al. 2022). Griffiths et al. (2022) explain validity as how accurate the results are in the conducted study, in relation to the research question. To optimize the validity in our bachelor's thesis study, we chose to utilize semi structured interviews. This interview method was chosen because it would enable us to get more detailed and in-depth information about our topic within data analysis and AI, where the respondents' experience and knowledge played a central part. Every respondent was carefully picked based on the premise that they worked closely with data analysis and AI. Therefore, we consider the respondents as relevant to contribute to our research question and demonstrate validity.

Reliability is described as the trustworthiness of the results acquired from the study (Griffiths, et al. 2022). Jacobsen (2002) explains a concept called the interview effect. Asking primarily open-ended questions is easier to conduct in person, and states that the atmosphere in which

an interview takes place has an impact on the reliability of the responses. Being in an unknown situation might cause you to react differently compared to being in a familiar setting (Griffiths, et al. 2022). As a result, we prioritized having as many as possible in-person interviews located in the respondents' office, and having online interviews as the second option, in order to optimize the reliability of our interviews.

3.5 Ethical considerations

The ethical aspect during the research is a crucial part to consider during the whole study process, from setting up the interviews to managing and analyzing the data. Griffiths et al. (2022) mention 3 key aspects one should take into account, namely informed consent, right to privacy, and requirements for correct presentation of data.

Informed consent (see Appendix B) was ensured by all respondents by voluntarily choosing to participate in the interview. This was achieved through us reaching out to the respondents informing them about the subject the interview covers, as well as giving the opportunity to create an understanding of the purpose of the study. This accumulated in a consent form being signed, where all information about the interview was included, to reassure no crucial information was missing (Griffiths, et al. 2022).

The main concern when addressing respondents' right to privacy, is that the research cannot interfere with their private life (Jacobsen, 2002). With this in mind, we constructed our interview questions to be as general as possible, while still assisting us to answer our research question. Furthermore, the respondents' identity remains anonymous, ensuring that their privacy is respected and that they feel comfortable answering the questions (Griffiths, et al. 2022).

When conducting qualitative research, it is vital to present the accumulated data in an adequate way to address the last key aspect, correct presentation of data (Griffiths, et al. 2022). To address this, all the interviews were recorded and transcribed. After the interviews, the transcriptions were offered to the respondents to assure correctness and transparency of the interview.

4. Empirical Findings

4.1 Theme 1: Data Analytics

The collected data showed that the respondents' analytics work encompassed both numerical and non-numerical data. As a result, the analytics processes were divided into either quantitative or qualitative work, in order to optimize our comprehension of the methodologies used by the respondents in their everyday work. The findings show the various methods are utilized based on the type of data, with varied analysis techniques employed for numerical versus non numerical tasks.

4.1.1 Subtheme: Numerical data

The analytics procedures used by all six responders were consistently comparable, with all of the respondents utilizing both predictive and prescriptive analytics for their numerical data. All the respondents utilized quantitative data analysis to enhance decision-making by leveraging data across various domains, by analyzing and evaluating the performance of various applications.

Respondent 1 described their numerical data analytics process as “[..]quantitative data wise, if we're measuring my chatbot performance like an AI agent, one of the big things that we look for is pretty straightforward. You know the first one's volume, you know. How are people using it? How many times is it triggered?” (R1: p.50). To evaluate further on the performance of the chatbot Respondent 1 mentions a key metric: “[..]the number one metric for the chat, is what is called containment rate and that is the percentage of times that customers talk with the chat bot and just end it there.” (R1: p.50).

Respondents 3 and 2 describe their quantitative works as “*I have worked with data analysis by taking data downloads from APIs and turning it into statistics in various analysis platforms.*” (R3: p.71). Respectively R2 “[..] when it comes to the product itself, it's a bit more quantitative, where we check performance. How well things go when we do updates, what kind of effects does it get from that?” (R2: p.63).

Respondent 4 works with collecting internal and historical company data to further develop statistical models “[..]access to historically conducted controls within a certain area where we also seek to build, develop or train some type of statistical model with an AI model” (R4: p.76).

Respondent 5 uses SQL queries to extract data from tables from Snowflake to thereafter work with analysis of the dataset “*The final sample data set, we try to do the exploratory EDA analysis. As I said, exploratory data analysis and that's the process how we*” (R5: p.87).

Respondent 6 performs their data driven performance measurement by analyzing the customers behavior “*You follow customers' behavior for a year and what you did 4-5 years*

ago was to collect all the data from various sources and then analyze it in Excel to optimize campaigns.” (R6: p.95).

In conclusion, our analysis reveals a clear distinction in the analytics processes employed by respondents when dealing with numerical versus non-numerical data. For numerical data, respondents consistently utilized predictive and prescriptive analytics to enhance decision-making, employing various quantitative analysis techniques such as statistical modeling, data extraction, and exploratory data analysis. Each respondent showcased a unique approach tailored to their specific domain, whether it be evaluating chatbot performance, product updates, statistical modeling, or customer behavior analysis. This demonstrates a comprehensive utilization of quantitative methodologies to extract actionable insights from numerical datasets, highlighting the importance of data-driven decision-making in various domains.

4.1.2 Subtheme: Non-numerical data

The gathered data showed that the main techniques for analyzing non numerical data were descriptive and diagnostic. Meaning that all of these analysis processes conducted involved identifying patterns or correlations of the datasets to gain enhanced insights.

Respondent 1 analyzes the interactions of the AI chatbot and emphasizes the value of analyzing non numerical data *“[...] that's where the money's at, I'll say. The golden egg, I'll say. Cause. Yeah, it's great to see, like more people are using it, you know, thousands of people having conversation with the chatbot. Great. What are they talking about? That's where the challenge comes in.” (R1: p.50).* Respondent 1 brings up the challenge to quantify the data in the AI chatbot since it is non numerical *“ So when we went from traditional chatbot to generative chat bot, that's one of the thing that we lost for short amount of time is if you ask me what are the top 10 questions people ask well with the traditional chatbot it's easy cuz it's just button Qliks so you can just quantify it in the data.” (R1: p.51).*

Respondent 2 works with identifying how their customers response and react to their marketing campaigns *“How well it performed, and, like, people click on it, if someone clicks on it - bought the product and then you get that information back that gives them better support to like how they're going to work in the future.” (R2: p.63).*

Respondent 3 works with the data that provides information about other data in databases *“I have worked with metadata and checked how things generally happen within a database.” (R3: p.71).*

Respondent 4 explains that much of the analytics work starts in utilizes a mix of different techniques and methods from in depth knowledge within a certain domain to identify potential errors in declarations *“A mix of different techniques and methods to identify the declaration. The actions lack with a higher risk of error is one of the methods is actually or part of the method is perhaps more experience driven.” (R4: p.77).*

To conclude, our analysis of the analytics processes pertaining to non-numerical data highlights the prevalence of descriptive and diagnostic techniques among the respondents. These methodologies aim to uncover patterns and correlations within the datasets, offering deeper insights into various domains such as AI chatbot interactions, customer responses to marketing campaigns, metadata analysis, and error identification within declarations. The

utilization of descriptive and diagnostic techniques underscores the importance of gaining nuanced insights from non-numerical data, contributing to informed decision-making and strategy development across various domains.

4.2. Theme 2: AI-based tools in Data Analytics

The analysis of the collected data reveals that all of our respondents use AI-based tools in their analysis work. Our subsections are based on the following findings: AI tools with specific purposes, the use of GPT, as well as AI in business intelligence.

4.2.1 Subtheme: AI tools with specific purposes

We found that 5 of our 6 respondents used a specific purpose AI-based tool designed to assist a specific workflow in their analytical work. Respondent 4 works at a government agency, Respondent 1 and Respondent 5 work at a worldwide enterprise, whereas Respondent 2 and Respondent 6 work at small to medium companies, yet they all use a specific purpose of self-developed AI-based tool for their analytical work.

Respondent 4 explained that he works with *“Various applications of AI in what is now referred to as AI.”* and they are developed to fit their specific needs *“We primarily develop a lot ourselves connected to what lies within what is risk assessments.”* (R4: p.77). In certain cases, Respondent 4’s organization would develop their AI based tools from scratch, and in other scenarios they start by implementing open-source code which they build upon. For example, they have an AI tool which *“classifies incoming emails to the Swedish Tax Agency, meaning we attempt to determine the subject and allocate it to the appropriate expertise in order to respond to the questions we receive,”* (R4: p.78), underlining the one-dimensional, specific aspect of the tool.

R5 uses a machine learning apps platform called databricks, where all of their code is ingested and synchronized with their GitHub repositories. Once you start writing a cell, there is an AI assistant embedded within their notebook software itself, that generates code when you start typing. *“Once you give a question like a pseudo code or something like that, it automatically generates the code for you.”* (R5: p.89). This means that they would only get code suggestions based on code written for data analytics.

Respondent 1 works at the same company as Respondent 5, but in a different department, and is responsible for developing a chatbot *“I own and operate and maintain our AI chat bot. So, for example I deployed it, managed it, make improvements on it and also did reporting on it.”* (R1: p.49). His analytical work comes in terms of analyzing the chatbot, meaning that the AI based analytical tool is only developed to cater to the needs of the chatbot. Respondent 1 analyzes the chatbot's interactions with the user, which helps him understand the volume and usage of the chatbot, but also what the users mainly discuss, *“[we] utilize GPT 4 model to basically read through the conversation and the GPT will give you a summarization.”* (R1: p.52).

Respondent 6 is a self-employed entrepreneur who has worked with developers who constructed an AI based tool from scratch for the specific purpose of analyzing price elasticity, *“The process of trying to find the optimal pricing is called price elasticity, and one*

way to test and evaluate it historically is by using machine learning to analyze customer data related to pricing." (R6: p.108).

Lastly, Respondent 2 uses a generative AI that analyzes and generates text based on the customer who is visiting the page. The text would for example look different if an older person wanted to buy a phone online compared to a younger person, *"Being able to develop texts and messages specifically targeted towards you, the age group you belong to, and your level of technical expertise."* (R2: p.65). The tool is designed for their customers and their respective segment.

In conclusion, our analysis highlights a significant trend among respondents, with five out of six utilizing specific-purpose AI-based tools tailored to their analytical workflows. Despite varying organizational contexts, from government agencies to small to medium-sized enterprises, respondents rely on self-developed or customized AI-based tools to enhance their analytical processes. These tools are designed to address specific needs within their domains, such as email classification, chatbot analysis, customer segmentation, and price elasticity analysis, emphasizing the importance of customization and innovation in leveraging AI technologies for efficient decision-making.

4.2.2 Subtheme: The use of GPT

The analysis of our collected data suggests that Generative Pre-trained Transformers (GPT) can work well as an analytical tool, even though the tool itself may have multiple other purposes.

Respondent 1 explained how they underwent a transformation from an algorithmic code-based chat bot to a GPT model, which hindered them from analyzing and evaluating the overall performance of the chatbot due to lack of transparency. By solving this, they implemented AI features that helped analyze certain keywords, making them quantifiable. *"Now you can, you know, kind of quantify it, OK? Like the tag is password reset. That's and then you can quantify it to see if that's one of the top questions you're being asked"* (R1: p.52). Then after analyzing the volume of the keyword "password reset" the GPT can analyze whether the different password resets were successful amongst their users or not *"It's going to label it as the issue was resolved or issue couldn't be resolved and why [it could not be resolved]"* (R1: p.52).

Respondent 4 revealed that they are exploring the possibility of implementing GPT in their institution but are not using it to date. *"GPT or GPT-4, if possible, for us to use as a Swedish authority, I would say we are in an exploratory stage right now."* (R4: p.78).

Respondent 3 and 5 specifically used ChatGPT, a GPT developed by the company OpenAI, for their work. They used it, not for the analysis itself, but rather to facilitate their analytic work by generating SQL queries and code snippets. Respondent 3: *"I use a lot of ChatGPT to simplify data analysis for me in general. It's mainly about checkability."* (R3: p.72), and Respondent 5: *"It can easily format it for me and with the right indentation and all of that" [...]* *"I define a use case for it, it will actually build that entire step for me. Like if I say hey, build a step-to-step process to build a binary classification model"* (R5: p.89). They all emphasized that they wouldn't use real data in ChatGPT under any circumstances. Respondent 3 claimed that GPT was not a collaborator with his company, but rather

something certain employees would use based on personal preferences. Respondent 5 also used anthropic cloud AI as well, for the same purpose as ChatGPT.

In conclusion, our data analysis suggests that Generative Pre-trained Transformers (GPT) show promise as analytical tools. Respondents integrated GPT models, like ChatGPT, to streamline workflows, though some are still exploring implementation. While GPTs offer automation and assistance, their use is cautious, prioritizing data privacy and security.

4.2.3 Subtheme: AI in Business Intelligence

All respondents had experience using BI tools. We asked each of them: Have you ever used business intelligence tools such as Power BI? Would you want BI to be more integrated with AI?

In response to the first question, they provided similar answers regarding the specific BI tools they've used. Respondents 1, 5, and 6 used Tableau for visualization. Respondents 1, 3, 4, and 5 have utilized Qlik and QlikSense. Respondent 3 was the only respondent primarily using Power BI. In response to the second question, the respondents shared different perspectives based on their experience and how they perceive further AI implementation in BI tools.

Respondent 2 emphasized the benefits of precision and higher rate in AI compared to BI's current state: *"I see advantages with AI compared to BI. BI runs slowly. It has nightly runs, and you have to know exactly what you need, so it's very precise, whereas with AI, it's faster."* (R2: p.66). However, Respondent 2 explains further that AI would work as an addition, and not a substitute for the current human-computer interaction in BI systems: *"Then, the downside as I see it is that... To extract something, like 'give me a campaign' based on those [customers] with the lowest sales, so we can have a discount campaign with them, you know. It's very precise. If you ask an AI the same thing, you don't know exactly what you'll get, so that's where the difference lies."* (R2: p.67). The insinuation is that existing AI cannot replace that part of BI.

Respondent 1 touched upon the same subject as Respondent 2, but through a more futuristic lens. Respondent 1 explained that companies are trying to implement an analytically based AI, *"Tableau, Salesforce, even power BI. They're all trying to add it in there, even Microsoft trying to add in, like, probably copilot into power BI and it's very soon if not already."* (R1: p.55). Similar to Respondent 2's view, Respondent 1 explains that the AI would make assumptions based on the input of the user. For example: *"Hey, what's my sales like in the last two quarters? What's the increase and decrease? It was like, sure, your sales have decreased in 5% in the Americas region."* (R1: p.55).

Lastly, Respondent 6 claimed that today's BI could become more intuitive, where he thinks that AI could play a role in decreasing the workload of the analyst: *"For example, having the data in front of you and simply saying: 'I want this data presented in a timeline, with this on the x-axis and the impact from rewards on the y-axis, and compare multiple years consecutively.'" (R6: p.100). Respondent 6 implies that this will help the analyst in his process, whereas Respondent 1 implies that these types of AI advancements may replace human analysts. Respondent 1: "[with this implementation] you wouldn't need to hire our own data analyst or business intelligence employee" [...] It's not gonna be a person's opinion on the data, but more of large language models, you know, assessment of the data and telling you what it thinks is gonna be."* (R1: p.54).

In summary, respondents commonly use BI tools like Tableau, Qlik, QlikSense, and Power BI. Regarding AI integration, opinions varied. Some emphasized AI's speed and precision over traditional BI processes, while others expressed concerns about AI's ability to replicate human-driven queries accurately. Insights into future integration envisioned AI enhancing intuitiveness and reducing analyst workload. Overall, while some see AI as complementary to human analysis, others suggest it may eventually replace certain analytical functions, highlighting the need for careful consideration of AI's role alongside human expertise in BI.

4.3 Theme 3: AI benefits

We asked our respondents “based on experience, what are the benefits you see with the use of AI in data analysis?” This allowed for a variety of responses without any specific direction. All the respondents could highlight at least one specific benefit but alluded to different perspectives. Some were more through the lens of their organization, and some more general.

Respondent 1 mentions the advantages he’s observed through the expanded implementation of AI in his chatbot. The previous chatbot had the following issue: *“We just assumed that if at the end of the answer they didn't ask for a live agent help, we just assumed that the issue was resolved,”* (R1: p.52). Now, the AI contextualizes the interaction, providing insights in why the user’s errand wasn’t resolved: *“Now you have this AI intelligence assessing, reading and understanding the context of those conversations because it is trained on our products.”* (R1: p.52).

Respondent 4 also gives insights into the benefits of his specific work: *“Work with various types of co-pilots to assist me in writing code better and more efficiently.”, but also “Transform unstructured data into a more structured format, allowing me to tackle the problem more effectively or more quickly.”* (R4: p.80). Respondent 4 continues, explaining that in the past, they spent money translating over 1000 pages across various documents, only to find that only 2 or 3 pages were relevant. By implementing AI, they solved this problem: *“Through machine translations, we actually save time and become more relevant in the materials we then forward for authorized translation.”* (R4: p.81).

Other respondents stated more general benefits in how AI has helped them to date. Respondent 2 emphasized the higher rate and flexibility of AI in data analytics, especially in working with large datasets: *“With AI, things move faster, and you can be much more flexible. That adaptability is what I see as crucial when dealing with large volumes of data. It's quite easy to produce results and conduct tests, so it's much more lightweight.”* (R2: p.66). Respondent 5, similar to Respondent 2, didn’t point towards a specific workflow but rather emphasized the benefits in general. However, Respondent 5’s benefits were less in line with high rates, but rather accuracy and minimizing human errors: *“It will reduce the risk of human error or the bias in data analysis. So, it's more accurate and it's more consistent, I feel and. Obviously, improvement of predictive capabilities.”* (R5: p.91).

Lastly, Respondent 3 and 6 combined their experience of AI in data analytics with insights in the capabilities that future AI may hold. Respondent 3 pointed out the benefit of AI’s ability to structure data: *“I believe that AI can, in a very effective manner, in the future, categorize big data and structure it very well so that it can be analyzed initially.”* (R3: p.73). Respondent 6 specifically pointed out the benefits in speed and simplicity that comes along with AI: *“The two main advantages are speed and simplicity. [...] One can manually perform this type of analysis, one by one. And, it becomes an enormous immediate task if one were to depict*

reality. It's like finding a needle in a haystack. But with AI now, so to speak, one can input these parameters and quickly obtain an analysis of the data." (R6: p.101).

However, both Respondent 3 and 6 stated that the benefits with AI are assisting the human analyst, rather than replacing human intelligence. R3 regarding AI's analytical capability: *"The analysis part, I'm not sure if it [the AI] will fix it as well [as a human]"* Respondent 6 specifies that AI is meant to save time for the analyst, which frees up time, not replace the human: *"If one can reduce the time spent on cutting and pasting and instead focus more on thinking and analyzing, it becomes a significant gain. Then one can free up much more time to truly think and contemplate."* (R6: p.101).

In conclusion, our respondents highlighted various benefits of using AI in data analysis, ranging from specific improvements within their workflows to broader advantages such as speed, accuracy, and efficiency. Some emphasized AI's ability to contextualize interactions, transform unstructured data, and enhance productivity by assisting in coding tasks. Others emphasized AI's speed, flexibility, and predictive capabilities, as well as its potential to structure big data effectively for analysis. While acknowledging AI's potential, respondents also emphasized the importance of AI as a tool to assist human analysts rather than replace them, underscoring its role in saving time and enhancing analytical capabilities. Overall, the insights provided by respondents demonstrate the multifaceted benefits of integrating AI into data analysis processes.

4.4 Theme 4: AI Concerns

The analysis of the gathered data regarding AI concerns revealed three subthemes that surfaced as primary concerns and topics of conversation among respondents, which were data challenges, limitations, and risks associated with working with AI.

4.4.1 Subtheme: Data challenges & Risks

Based on what the respondents said, the main challenge for data analysts working with AI is the data. Even though the respondents pointed at the same problem, they come from different perspectives.

Respondent 1 explains the core problem for all data analysts lies in the quality of data *"I think the biggest challenge is the same challenge that all data analysts face is the quality of the data."* (R1: p.56). Why working with good data is so crucial is because *"[.]Bot is only as good as the training from the data."* (R1: p.56). Therefore, a large emphasis went in to assuring good data quality of the AI chatbot before launch.

Respondents 3, 5 and 6 share all the same challenges handling the data in AI tools. Respondent 3 explains that as a professional working within a company, handling certain data with AI tools is illegal *"Right now you cannot use your own data, company data, customer data in general, in AI tools that are not privatized"* (R3: p.73). Respondent 3 explains further that perhaps illegal usage will not have an instant impact but can have devastating effects in the long term, *"Maybe no one will see it for a little longer, but if there is a breach, it can be very bad."* (R3: p.73).

Respondent 5 shares the same concern when handling data with AI tools even though they don't have experienced it directly, *"I haven't seen it yet personally, but I feel like it could pose problems. Maybe for privacy and our data rights that we have"* (R5: p.92).

Respondent 5 discusses further about potential risks of data misuse *"It could hurt our reputation like the image that Qlik has in the market right now. Qlik has a really good name in the market. In the Gartner chart and all of that, right. So, I think that the reputation will be damaged because of this and also the ethical concerns that we have."* (R5: p.93).

Respondent 1 explains certain industries where AI tools are not allowed because of the strict security rules *"For example, our healthcare industry is where big concern with that is like. We don't want any AI features because we're worried this is public health data of our customers and all like patients that has one of the strictest rules on security."* (R1: p.57).

To prevent misuse of data Respondent 6 explains that they delete sensitive information and only keep the relevant data when they use customer data within their own developed AI models. R6 explains that even though security measures are implemented, they still cannot guarantee complete secure data. *"We remove sensitive information such as name and address and only keep the data we need for analyses. But there is always a risk that data can be identified, depending on how much data you have access to"* (R6: p.101).

In conclusion, the primary concern voiced by respondents regarding AI in data analytics revolves around the quality and legality of the data being used. Across various perspectives, the focus remains on ensuring data integrity and privacy. Challenges include ensuring data quality for effective AI training, navigating legal restrictions on data usage, and mitigating risks of data misuse, especially in sensitive industries like healthcare. While measures such as data anonymization are employed to enhance security, complete data security cannot be guaranteed. Overall, these concerns highlight the importance of robust data governance and ethical considerations in AI-driven data analysis practices.

4.4.2 Subtheme: Limitations

AI can identify trends, but not understand them to the extent human intelligence would. Even though there are lots of benefits with using AI as a tool, there is still a need to monitor the results generated from AI. Respondents 2, 3 and 6 all emphasize the constant need to always oversee, and quality assure the generative AI content. R2 explains it as follows, *"The challenges are, it's precisely this matter of knowing if it's right, so if I get a suggestion, it's if we look at the generative AI that we've had, it generates texts, sort of. I can go in and verify and do random tests, but you don't know that everything will be fine, like, and it's the same if you get a result"* (R2: p.67). Respondent 6 emphasizes that the incorrectness of the AI is a matter of probability *"AI can't be 100% safe, but to claim it is wrong. It is a matter of probability, and it may be possible to provide a guideline for the probability based on the information available and the analysis performed. One should be aware of how the system works. It is not about absolute truths but about probabilities"* (R6: p.103).

Respondent 4 shares also the sentiment of needing to evaluate AI content, but also brings up the need of putting the AI content in context of things, *"[...] you always need to evaluate or test what you get from an AI system to put it in context [...]"* (R4: p.81). Respondent 6 shares the same sentiment and emphasizes the importance of analyzing the data within the correct context, *"If I see that a lot of red pants were sold last year and much less this year, I have to*

think about why. I don't think the data in itself is very helpful there, it's more about general trends or psychological factors in people that influence. What it is due to is usually unclear, and I think there is still a long way from being able to use AI and understand such things” (R6: p.100).

In conclusion, while AI offers valuable insights and trends, it falls short of human comprehension. Respondents underscore the necessity of continual oversight and quality assurance in interpreting AI-generated content. Acknowledging the inherent limitations, particularly in accuracy and contextual understanding, is crucial for effective utilization of AI-based tools.

4.4.3 Subtheme: AI hallucination

AI hallucination is a common event and can lead to negative impacts in several different aspects. Respondent 1 described AI hallucination as their main concern when deploying their AI solution *“That was our biggest concern when we were trying to launch AI assistant is how do we prevent hallucinations from happening?”* (R1: p.58).

Respondent 2 compared the AI confabulation to how certain people act *“Some men who would never admit that they didn't know something, like making up an answer, but a normal person would say, like, no, but what I don't know, I'll have to look it up and get back to. Like between an AI could make something up and if you don't have source criticism, it could be that you take it for as if it were the truth, so that's also how you know you're interacting with an AI.”* (R2: p.69).

Respondent 3 explains that the whole point with data analytics is that the data is correct, if not then there is no point *“Yes, but the whole thing with data analysis is that the data must be correct [...] If it doesn't, then it's no good. It is just as well to throw away the reports.”* (R3: p.74). Meaning, using AI to produce data for analysis would not be a wise move if the accuracy of the data it generates, cannot be guaranteed.

To conclude, AI hallucination poses significant risks across various domains. Concerns voiced by respondents highlight the importance of addressing this issue to ensure the reliability and integrity of AI-driven solutions. Without verifiable accuracy in generated data, the efficacy of AI in analytics is compromised, emphasizing the need for robust quality control measures.

4.5 Theme 5: Accountability

For our last question, we presented the respondents with a hypothetical scenario regarding accountability in regard to faulty AI. The scenario was the following: A vendor provides a consultant with an analytical AI tool. The consultant uses the tool to help their customer with an analysis. The customer invests money based on the consultant's analysis. However, the analysis was wrong due to a mistake made by the AI tool which made the customer lose money. Who is to blame in this scenario? The vendor, the consultant, or the customer? The emerged finding was that our respondents had different views and insights on this matter all around.

Respondent 1 made it clear that the responsibility should be shared across all 3 parties. You cannot blame the developers 100% since that would be harmful to the AI industry as a whole and inhibit development: *“If you make the developers liable 100% all the time, that risk is*

gonna get so big that it's gonna start to actually shrink the AI industry.” (R1: p.60). Respondent 2 made a reference to self-driving cars and accountability regarding AI in other industries containing AI: “That's one of the reasons they haven't released them (self-driving cars), because the issue of liability cannot be resolved, you know.”. Furthermore, R2 touched upon AI as a whole implying that those who use AI should know better than relying on it fully. “AI is uncertain, so I think that's also a reason why, in some cases, one doesn't want to rely on it entirely, one needs to verify it.” (R2: p.69).

Respondent 4 and 5 were shifting their focus more towards themselves and their overall responsibility as an organization. Respondent 4 explained that, as a Swedish authority, they need to be held accountable for the actual responses they leave to citizens and other governmental employees. Respondent 4: *“So in that way, we own that question ourselves, and we will never be able to point to anyone else.” (R4: p.82). Respondent 5 had a similar view, but more in line with persevering the brand's quality of her organization: “There could be a problem like an erosion of trust, the loss of trust. I mean once the trust is lost, you can never get it back, right?”. Respondent 5 mainly placed the responsibility on themselves as consultants as well as the vendor: “I think it should be mostly us and the vendor. Us and the vendor should be taking the responsibility, yeah.” (R5: p.93).*

Furthermore, quality assurance was brought up by both. Respondent 4 pointed to continuous quality assurance: *“There are a variety of technologies and methods to limit those harmful effects, so to speak, or limit the system's ability to respond. In some cases, one can work with various testing techniques to avoid the model responding even though it doesn't have an answer.”, and Respondent 5 emphasized the quality assurance work that needs to be done prior to implementing an AI vendor's product: “You know, go ahead with the tool if you like it or not. If it passes all your quality checks or not.” (R5: p.86)*

Respondents 3 and 6 explained that there was no easy answer to the hypothetical scenario. It was rather a matter of what's been discussed and communicated between the different parties. Respondent 3 asked a counter question: *“Was it our (consultant's) initiative to use AI, or was the client also involved in it?” (R3: p.75), we replied “it was your initiative”. Respondent 3 stated that the consultant should take the most part of the responsibility in that case. Respondent 3: “Yeah, but ultimately, regardless, it's our fault, because as a consultant and as an analyst and data engineer, you have to be on top of things, you have to take responsibility. Your competence is your asset, the owners' capital, and your resource.” (R3: p.75). Lastly, Respondent 6 claims that it essentially all comes down to what's communicated. If the customer is aware that the analysis is AI based, then they should take most of the responsibility given that AI can only be a probability, not a set fact. If, however, the consultant used AI without the customer's knowledge or consent, it is fraud: “If the customer is aware that the analysis is based on AI and still proceeds with the decision, they have only themselves to blame if things go wrong.” (R6: p.104).*

In conclusion, the issue of accountability surrounding faulty AI presents a complex landscape with varied perspectives. While some respondents advocate for shared responsibility among developers, consultants, and customers, others emphasize organizational accountability and the importance of quality assurance measures. Ultimately, the allocation of blame hinges on clear communication, informed consent, and the proactive management of risks inherent in AI utilization.

5. Discussion

In this chapter, we discuss our research findings in relation to previous literature and the selected theory of sensemaking to extract insights. Through the application of the theoretical framework to our findings, we reply to the posed research question, and we achieve a better understanding of the perceptions and experiences of data analysts regarding the use of AI-based tools and technologies to their everyday work. The following Table 4 encapsulates our research findings and delineates the themes that address our research question.

Table 4: Findings Addressing the Research Question

Research Question	Findings
How do data analysts perceive and experience the usage and impacts of AI in data analytics?	Theme 1: Data Analytics Subtheme 1.1: Numerical Data Subtheme 1.2: Non-Numerical Data
	Theme 2: AI-Based Tools in Data Analytics Subtheme 2.1: AI Tools with Specific Purposes Subtheme 2.2: The Use of GPT Subtheme 2.3: AI in Business Intelligence
	Theme 3: AI Benefits
	Theme 4: AI Concerns Subtheme 4.1: Data Challenges and Risks Subtheme 4.2: Limitations Subtheme 4.3: AI Hallucination
	Theme 5: Accountability

The bachelor's thesis findings showed that all participating data analysts recognized the benefits of integrating AI into data analytics. Some highlighted how AI can reduce errors and bias compared to human analysts, while others highlighted its cost-saving potential. However, the strongest correlation between them in terms of benefits was the efficiency and time saving aspect. The analysis needs to be done whether or not AI is used, and while product improvement may happen even in the absence of AI, its usage speeds up the process considerably. While our findings showed that the data analysts were proficient in SQL queries and Python coding, AI performs these tasks faster, allowing them to allocate time to areas where AI struggles, the ability to see the whole picture. Our bachelor's thesis suggests that AI

can either replace specific workflows or serve as a complementary assistant. However, it became clear that AI cannot completely replace human analysts altogether in its current state.

Regarding specific workflows, the findings showed that the majority of data analysts utilize software designed to either replace or assist specific processes. For example, one data analyst used a self-developed translator tailored to their workplace and used about ten different AI-based tools for their analytical job. They decided against utilizing Google Translate because of privacy issues, which was one common point made by respondents when using open AI for real-world data analysis. Usage of multi-purpose AI software like ChatGPT or Google Translate immediately compromises data integrity.

Further findings showed that the risk of using AI multi-purpose software that analyzes the bigger picture, it can be less transparent and harder for humans to detect where the fault lies, when it inevitably occurs. By having software that analyzes a specific workflow, you increase your control of the analyzing process. Therefore, creating software specifically designed for particular workflows not only improves control and transparency but also reduces the dangers related to multi-purpose AI tools. Our findings showed that the data analysts take advantage of machine learning benefits by using self-developed, specialized AI-based solutions to train software specifically for their data analysis requirements. Overall, they used a wide range of tools, including code generators, BI, GPT models, AI helpers, and visualization tools.

Where there was consensus in regard to the beneficial aspect of AI, the risks, challenges, and limitations were more of a broadly discussed subject. Along with limitations in AI's current capabilities, concerns were expressed about data breaches and integrity. The findings also showed that the data analysts acknowledged the need for further education to remain competitive in the face of AI advancements. Others, however, did imply that the work of data analytics is harder to be replaced by AI, and were more optimistic overall. The findings highlighted AI hallucination and how important it is to implement quality assurance procedures to prevent it. In addition, findings repeatedly showed a concern of accountability; that is, data analysts need to take their own responsibility, and need to communicate with various stakeholders in the work of implementing and optimizing artificial intelligence. AI can make mistakes, just like humans.

Our findings showed that AI in data analytics is in a transitional phase, with data analysts still heavily relying on manual workflows for analyzing numerical and non-numerical data. While none of them use AI extensively in their overall analytical work, it is utilized to a greater extent in specific workflows. The general attitude was cautiously enthusiastic, weighing the potential of AI against its risks and challenges. While some believe AI will eventually replace human analysts, others are still reluctant in putting their faith in the AI's ability to see the whole picture. The findings are further discussed in the following section with the help of the theoretical framework.

5.1 Data Analytics

Investigating how data analysts conduct their data analysis processes provides an essential foundation for addressing our research question on how data analysts perceive and experience the usage and impacts of AI in data analytics. By getting a more in depth understanding of the various methodologies and tools in data analytics, we gained a comprehensive understanding

of the data analytics landscape. With this knowledge we could dig into potential opportunities, challenges and implications associated with implementing AI in these processes.

The findings from our bachelor's thesis research shed light on the various approaches used by data analysts to analyze numerical and non-numerical data, displaying a range of different methods and tools that they apply during their analytical processes. The steps described in the literature review including data discovery, storage and preparation, modeling, evaluation, and communication, provide a structured framework that captures key components of data analytics. These steps serve as guiding principles for analysts when reviewing and analyzing datasets in search of insightful conclusions.

5.1.1 Numerical Data Analytics

Numerical data analytics, as described by Kotronoulas et al. (2023), is the process of analyzing quantitative data using a variety of statistical techniques. This is supported by our findings, which show that data analysts use numerical data to analyze trends, assess performance, and make insightful conclusions. By utilizing predictive analytics and prescriptive analytics, analysts can extract valuable insights from numerical datasets. For example, in order to extract insights from data, respondents reported procedures including statistical techniques, SQL queries, and performance evaluation indicators.

Kotronoulas et al. (2023) and our findings emphasize the importance of quantitative data analytics when dealing with large data sets. Without the use of statistical methods and techniques the possibility of speculation runs high, ultimately resulting in non-trustworthy results and conclusions. This said, quantitative data analytics plays a crucial role in research, bridging the gap between raw numerical data and useful insights.

5.1.2 Non-numerical Data Analytics

Whereas quantitative data analytics focuses on numerical data, qualitative data analytics analyzes and interprets non-numerical data in a methodical manner to identify themes, patterns, and interpretations (Sandelowski, 2000). This aspect of data analytics involves using descriptive and diagnostic analytics to find subtle insights that may be overlooked by using only quantitative methods. Our findings highlight the value of qualitative data analytics in understanding customer behavior, improving products, and optimizing marketing strategies. The findings also showed the challenges associated with quantifying non-numerical data, particularly in the context of AI chatbots where interactions are diverse and complex.

Our findings highlight the distinctions and similarities between numerical and non-numerical data analytics, showing that although the methods vary, the overall goal of enhancing data remains the same. Non-numerical analytics involves identifying patterns and trends in order to gain insight, numerical data analytics primarily deals with quantitative measures and performance indicators. Respondents described various procedures and methodologies for each kind of analysis, demonstrating the variety of data analytics approaches.

5.2 AI-based tools in Data Analytics

5.2.1 Tailoring easy to use AI to fit specific analytical needs

Our findings show that respondents from a range of organizations emphasized the importance of using AI technologies that are tailored to fit particular workflows in their fields. For instance, Respondent 1, who is in charge of monitoring and evaluating a chatbot, emphasized the significance of AI tools customized for the chatbot's users. This observation is consistent with Huawei's (2023) talk about machine learning techniques, which let computers learn from data and automate processes that are particular to analytical needs. Machine learning should be delimited correctly.

Similarly, Respondent 5 discussed their utilization of a machine learning apps platform explicitly designed for data analytics tasks. An embedded AI assistant in this platform creates code that is tailored to their analytical requirements. In their elaboration on the function of AI assistants, Rogers et al. (2023) highlight how these tools can be used to provide advanced data science capabilities that are customized to common data analysis activities, which can improve collaboration and productivity.

Respondent 4 described the development of AI tools tailored to risk assessments within their organization, emphasizing a customized strategy for dealing with particular analytical difficulties. This is similar to Liu and Downing's (2024) explanation of how no-code AI tools are becoming more and more important for evaluating unstructured data, allowing users to derive insights even without a data science background. Respondent's 4 organization's approach to filtering out unnecessary documents or emails was described as one of the first machine learning tools (Huawei, 2023), yet still used today, suggesting that AI in analytical work can be both complex and less complex.

Respondent 6 utilized his expertise in price elasticity to set up the development of a specific AI tool assisting customers in determining product pricing on a continuous basis. Respondent 6 does not have a data scientist background, nor do his customers in general, underlining the fact that AI assistants can be used widely. Respondent 6 also claimed that excel today is in a state where the input part of the analysis takes too long in contrast, Respondent's 6 AI-based tool, though based on advanced machine learning and AI, was deemed easier to use than Excel, aligning with the findings of Liu and Downing (2024) and Rogers et al. (2023).

5.2.2 Using GPT in Data Analytics

The integration of Generative Pre-trained Transformers (GPT) into data analytics has emerged as a promising field, based on our findings. Despite having other purposes, GPT models have been shown to be useful analytical tools in a variety of situations. For example, R1 explained how to convert an algorithmic code-based chatbot to a GPT model, demonstrating how the model may help with performance analysis and evaluation. GPT improved comprehension of user wants and replies by enabling measurable insights into user interactions through the implementation of AI elements like keyword analysis. Additional research backs up the usefulness of AI-driven data analysis tools like GPT. Sen et al. (2023) explored the utility of

computer-generated (CG) qualitative data analysis, demonstrating CG's ability to efficiently categorize and interpret data across multiple categories. The study highlighted CG's adaptability, multilingual capabilities, and quick, precise analysis, positioning it as an asset for qualitative research analysis. Gowrabhathini et al. (2024) emphasized the benefits of integrating NLP techniques within BI systems, similar to the above.

ChatGPT is a type of GPT with many purposes. The flexibility with which ChatGPT may generate codes, groups, and themes from data highlights the promise of this tool for general data analytics (Sen et al., 2023). The model's natural language processing skills enable it to read and evaluate data effectively, providing insightful analyses to support diverse analytical goals. ChatGPT's ability to quote directly from qualitative text also enhances the dependability and transparency of the analysis process,

Our empirical findings found that two of our respondents, Respondents 3 and 5, specifically use ChatGPT in their everyday work. Respondent 3 also stated that many at his company use it as a personal preference in their analytical work. Respondent 3 and Respondent 5 discussed ChatGPT's ability to generate SQL queries and code snippets to speed up their data analysis procedures. ChatGPT's natural language processing skills were useful in streamlining general data interpretation as well as the creation of analytical components (code, models). ChatGPT was not used for analysis directly, but rather as an assistant in the analytical work.

Lastly, continuous machine learning and improvement in AI algorithms contribute to the ongoing enhancement of AI-powered products as supported by Murphy & Feeney (2023). Respondent 1's company is using a GPT suited for their needs, and not ChatGPT, which makes the training of the software secluded to their customers alone. This iterative learning process enhances the accuracy and efficiency of AI-driven data analytics tools.

5.2.3 Business Intelligence (BI) can benefit from AI advancements

For businesses looking to use data-driven insights for strategic decision-making, business intelligence (BI) is a vital component. Our investigation into the relationship between artificial intelligence (AI) and business intelligence (BI) produced new findings about the possible advantages of AI advancements for enhancing BI capabilities.

Our results demonstrate how widely respondents used BI tools, underscoring the crucial function these platforms play in data analytics whether you work in a corporation or authority. The respondents overall demonstrated a strong desire to use AI to improve BI features.

Research by Edge, Larson, and White (2018) aligns with our empirical findings in the literature, highlighting the difficulties brought about by the abundance of unstructured data sources and the requirement for AI-enabled BI solutions to manage this complexity. In their support of AI integration into BI systems, Edge et al. (2018) point to the technology's ability to draw insights from unstructured data streams and enable data-driven decision-making.

In line with Gowrabhathini et al. (2024) findings, our findings emphasized AI's potential to increase BI procedures' accuracy and efficiency. Gowrabhathini et al. (2024) underscore the advantages of integrating Natural Language Processing (NLP) techniques with BI systems, enabling automated data processing and analysis operations. Respondents 2 and 6 both claimed that they have yet to see this but would like to in the future. Respondent 2 visualizes

an AI driven BI tool that can provide sales campaigns to customers based on the data provided. Respondent 6 saw NLP in BI as a way of saving time for the analyst. Respondent 6 suggests that the analyst's role is to analyze the bigger picture, whereas AI's role is to provide the bigger picture. Gowrabhathini et al. (2024) findings align with our findings, which show respondents' optimism about AI-driven BI systems as intuitive analytics tools capable of reducing analyst workload and streamlining decision-making processes.

Lastly, research by Liu and Downing (2024) reflects on how no-code AI technology, particularly in text analytics, has great potential too. Respondents' excitement for AI-driven BI solutions that streamline data analysis activities without requiring a deep understanding of programming is reflected in this. Respondent 1 discussed how his company uses GPT to extract key phrases in their non numerical analysis. As a data scientist, respondent 1 has the tools to implement such software. However, Liu and Downing (2024) suggest that this type of technology can be implemented as no-code into existing BI tools. These technologies allow enterprises to fully utilize their data resources in an efficient and effective manner by enabling individuals to evaluate unstructured data without requiring considerable programming knowledge as said by Liu and Downing (2024).

5.3 AI Benefits

5.3.1. Enhanced Efficiency and Time Savings

Professionals benefit greatly from increased efficiency and time savings by integrating AI into data analytics as supported by Murphy & Feeney (2023). Our findings emphasized how AI might revolutionize operations by automating tedious jobs and optimizing workflows. For example, R1 mentioned how AI contextualizes conversations in chatbots, resulting in more effective problem solving. Furthermore, Respondent 4 demonstrated how AI makes it possible to filter out a lot of unnecessary data into a structured manner, which speeds up the analytical process. These findings align with previous research emphasizing AI's role in automating tasks like data entry, categorization, and analysis by Martinez (2019).

5.3.2 Enhanced Precision and Reliability

The increased consistency and accuracy of insights obtained from large datasets is another important advantage of artificial intelligence in data analytics as found by Murphy & Feeney (2023). Our findings emphasized how AI may reduce biases and errors made by humans, leading to more accurate analyses. The findings of our bachelor's thesis study highlighted AI's ability to reduce human errors and biases, resulting in more reliable analyses. In contrast, Murphy & Feeney (2023) considered bias and fairness in AI to be a concern. Since AI receives its training based on data from the real world, human intelligence and artificial intelligence seem to be related in terms of bias. If the data is biased, then the AI software would as well, suggesting that there is no easy way around bias (Murphy & Feeney, 2023).

Respondent 2 points out that AI algorithms are excellent at predicting the future by identifying patterns and trends in past data. These findings support earlier study, which highlights AI's contribution to improving decision-making accuracy through data analytics as found by Martinez (2019).

5.3.3 Complementary Role of AI in Elevating Human Analysts

Martinez (2019) states that there is a difference between human intelligence and artificial intelligence. AI simulates human behavior and intelligence and is therefore hard to define and compare to human intelligence. According to respondent 6, AI has great capabilities in exceeding human intelligence in terms of speed and structure, whereas other parts in human intelligence are harder to replicate. The findings of our bachelor's thesis study emphasized that AI serves as a supportive tool for human analysts rather than a replacement R3 and Respondent 6 demonstrated how AI can efficiently organize and classify large amounts of data, supporting human analysts in their work. Furthermore, as stated by Respondent 6, AI frees up professionals to concentrate more on analysis and critical thinking by automating repetitive jobs. The relationship between AI and human intelligence aligns with the broader understanding of AI as a complementary tool rather than a substitute for human expertise as supported by Martinez (2019).

5.4 AI Concerns

5.4.1 Data challenges & Risks

Murphy and Feeney (2023) have pointed out that as AI algorithms become more widely used, the risk of data breaches and illegal access increases. These worries were aligned with our findings, which underlined the significance of putting strong data protection procedures in place to secure sensitive information.

Our research findings uncovered a different aspect of data management difficulties, particularly with regard to data quality. The primary challenge encountered by data analysts was described by Respondent 1, who emphasized the widespread problem of data quality in a variety of companies. Although respondents acknowledged the importance of having clean data for enabling efficient AI-driven analytics, many gripped about the lack of perfectly clean data sets. According to Respondent 1, using AI technology for decision-making processes is severely hampered by the lack of adequate data.

Respondent 3, 5 and 6 also shared the same sentiments regarding managing data within AI frameworks, but from different perspectives. Respondent 3 highlights the potential legal repercussions of mishandling business data in AI models, which can lead to long term implications. Respondent 5 shares the same opinion and emphasizes the reputational and ethical issues associated with data exploitation, which might damage the company's reputation and damage consumer trust. Furthermore, respondent 6 describes the practical steps taken to reduce data risks, such as anonymization. Even though security measures were implemented, respondent 6 still acknowledges the ongoing difficulty in attaining total data protection. These viewpoints demonstrate the intricacy of managing data within AI environments.

5.4.2 Limitations

One key concern of AI usage is the overreliance on the technology (Murphy & Feeney, 2023). Our findings highlighted the limitations of AI usage: its inability to interpret data in the nuanced way that humans do. Although AI is excellent at identifying patterns, it frequently struggles to comprehend the meaning or underlying context of such patterns. In evaluating

AI-generated material, every respondent underscored the value of human oversight and contextual analysis.

Our findings emphasized the necessity of placing AI-generated content within the right context to derive relevant interpretations. Respondent 6 also highlighted the importance of analyzing data patterns within the context of psychological and social factors that impact human behavior, in order to optimize the analyzing process.

Together, these viewpoints highlight the complex interactions that occur when AI and human judgment are used to derive valuable insights from data. Furthermore, there is concern that an over-reliance on technology may result in an undervaluation of human experience and judgment. While data analytics and AI are powerful tools for problem solving and making informed decisions, professionals have to maintain a balance between using technology and retaining the human-centric aspects of their job as supported by Murphy & Feeney (2023).

5.4.3 AI Hallucinations

Furthermore, there are difficulties in data analytics due to the phenomena of AI hallucinations. Beutel et al. (2023) describe AI hallucination as the term used to characterize situations in which outcomes produced by AI significantly depart from reality or the intended context. The respondents conveyed concerns over the precision and dependability of insights offered by AI generated content, underscoring the need for precise data and the necessity of stringent quality control protocols.

AI-generated data could risk the validity and reliability of analytical insights in data analytics. Respondent 1 described AI hallucination as their primary worry while deploying their AI chat bot, highlighting the pivotal role of addressing this challenge in ensuring the reliability of AI-driven analytics initiatives (Beutel et al., 2023). Additionally, respondents stressed the need for critical thinking and quality assurance mechanisms in mitigating the hazards connected with AI delusion. For example, Respondent 1 and 5 created their own security techniques, such as context-based approaches and grounded generation, to prevent AI hallucination and improve the accuracy of AI-generated insights.

5.5 Accountability

Our findings showed a range of opinions about who should bear the responsibility in situations where AI is flawed. For example, Respondent 1 emphasized the significance of shared accountability between the vendor, consultant, and client, pointing out possible consequences for the AI sector if developers were held exclusively responsible. This view is consistent with that of Sattlegger, Klievink, and Tan (2022), who claimed that putting all the blame on developers could be harmful, and that other stakeholders should take responsibility. Respondent 3 emphasized the responsibility of consultants, particularly due to their expertise and role in utilizing AI tools, implying that the developer of the AI tool cannot be solely to blame.

respondents emphasized the significance of thorough testing and quality assurance to prevent AI-related errors. This is in line with the recommendations of Sattlegger et al. (2022), who support including accountability conditions into the design process in order to reduce potential harm. Respondents 4 and 5 touched upon quality assurance. Respondent 4 discussed it as an ongoing process, whereas Respondent 5 claimed it was something that should be done

prior to implementing AI in your organization. Respondent's 4 role could be considered both as a head developers and consultants, based on our provided scenario, since their organization processes both, whereas Respondent 5 is a consultant. The findings suggest that it is not only the developer's role to assure quality in the AI product, but the consultant as well. Respondent 3 and Respondent 5 also discussed upholding the quality of their respective brands, meaning that it is pointless to point fingers at the developers, since their customers are only interested in what they have been promised by the consultants.

The literature on accountability in decision-making AI technologies suggests that both humans and AI agents are held accountable for negative outcomes to the same extent, even though humans are generally attributed greater responsibility (Lima, G., Grgić-Hlača, N., & Cha, 2021). Consultants may use AI for their analysis to save time but have to realize that they are still held accountable if something in the analysis goes wrong. Respondent's 6 reply regarding this was focused on transparency between the different parties. This emphasizes the importance of understanding stakeholder perceptions in deploying AI systems, especially in high-stakes situations as also supported by Lima et al. (2021).

5.6 Discussion of the Findings with Sensemaking Theory

Sensemaking theory (Kaur et al., 2022) provides a framework on how data analysts perceive and experience the utility and usability of AI tools in data analytics. The theory explains how individuals make sense of their environment in order to navigate complexity and ambiguity. So, in our bachelor's thesis, sensemaking theory offers insight on the cognitive processes data analysts engage in while working with AI-driven data analytics tasks.

Interpretation of Ambiguity: When working with complicated datasets and new technology, like AI, data analysts frequently run into ambiguity and uncertainty. The first step in the theory is to attempt to understand and make sense of this ambiguity, which in the case of our bachelor's thesis findings is to understand the implementation of AI in the data analytics field.

Social Interaction and Group Interpretations: Making sense of perplexing situations requires people to connect with peers and rely on collective perceptions, which makes sensemaking fundamentally a social activity. Our findings showed that in this context of data analytics, analysts use collaborative sensemaking procedures to better comprehend the implications and uses of AI tools inside their work processes. This was shown by numerous of our responders, who both led and participated in various workshops within the domain area.

Enactment of Environment: Sensemaking involves not only interpretation but also action to actively change one's environment. Our bachelor's thesis findings showed that for data analysts this means adapting existing analytics processes and software to integrate AI tools.

Retrospective Reflection: The sensemaking process requires reflection on past experiences. In our bachelor's thesis findings showed that data analysts use their prior expertise working with data analytics to pinpoint possible areas for improvement using AI techniques.

Ongoing Sensemaking Process: Sensemaking is an ongoing and iterative process in which people constantly interpret new information and improve their knowledge over time. Our findings showed that as data analysts work with AI-driven data analytics tasks, their opinions about the tools' value and usability may change as they gain more experience.

Surprise and Mindfulness: Unexpected occurrences, such as encountering limitations or unexpected outcomes with AI tools, may prompt data analysts to engage in sensemaking processes to restore coherence. Mindfulness, or awareness of one's surroundings and actions, plays a crucial role in navigating surprises and adapting analytical approaches accordingly. Our findings showed that analysts' perceptions of AI tools are shaped by their ability to effectively address surprises and maintain coherence in their analytical workflows.

Driven by Plausibility: Sensemaking is driven by plausibility rather than strict accuracy, as individuals prioritize explanations that align with their values and views. Our findings showed that in the context of data analytics, analysts assess the usefulness and ease of use of AI tools based on how well these tools align with their analytical goals and objectives. Perceived usefulness may be influenced by the extent to which AI tools support analysts preferred analytical approaches and methodologies.

To gain a greater understanding on how data analysts comprehend and make sense of AI-driven technologies in their data analytics work, the sensemaking theory provides an essential framework. We are able to better identify and learn about the factors that affect a data analyst's perception of the usability and value of AI in data analytics, by focusing on the cognitive processes involved in sensemaking.

6. Conclusion

6.1 Conclusions

The bachelor's thesis examines the usage of Artificial Intelligence (AI) in data analytics. The purpose of the bachelor's thesis was to explore how data analysts perceive and experience the implementation and impact of AI-based tools and technologies in their everyday work regarding data analytics. Data analysts' experiences refer to their experiences from using AI in their analytical work, whereas perceptions refer to data analysts' general views on the topic based on their expertise. By exploring their perspectives and experiences, this bachelor's thesis study aimed to provide valuable insights on how the use of AI affects data analytics nowadays. For this purpose, we posed the following research question: How do data analysts perceive and experience the usage and impacts of AI in data analytics? For this, we conducted qualitative research where data were collected through semi-structured interviews from six data analysts working in companies located in Sweden and the United States. The collected data were analyzed thematically to conclude five themes. The five themes represent the empirical findings of this bachelor's thesis research and were followingly discussed with the help of the theoretical framework. The theoretical framework consists of the main concepts of data analytics, artificial intelligence (AI) and AI in data analytics, AI-based tools and technologies, and accountability in decision-making AI tools, and the sensemaking theory.

The research findings show that AI benefits in data analytics are acknowledged by data analysts, particularly in error reduction, bias mitigation, and time efficiency. However, while AI accelerates certain tasks, it cannot entirely replace human data analysts. The findings show that data analysts predominantly utilize specialized software for specific workflows, emphasizing control and transparency to mitigate risks associated with multi-purpose AI tools. The use of self-developed, tailored AI solutions enhances machine learning benefits for data analysis. Despite enthusiasm for AI, concerns persist regarding limitations, data integrity, and the need for ongoing education. Accountability remains a central finding, highlighting the necessity for responsibility and effective communication among stakeholders. While AI adoption in data analytics is increasing, manual workflows still prevail, reflecting cautious optimism and ongoing assessment of AI's capabilities and challenges. Sensemaking theory aids in understanding analysts' cognitive processes in navigating AI-driven data analytics tasks, emphasizing continual adaptation and social interactions. Analyzing the cognitive processes associated with comprehending insights produced by AI helps to clarify how analysts manage uncertainty, rely on social interactions, and adapt their analytical approaches continually.

6.2 Contributions

This bachelor's thesis research contributes both theoretically and practically. The theoretical contribution lies in the exploration of the dynamics surrounding the adoption of AI-based technologies in data analytics, including the examination of the benefits, limitations, and challenges associated with their integration. This includes insights into how data analysts perceive AI's role in error reduction, bias mitigation, and time efficiency, as well as their concerns regarding data integrity, limitations, and the need for ongoing education. Additionally, the application of sensemaking theory to understand analysts' cognitive

processes in navigating AI-driven data analytics tasks adds to the theoretical understanding of this area.

The practical contribution of the bachelor's thesis research outcome involves providing insights into how AI-based technologies can be effectively utilized in data analytics workflows. This includes understanding data analysts' preferences for specialized software for specific workflows, their emphasis on control and transparency to mitigate risks associated with multi-purpose AI tools, and the use of self-developed, tailored AI solutions to enhance machine learning benefits. Furthermore, the emphasis on accountability and effective communication among stakeholders underscores practical considerations for successful integration and utilization of AI-based technologies in data analytics.

6.3 Future Research

Our bachelor's thesis provides a comprehensive exploration of the integration of AI in data analytics within a broad context. However, there are opportunities for future research to focus more on specific aspects of this subject. One approach would be to research more about the accountability and ethical concerns related to AI-driven analytics. This would involve looking into the moral questions of AI, data handling procedures, and decision-making techniques. It would also involve looking into ways to make sure AI-driven analytics systems are transparent and accountable.

Alternatively, future research could opt to explore the functionalities and capabilities of AI-driven analytics software in more depth. This would involve conducting detailed analyses of specific AI tools, their algorithms, and their applications in various domains, but also specific analytical workflows. Through a detailed examination of AI software functions, researchers can acquire a deeper understanding of the workings of these tools, their advantages and disadvantages, as well as potential effects on data analysis procedures and results.

References

- Aldboush, H.H.H. & Ferdous, M. Building Trust in Fintech: An Analysis of Ethical and Privacy Considerations in the Intersection of Big Data, AI, and Customer Trust. *International Journal of Financial Studies*, Vol 11, Iss 3, p 90. MDPI AG. Available online: <https://doaj.org/article/c38e5fcf2bea468ea55ba182a8758b7a> [Accessed 14 April 2024].
- Annasingh, F. & Sesay, J.B. (2022). *Data Analytics for Business Foundations and Industry Applications*. Routledge, Taylor & Francis Group.
- Bajec, M., & Krisper, M. (2005). A methodology and tool support for managing business rules in organisations. *Information Systems*, 30(6), pp. 423-443.
- Beutel, G., Geerits, E., & Kielstein, J.T. (2023). Artificial hallucination: GPT on LSD? *Critical Care*, 27(1). Available online: <https://doi-org.ludwig.lub.lu.se/10.1186/s13054-023-04380-2> [Accessed 17 April 2024].
- Bowden, J.A. (2004). Capabilities-driven curriculum design. In C. Baille & I. Moore (Eds.), *Effective Learning and Teaching in Engineering* (pp. 36-47). New York: RoutledgeFalmer.
- Dolan, H.R., Alvarez, A.A., Freylersthe, S.G., & Crane, T.E. (2023). Methodology for Analyzing Qualitative Data in Multiple Languages. *Nursing Research*, 72(5), pp. 398-403. Available online: <https://doi.org/10.1097/NNR.0000000000000673> [Accessed 12 April 2024].
- Edge, D., Larson, J., & White, C. (2018). Bringing AI to BI: Enabling Visual Analytics of Unstructured Data in a Modern Business Intelligence Platform. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*.
- Ertel, W. (2017). *Introduction to Artificial Intelligence*. 2nd ed. Springer International Publishing: Imprint: Springer.
- Feeney, O., Murphy B. (2023). AI, Data Analytics and the Professions. In *The Future of Work: Challenges and Prospects for Organisations, Jobs and Workers* (pp. 35-51). Cham: Springer International Publishing.
- Gupta, M. (2023). Advances in AI: Employing Deep Generative Models for the Creation of Synthetic Healthcare Datasets to Improve Predictive Analytics. In *2023 International Conference on Communication, Security and Artificial Intelligence (ICCSAI)*, pp. 1026-1030. IEEE.
- Halawi, L., Clarke, A., & George, K. (2022). *Harnessing the Power of Analytics*. 1st ed. Springer International Publishing: Imprint: Springer.

Porter, M.E., Davenport, T.H., Daugherty, P. and Wilson, H.J. (2018). *HBR's 10 Must Reads on AI, Analytics, and the New Machine Age (with bonus article "Why Every Company Needs an Augmented Reality Strategy" by Michael E. Porter and James E. Heppelmann)*. Harvard Business Press.

Huawei Technologies Co., Ltd. (2023). *Artificial Intelligence Technology*. 1st ed. Springer Nature Singapore: Imprint: Springer.

Iansiti, M. & Lakhani, K. (2020). *Competing in the Age of AI: Strategy and Leadership When Algorithms and Networks Run the World*. Harvard Business Review Press.

Jacobsen, D.I. (2002). *Vad, hur och varför: Om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen*. Translated by Gunnar Sandin, edited by Caroline Hellström. Lund: Studentlitteratur.

John, R.J.L., Bacon, D., Chen, J., Ramesh, U., Li, J., Das, D., Claus, R., Kendall, A., & Patel, J.M. (2023). DataChat: An Intuitive and Collaborative Data Analytics Platform. In *Proceedings of the ACM SIGMOD International Conference on Management of Data, SIGMOD 2023*. Association for Computing Machinery.

Lima, G., Grgić-Hlača, N., & Cha, M. (2021). Human Perceptions on Moral Responsibility of AI: A Case Study in AI-Assisted Bail Decision-Making. Available online: <http://arxiv.org/abs/2102.00625> [Accessed 17 April 2024].

Liu, C. & Downing, C. (2024). Using Text Analytics AI Insights in Microsoft Power BI Desktop to Score Sentiments, Extract Key Phrases, and Discover Unstructured Data Patterns. *Journal of Information Systems Education*, Winter 2024, Vol. 35, Issue 1, pp. 48-55. Available online: <https://eds-p-ebSCOhost-com.ludwig.lub.lu.se/eds/pdfviewer/pdfviewer?vid=22&sid=336ff163-2d06-4f99-9649-7dfd9a06d366%40redis> [Accessed 17 April 2024].

Martinez, R. (2019). Artificial Intelligence: Distinguishing between Types & Definitions. *Nevada Law Journal*, Vol. 19, Iss. 3, pp. 1015-1042. Available online: <https://heinonline-org.ludwig.lub.lu.se/HOL/Page?handle=hein.journals/nev19&div=32> [Accessed 17 April 2024].

Oates, B.J., Griffiths, M., & McLean, R. (2022). *Researching Information Systems and Computing*. 2nd ed. SAGE.

Russell, S.J. & Norvig, P. (2021) *Artificial Intelligence: A Modern Approach*. Fourth edition global edition. Pearson Education Limited.

Saivasan, R. & Lokhande, M. (2023). Exploring Use Cases of Generative AI and Metaverse in Financial Analytics: Unveiling the Synergies of Advanced Technologies. *International Journal of Global Business and Competitiveness: A Journal of Global Institute of Flexible Systems Management*, 18(Suppl 1), pp. 77-86. Springer Nature Singapore. Available online: <https://doi.org/10.1007/s42943-023-00082-2> [Accessed 16 April 2024].

Sattlegger, A., Van den Hoven, J., Bharosa, N., Hagen, L., Solvak, M., & Hwang, S. (2022). Designing for Responsibility. In Proceedings of the 23rd Annual International Conference on Digital Government Research (dg.o 2022), Virtual Event, South Korea, pp. 499-499.

Sen, M., Sen, S.N., & Sahin, T.G. (2023). A New Era for Data Analysis in Qualitative Research: ChatGPT! Shanlax International Journal of Education, 11(1), pp. 1-15. Available online: <http://www.shanlaxjournals.in/journals/index.php/education> [Accessed 17 April 2024].

Zheng, Y. (2019) A comparison of tools for teaching and learning data analytics. In Proceedings of the 20th Annual Conference on Information Technology Education (SIGITE 2019), pp. 160. Association for Computing Machinery, Inc.

Appendices

Appendix A. Interview Guide

Opening Questions

1. Can you please state your name and current role, and how long you have been working in that role?

Main Questions

2. Could you share with us some examples of your work (How do you work with data analysis in your everyday work)?
 - a. How do you work with quantitative data analysis?
 - b. And qualitative data analysis? Non numerical data.
3. Do you use AI-based tools for your work (the data analysis)? Could You Share some examples with us? (maybe explain what AI-based tools they use for that).
4. Have you ever used an AI assistant, such as GitHub pilot? Yes or No?
5. (If yes, what are your takes on it and do you see it being useful data analysis?)
6. Have you ever used generative AI, such as ChatGPT? Yes or No?
7. What are your thoughts about AI assistance when handling big data?
8. Have you ever used business intelligence tools such as PowerBI? Would you like to see it more AI integrated?
9. Based on your experience, what are the benefits you see with the use of AI in data analysis?
10. Based on your experience, what are the concerns/challenges seen with the use of AI in data analysis?
11. What is your opinion on AI in terms of safety in relation to data analysis? And sustainability in relation to data analysis?
12. What is your opinion on AI in terms of ethical concerns in relation to data analysis?
13. What are your thoughts on AI confabulation (AI software that is 100 percent certain of its thing when wrong, rather than admitting that it doesn't know the answer)?
14. A vendor provides a consultant with an analytical AI tool. The consultant uses the tool to help their customer with an analysis. The customer invests money based on the consultant's analysis. However, the analysis was wrong due to a mistake made by the AI tool which made the customer lose money. Who is to blame in this scenario? The vendor, the consultant or the customer?

Closing questions

15. Is there something else you would like to add in our discussion that you think we missed out?
16. Would you like us to send you the transcription of our interview?

Appendix B. Informed Consent Form

Artificial Intelligence in Data Analytics: A Qualitative Study on how Data Analysts perceive and experience the usage of AI

By **Tony Do & Adam Jacobson**

I have been given information about Artificial Intelligence in Data Analytics: A Qualitative Study on how Data Analysts perceive and experience the usage of AI and discussed the research project with Tony Do & Adam Jacobson conducting this research as a part of a Bachelor's thesis in Information systems supervised by Niki Chatzipanagiotou.

I understand that, if I consent to participate in this project, I will be asked to give the researcher a duration of approximately 45 minutes of my time to participate in the process. I understand that my participation in this research is voluntary, I am free to refuse to participate and I am free to withdraw from the research at any time. By ticking this checkbox below I am indicating my consent to participate in the research as it has been described to me. I understand that the data collected from my participation will be used for my thesis only and will be deleted after completion, and I consent for it to be used in that manner.

Name:

Email:

Telephone:

Appendix C. Interview 1

Interview 1 (R1), Qlik representative, Length of interview: 40 minutes, X = respondent.

Tony:

So can you please state your name and current role and how long you've been working in that room?

X :

Sure. My name is —. I'm the principal technical program manager for AI and chat bot. I've been at Qlik in this technical program manager position for about three years now, and yeah, three years. But before that I also worked at another tech company called Citrix as a data analyst for about two years.

Tony:

Cool. So in this message paper we are looking for and respondents that have been working with the data analytics within AI. And so I'm not quite sure if you have done it directly or maybe you said previously that the first position was strictly data analytics and and the current one is a, but it's the integration between those two.

Tony:

So can you give us some examples of your work and how you deal work with the data and analysis and AI?

X :

Sure. But so for example, right now as a technical program manager, I own and operate and maintain our AI chat bot. So I deployed it, manage it, make improvements on it and also do reporting on it. So that's where I think it comes into play.

Tony:

Yeah.

X :

What you guys are looking for, like the combination of AI and analytics that comes into play, think I have a couple of things I can share from what I do currently or experience this with on how I see the impact of AI and analytics and especially working for a company like Qlik who is a data analytics software company and we're implementing a lot of these AI features in. I do hear what we're working on, what we're investing in, what we're developing, that I'm kind of partaking in as well helping them develop the new features.

Tony:

Yeah.

X :

That includes AI and such that I can't go too detail into, but I do can't share some examples so.

Tony:

Perfect. OK.

Tony:

So if we say we're allegations regarding the analysis work, how do you work with quantitative data analysis? Can you give us some examples?

X :

Sure. Like quantitative data wise, if we're measuring my chat box performance like AI agent, one of the big things that we look for is pretty straightforward. You know the first one's volume, you know. How is it are people using it? How many times is it triggered? You know from which location is it being triggered the most often and such, and it also going into measuring the. Customer experience by looking at the CSAT score, the surveys that people fill out after talking with the chat, but are they happy or are they not? What is the overtime trend after making improvements, especially the number one metric for the chat, is what is called containment rate and that is the percentage of times that customers talk with the chat bot and just end it there. They don't ask for live agent help or to submit a ticket to Qlik support to for further assistance. No, they came in. Talk just with the chat bot, the AI assistant, and then leave afterwards. So that's like a biggest metric. So that's like few of the examples of using quantitative data in my work.

Tony:

Alright, how about the the non numerical data such as the qualitative?

X :

Sure. So that's, that's where the the money's at, I'll say. The golden egg, I'll say. Cause. Yeah, it's great to see, like more people are using it, you know, thousands of people having conversation with the chatbot. Great. What are they talking about? That's where the challenge comes in, and I'm sure I can elaborate more and ask more questions about it down the line.

Tony:

Yep.

X :

I read your list of questions, so I expect some of them will ask follow up.

Tony:

Yeah.

X :

So the chat bot utilizes large language model such as GPT to become a generative chat bot, so it's not a chat bot that you just Qlik buttons in like how can I help you select from above? No, it's like you type in anything and the chat bot is trained on thousands and thousands of our articles and content within Qlik. So it's like a Qlik expert. Now the problem is is great that we see increase in let's say the conversation volume, but what are people asking? You know, that's the way the qualitative data comes in because those are not numerical. Those are just. Conversational data. It's like your text message. You know, data, there's like going back and forth between the user and the chat button. So how do you measure that? You know, how do you quantify that? You know, it's like, what are ways you can use that data and that's where the challenge has been before. And there are

a couple new ways we're tackling it that is allowing us to kind of get a better reporting on it.

Tony:

Yeah.

X :

When we went from traditional chat bot to generative chat bot, that's one of the thing that we lost for short amount of time is if you ask me what are the top 10 questions people ask well with the traditional chat bot it's easy cuz it's just button Qliks so you can just quantify it in the data. But for generative chat bot like chat should be T like how do you quantify it?

Tony:

Yeah.

X :

It it's hard, you know, but how we go about resolving that before I give it back to you is we actually reutilize tech. Appt.

Tony:

OK.

X :

To help us quantify that which is, which is kind of funny, we get the data from GPT and we add GPT to quantify it for us and I can go more into depth, but I would like to pass it back to you for any kind of next questions.

Tony:

First, yeah, yeah.

Jacobson, Adam - STA:

Of course.

Jacobson, Adam - STA:

So we'll continue with interview.

Jacobson, Adam - STA:

As with send it to you, but as you mentioned with we've covered some areas already like like, but uh, there's question obviously is do you use AI based tools for your work. That being said, there analysis.

X :

Umm.

Jacobson, Adam - STA:

So essentially you covered the perspective of the user where they would, uh Qlik into certain fields when using your chatpad.

X :

Umm.

Jacobson, Adam - STA:

But for your analysis specifically, do you use any sort of AI based tools for the analysis itself?

X :

Umm.

X :

Umm. Umm.

Jacobson, Adam - STA:

And one question is obviously have you ever used an AI assistant such as GitHub pilot which works with coding, but in this case for the analysis itself?

X :

Let's see, I used a based tools.

X :

I wouldn't say there's like a tool like the chat bot platform that we use does come with AI reporting tools, and that's where they I mentioned earlier, like it helps us to kind of show what are the, you

know, ways we can see that qualitative data for example this recently they released a feature because we were complaining about it's like hey it's great that we went to a generative AI assistant now but I don't have the transparency or insights I used to I don't know what the top questions are I don't know what the question. People ask that my chat bot doesn't know, and that's where they the feature that they developed by working with their customer and going from their feedback is now they have they utilize GPT 4 model to basically read through the conversation and the GPT will give you a summarization and then a tags of what the topic is and if that using like using that. Now you can, you know, kind of quantify it, OK? Like the tag is password reset. That's and then you can quantify it to see if that's one of the top questions you're being asked. And a great thing in also added in was it can assess the conversation and tell you depending on the AIS intelligence, if it's going to label it as all the issue was resolved or issue couldn't be resolved and why. So using that we get a different metric of OK AIS. Well, basically assessment of the containment rate. Did the customer get this answer? Yes or no. You know that used to be a very straightforward thing with the previous chat bot. We just assumed that if at the end of the answer they didn't ask for a live agent help, we just assumed that the issue was resolved, but now it's a lot more accurate because now you have this AI intelligence assessing, reading and understanding the context of those conversations because it is trained on Qlik products to say, OK, they asked about this error code and they answer that we provided was exactly to do with how to fix that error code. This is stamp of approval. It is resolved, so that is some of the reporting capabilities that I use within the chat bot platform. Besides that, I will uh for AI based tool. It's more of me, personally. Because I use Python, uh, throughout my work, sometimes working with the chat bot and it's always great to use, you know Python for a couple works and that's where I kind of try to try to come up with my own solution before they came up with this new feature was utilizing natural language processing to assess what are like the top common clusters of topic within that historical data of let's say, last three months the conversation

data. So and then you asked, have I ever used AI system tool? Yes, because I operate the chat box. That's our AI assistant.

Jacobson, Adam - STA: I mean, yeah, of course I would consider that to be.

X : Right.

Tony: What?

Jacobson, Adam - STA: Yeah, but it's still me personally.

Jacobson, Adam - STA: But like is it so it's not like bot software that you have integrated into your system.

Jacobson, Adam - STA: It's like built from scratch in your system, so to speak.

X : Almost from scratch, where the Bender made it easy. It's like a zero code low code environment, but then you get to build your own AI assistant from scratch. Yeah, they just provide the engine for it.

Jacobson, Adam - STA: I see.

Jacobson, Adam - STA: And do you have?

Jacobson, Adam - STA: Do you have a name for it? X : Yeah.

X : The company is called ADA ADA Dot CX is their website. If you just search ADA, you'll go to.

Tony: Crypto yeah. Jacobson,

Adam - STA: Thank you. OK.

Tony: Yeah. X : I believe the different association for Disability, I think, but ADA dot CX, they provide chat bot basically platform where you can build your own AI system.

Jacobson, Adam - STA: Boo uh, I guess we'll continue down, but some of these errors and all have been covered.

Jacobson, Adam - STA: Uh, we haven't really touched upon big data, or at least we have to some extent, but that depends on how you define big data.

X : Umm. Mm-hmm. Jacobson, Adam - STA: So I'm talking about the big the 5V's, so velocity in volume and so on and so forth.

Jacobson, Adam - STA: Would you consider your work to be big data?

X : Umm, yes and no. I've always been kind of like on the skeptical side of big data, big data because, you know, it's just skeptical side of me who gets annoyed with all these lot of changing trends and people love putting new labels to make it seem like they know what they're talking about, even though it's all the same thing, like big data is like, sure, you can call it big data because it has so much stuff in it. Like you said, velocity, volume and all that. So yeah, I'll say yes, but it's just my personal opinion. I think it's overrated. Hey, that's just data, it's.

Jacobson, Adam - STA: Being here, no, it doesn't seem to be like a great song between big data and data.

Tony: Yeah.

Jacobson, Adam - STA: Like, where's the threshold?

X : Right.

Jacobson, Adam - STA: Yeah, And would you like to continue, Tony?

Tony: Sure. Tony: So and this is the tangent tool and Qlik is one of the main product products.

X : Umm.

Tony: But do you use? Tony: Yeah, strictly Qliks, business intelligence or the other probabilities and and how's the integration with these tools?

X : Sure. So before I joined Qlik working as a data analysis might actual preference was Tableau.

Tony: Yeah.

X : I used tableau for years and then I joined Qlik and then learned about Qlik sense and stuff.

Tony: Yeah. X : But it's hard. I'm not biased, this is just my assessment, but it's hard to do that wearing this T shirt, but I think Qlik sense has a lot more AI features than Tableau.

Tony: Yeah, yeah.

X : Yeah, there's a lot of new features like AI features being added in.

X : One cool feature that or trend that I see a lot of BI tools going into is adding in a business intelligence AI into it where you wouldn't need to hire our own data analyst or business intelligence employee.

X : No. Like one of the thing like is trying to develop an all their BI are trying to implement is adding your own GP T into your analytics app where you ask that let's say analytics bot.

Hey, what's my sales like in the last two quarters? What's the increase and decrease? It was like, sure, your sales have decreased in 5% in the Americas region, but increased 15% by you know blah blah blah and the EMEA region, you know like building that type of AI intelligence into a business intelligence platform. I see that it's where people are going into right now

and I know that's something that Tableau, Salesforce, even power BI. They're all trying to add it in there, even Microsoft trying to add in, like, probably copilot into power BI and it's very soon if not already. We're now, it's not gonna be a persons opinion on the data, but more of large language models, you know, assessment of the data and telling you what it thinks is gonna be.

Jacobson, Adam - STA: That's essentially why we wrote about this.

Tony: Yeah. X : Umm.

Jacobson, Adam - STA: This topic, like the intelligence part of the the artificial intelligence versus the human intelligence, and therefore we'd like to ask you, like, based on your experience, what are the main benefits with the use of AI and data analysis?

Jacobson, Adam - STA: What comes to mind when you think of the benefits with AI in the analysis in the yeah.

X : I think one of the biggest benefit is kind of a concern too, but I'm excited for it because overall it's a good like a it's a good for the general public, although as a person who wants to stay in like data analytics and data science, field is kind of dangerous for me for my career path. But what I see is originally, if you wanted to do some forecasting prediction simulations, you would need to hire a data scientist. It used to be such a niche skill that requires years of training and studying and experience to do it right. But now, with AI being embedded into these analytics platform and being able to assess your data and make you know like it's own insights from it, I see in the future where like I said earlier, position like business Intelligence, data analyst or even data scientists, that's the biggest one that I think is gonna be changed because if you ask a startup company, hey, what's your projections for your numbers for next two years? Well, well, based on your road map and stuff like that or have you run a simulation on the viability of your product among the customer base that you're targeting? I have a data scientist friend and Microsoft who does this for a living and now Microsoft has to bank to fund those teams, you know, and really get a basically all seeing, you know, globe, you know, the glass globe, a fortune teller and these are what I kind of think data scientists are lots of times can be sort of like you know fortune tellers can tell slightly what might happen in the future. But it's like a multiverse. It Michael this way or that way we can only tell you what are the possibilities of it.

Jacobson, Adam - STA: Umm. X : But with AI going into these tools, business intelligence tools, I see that power being available for anyone, even business people who don't have the skills of business and data analytics or data science.

X : You know, it's just like how GPT didn't just become the new Google. X : No, it enabled non technical people to do technical work.

Jacobson, Adam - STA: Yeah.

X : You know you can use GPT to create your own website and just couple hours, whereas before you need to go to like a six month boot camp or it or like you don't know how to code, you can ask it to code for you.

Jacobson, Adam - STA: You.

X : You know that kind of like enabling the general public to another set of skills that was kind of like, not gatekeeper, but there was a big threshold to make it through. It's gonna open up very soon. That's what I kind of predict.

Jacobson, Adam - STA: So depending on whose perspective it could be considered beneficial or concerns essentially.

X : It could. I am about to start my Masters program for Computational analytics would go into data science and stuff, and I'm sure they'll cover that in my curriculum.

Jacobson, Adam - STA: Yeah.

X : It's like, hey, there's these new AI models coming out that can do lot of these four casting and simulations easily, you know, and I'm going more into Azure AI portfolios of their stuff to kind of implement into our work and projects and I see them like you can do these type of simulations and forecasting with a Qlik of a button. You just load your data in and say press Qlik and that's it, whereas before you need to go, you need to know how to do Python And R and develop your own code how to do it. X : You need to know how to use all those libraries and come up with the logic yourself and boom that was a data scientist role in this full time job, but now they're like Qlikable button. X : It was like, hey, here's your top trends.

Jacobson, Adam - STA: Yes. Tony: Anything.

Tony: OK. Tony: So if we look at at use the AL in the analytics from another perspective. Tony: So basically experience, what are your concerns and challenges?

X : I think the biggest challenge is same challenge that all data analysts face is the quality of the data. X : Umm, I have yet to work for a company where the data is so clean. X : I'm like, wow, this makes my job so easy. X : And I was. X : I just need to use it like. X : No, it's like and it's. X : I've come to accept the fact that there's no such thing as perfect clean data. X : There's always discrepancies. X : There's always some cleaning to do before you use it for these tools, and that's one of the things I have to make sure and put in extra effort for my AI assistant chat bot launch. X : It's good. X : Yeah, it's just in chat. X : Bot is only as good as the training from the data, so I think that's one of the big kind of oopsie moment that I see a lot of companies and teams or even professionals out there who are now getting exposure into AI to learn that lesson first hand. X : And it's like you're a I will turn into a crapshoot if you don't have the good data to back it up. X : And that's one of the big challenges and best outside of a level beyond that is OK. X : There is a lot of these, you know, data that we need to use. X : Another challenge I face is if you're working for a corporate like say, even medium or large size large size for short is to get your hands on that data. X : But it comes with a lot of security concerns. X : It's there's a lot of political battle about it too, because I had to pull out so many teeth to get this AI assistant approved because there is lot of concerns security wise with large language model with AI that uses large language model cause it's unpredictable. X : You know you can put stuff in and you don't know what it's gonna use it for. X : You know, for example, not too many people are knowing or worried that open AI has a lot more data on you than you would think.

Tony: Right.

X : Even recently, with the whole scandal of their CTO on interview, like Squeamish to admit that they've been using YouTube videos out there.

Tony: Yeah.

X : Although she did admit it, so it's not an official statement, but from her look of her face that she didn't want to admit that she's they're open. X : AI is using all the available data that's on the Internet, regardless of what the security protocols are to train their models for the next product. X : And that's

why, like our internal security teams also concerned with these type of AI tools is what can you feed in and make sure it doesn't hold on to that information. X : Because if we're giving if customers are using our Qlik sense product with AI features, they're gonna be worried. X : For example, our healthcare industry is where big concern with that is like. X : We don't want any AI features because we're worried this is public health data of our customers and all like patients that has one of the strictest rule on security. X : You know, with data and where it can be used for. X : So I think that is a big concern.

Tony: Very interesting.

Jacobson, Adam - STA: Thank you so much very much. Alright, so we've covered safety to some extent. Do you have a say in sustainability in relation to data analysis?

X : Sustainability wise, like the job market, it's a position.

Jacobson, Adam - STA: Well, our our literature essentially covered, uh, stuff that necessarily aren't related to to the analysis itself.

Jacobson, Adam - STA: So I believe we will we we might eventually scratch that. Uh, next question covers AI confabulation as a professor in Lund. So easily put, it essentially means that AI software that sounds 100% certain of their thing, rather than admitting that it doesn't know the answer.

X : Yeah, yeah.

Jacobson, Adam - STA: And so for example, with Chachi, Bettie would have data from prior to no after it's launch, where the chibitty essentially gave answers that sound that correct but weren't so like. But considering what you said earlier, like it felt it feels like data scientists such as yourself needs like the need for them, will always be there in case these like sort of things happen.

X : Yep, Yep.

Jacobson, Adam - STA: So what are your thoughts on this? Jacobson, Adam - STA: Exactly. Jacobson, Adam - STA: That's the same thing.

Tony: Yeah, yeah.

X : Yeah, I can share a little bit about that as well cause, yeah, another term that they call is like hallucinations, you know, like the yeah AI hallucinations. Jacobson, Adam - STA: Yeah, yeah, yeah, yeah.

X : And that was our biggest concern when we were trying to launch AI assistant is how do we prevent hallucinations from happening? X : And I can share this design architecture that lot of these generative AI chat bots that they're selling goes for is what it's called, utilizing the method of grounded generation and what it is is you're we're using GPT 4 as the main engine or brains of it. X : But we're not using it for information. X : It's just using it for logical standpoint and you connect that brain to library of contents and you're designing this brain to only use these libraries you connected to. X : So with my chat, but my chat bot has a basic intelligence because it is, you know, connected to GPT 4, but it won't use any of this trained data from the GPT 4 model to create answers only to form sentences and to understand the content it reads the content it reads is the content I connect to that I as the the developer connect to our Qlik library of our content such as our product document of how to use our and install our product and then support articles which is how to break fix our products and with. X : That method it kind of prevents majority of the hallucinations is still can occur for and I go into that in a little bit. X : But yeah, using that grounded generation method, it kind of prevents a lot of that hallucination from using its pre trained data rather than the newly trained data that is verified. X : And this is a hallucination. X : Is a big concern for a corporate environment or industry. X : Whoever wants to implement, if it's just company wise, you're legally you're legally liable for your customers. X : You

know, if you if you're AI assistant gives the wrong answer to the customer and they're basically infrastructure goes down and they're lose business, they lose money, they lose clients and stuff and it's like. J

Tony: Yeah.

X : Even I would go after the vendor and you know, with a lawsuit, you know. X : So that's why with my AI like when I first proposed the idea of deploying AI assistant, that was the biggest concern they had. X : And even from the executive level was like, what if the AI gives wrong answer and it messes up the customer so bad that they leave? X : You know, aren't you liable for that? X : Then you know it's like, well, how do you prevent that? X : And there's like, grounded generation. X : But now here's where grounded generation still is not perfect. X : For example, we have. X : Here I Qlik multiple products with different names but very similar names. X : Qlik sense, Qlik view, Qlik and printing. X : But and a lot of them are like the cloud version or on Prem offline version. X : You know cloud customer comes using Qlik Sense cloud and ask I'm having this error. X : How do I fix it? X : My AI assistant can't tell if it is cloud or on Prem version because the customer never wrote it in their question. X : They just said I just have this error now. X : If they included in their question, I'm using Qlikonce cloud and I'm having this error. X : It will give the proper answer, but it does. X : I'm having

this error and that is totally justifiable on the customer side because they're just having the issue and they're just letting them know they didn't know. X : They didn't. X : They needed to provide all this extra information to make sure that you get the most accurate thing, and that's where, like the adoption of these GPT chat bots come into play. X : The general audience is still learning. X : You know the users, they don't know what's the best way to come up with a prompt and that's why there's people selling classes online of prompt engineering. J. X : How to write the best prompts and those people are probably making a lot of money right now for not much knowledge out of it in there. X : But yeah, going back to my point. X : So because the tap bot doesn't know if it's cloud version or on Prem version, it accidentally provided the on Prem version. X : Here's how you go about fixing that. X : Press tier press. X : I press there and the customers like there's no such thing. X : There's no button like that in my page it what's going on, you know?

X : Or like they try to follow it and turns out it's only for the on Prem version.

Jacobson, Adam - STA: Yeah.

. X : If you do that on the cloud, it shuts down as like you just broke my stuff even more like Oh my God, I'm about to lose my job because of this, you stupid AI, you know, like, so that they're definitely ohh big risk with it, especially in a liable field like corporate or sales.

Jacobson, Adam - STA: Umm.

X : Uh, it feels just for your own personal use. X : OK, give you a wrong service. X : So big one, you know, nobody got hurt, you know. Tony: Yeah. X : But if we're implementing AI to everywhere, including health industry, you know healthcare providers, that's where my biggest concern is I see the biggest money in it, but I see the biggest danger in it. X : If you're relying on AI and then it gave the wrong kinds of diagnosis, which I know there are people who are using chat, GPD to self diagnose and self treat themselves OK, you've wrongly diagnosed and you die from it or you get hurt from it or disabled like this open AI liable. STA: To. X : It's like probably not cause they got great lawyers.

Jacobson, Adam - STA: Yeah. X : So like this stuff.

X : But that's where my biggest concerns are with AI hallucinations is in couple settings. X : It's not as harmful. X : It's not as dangerous, but with everybody putting this sticker of AI and every single place they can think of, like eventually that kind of catastrophe is inevitable. X : I think it's gonna happen someday.

Tony: Interesting.

X : Right.

Jacobson, Adam - STA: Especially with companies to like chase monetization and like risking the the secured outfits.

X : Yep.

Tony: Yeah, more or less the same.

Jacobson, Adam - STA: So yeah, question 11, Tony, that essentially we've covered it, yeah.

Tony: Yeah, exactly.

X : Umm. Tony: OK, so we can go with the close equations and is there something else you would like to add in our election that I think that were missed out and comes to this domain?

X : Sure.

Jacobson, Adam - STA: Before we get into that, I'd like to ask you just in the query. Jacobson, Adam - STA: So question 11 was essentially what you talked about. X : Umm. Jacobson, Adam - STA: A company makes a business decision based on AI data analysis that was wrong. X : Umm. Jacobson, Adam - STA: Who should be held accountable? Jacobson, Adam - STA: Is it the vendor or is it the company? Jacobson, Adam - STA: Or is it?

Jacobson, Adam - STA: Obviously not the customer, but.

X : I think it should be kind of shared.

Jacobson, Adam - STA: Yeah. J

in : It should be shared across cause if you make the business liable.

X : Let's say if you make Qlik liable for is Qlik AI providing answers that might have been wrong and it caused a business impact on the customer side.

Tony: Yeah.

X : If you make the just the developers liable 100% all the time, that risk is gonna get so big that it's gonna start to actually shrink the AI industry where, like people are gonna be too squeamish. X : The risk is too big where one mess up from having a billion dollar lawsuit that's gonna bankrupt the company. X : Who's gonna wanna go into that?

Tony: Yeah.

X : You know, if we really want AI to flourish and wanna bring a greater good to the general public. X : Like a nonprofit. X : Kind of mission goal.

Jacobson, Adam - STA: Umm.

X : If we really wanna live together with AI and take all his benefits and can provide to ease our lives and help people around the world, I think the load of risk needs to be divided among everybody. X : Everybody is liable. X : Everybody has due diligence. X : They must do when utilizing AI tools and that's I think the best approach for it. X : You can't be like Openai. X : Whereas 100% is like the customer side, it's like hey, it's like disclaimer, this is just a model, don't follow whatever it is and

they're they're taking their precautions where you see even articles that they're censoring a lot of stuff that they're using prompts that it's a little bit controversial because they want to protect themselves from the law. X : And then eventually it's going to skew to a point where this open AI large language model, the greatness kind of shrunk down a little because they had to compromise for the risk and stuff. X : So to prevent too much compromise to make sure that we're making progress with AI to make benefits, I think to load should be yeah, spread across all parties. X : The users needs to know too and that should become the standard too. X : You come up with your own risk.

X : It's just like the Internet.

Jacobson, Adam - STA: Yeah.

X : Internet can't be liable.

Tony: Yeah.

Jacobson, Adam - STA: Yeah. X : The worldwide web is not liable for youth falling into scammers or identity theft and such.

X : You knowingly going into online that there's some risk to it, that your photos being posted on Facebook and Instagram Mike gets created by people on the dark web. X : You know, it's such so like it is scary, but you need to have your own self practice.

Tony: So the last question that was asked earlier. Tony: So is there something else that you would like to add to our discussion regarding this domain area?

X : I I think you guys have a great topic here. I get why you guys are writing your research thesis on this one, I think perhaps I would love to see more of as a next step might be discussion of not just AI, but AGI artificial general intelligence.

Tony: OK. Yeah. X : What does that have impact on the data analytics side?

X : Does that really wipe out the whole role of data analysts? X : How good can it be? X : You know, if we were to go with artificial general intelligence, I mean, of course, maybe not like maybe we still need people with PhD. X : But does this mean that really, unless you get a PhD that the human resource is not as valuable anymore? Tony: Umm. X : You know, like is it worth getting a bachelor's at that point, unless you're getting a PhD like it's AGI gonna wipe out the whole like university structure and reform it to where you need to be smarter than azi? X : And what's that thought of ACI? X : What's the IQ level? X : Let's say the first version of Avi is IQ 100, where it's just a general like high school graduates. X : Listen, and you can teach us stuff to become smarter. X : Ohh with it stops there.

Jacobson, Adam - STA: Umm.

Tony: Yeah.

X : Would it keep developing where IQ reaches 150, where it's like a Mesa level where like a PhD level person? X : Level person. X : If that's the case, then who? X : Like where? X : Where do you draw the line of ethics on human resource? Tony: It's crazy, yeah.

Jacobson, Adam - STA: Yeah. X : You know, cause that's crippling the human, you know, species is enabling, but crippling at the same time.

Tony: Yeah, yeah

X : The only 1% of people will be able to stand above it. X : You know, like that's kind of like ethical talk about the next level of artificial intelligence, I think would be an interesting thing to discuss.

Tony: Very interesting.

Jacobson, Adam - STA: Well, in that case we would just like to ask you if you would like to send you, we would like to send you the the transcription, but you're the one who's owning it.

Tony: Yeah. X : Yeah, no problem.

Jacobson, Adam - STA: So I guess it's vice versa.

Tony: Yeah. X : Yeah, I'll get the transcript.

Jacobson, Adam - STA: Yeah. X : And forward it to you guys right away.

X : When it becomes available.

Tony: Perfect.

Jacobson, Adam - STA: In that case, you may stop recording because the interview is over.

X: Sure.

Appendix D. Interview 2

Interview 2 (R2), Occto representative, Length of interview: 35 minutes, X = respondent.

Jacobson, Adam - STA Vår första fråga, kommer vi be dig uppge namn, roll och hur länge du har jobbat inom den rollen.

X - Jag heter ----- och min roll är väl cloud engineer och chefsarkitekt och den rollen har jag haft i olika bolag i 15, 20 år nästan. Och just nu jobbar jag på ett företag som heter OccToo som jobbar med att samla in massa olika datakällor och konsolidera det och exponera dem till olika system. Olika e-handelsplattformar och andra lösningar. Så det är mycket data, datadrivet det vi gör. Och så jag är också en av medgrundare till företaget. Jag vet inte om ni behöver någon mer info där kring det.

Jacobson, Adam - STA: Nej, det blir jättebra, Tony, ska vi köra varannan fråga?

Tony: Naturligtvis. Ja, jättebra. Vi skulle kunna dela med oss av hur ditt jobb ser ut vardagligen, hur ni jobbar med dataanalys.

X : Ja, vi har ju en plattform som vi har utvecklat där vi har två olika nivåer. Det ena är ju att vi har våra kunders data som vi suger in och som bearbetas och som används, det är ju en del så att ibland så är jag med och jobbar med våra kunder liksom med deras data. Man kan göra den, hur gör det bättre? Sen har vi ju också hur produkten utvecklas och liksom hur den tuffar på och vi har ju en produkt som betjänar många olika kunder. Och där är det ju inte kopplat till specifika kunder, utan det är systemet i sig. Hur det, hur det mår och hur det liksom presterar.

Tony: OK, och hur ser fördelningen ut med arbetet med då den kvantitativa dataanalysen, processen och den kvalitativa?

X : Ja, men jag skulle säga när det gäller själva produkten så är väl det lite mer åt det kvantitativa, där vi kollar prestanda. Hur bra saker och ting går när vi gör uppdateringar, vilka effekter får det liksom ifrån det? Det vi gör är att vi tar väldigt stora datamängder som vi kör igenom systemet liksom och vi vill få igenom så mycket som möjligt så snabbt som möjligt. Så där jobbar vi.

Där är det ju mer om att kunna se det kvalitativa än det kvantitativa. Det är ju det här, menar jag. Vi tittar på kundernas olika datakällor som vi jobbar med och förädlar där är det ju går vi in och försöker liksom komma till lite insikt och och hjälpa dem hur de ska strukturera sin data för att det ska komma ut rätt. Sen är det även också när datan vi tar in och sedan används den datan och när den datan används så ger ju det också insikter som att det är någon feedback loop som man gör en kampanj på rätt rätt. Hur bra presterade den och, liksom klickar folk på den, om någon klickar på den - köpte de produkten och då får man tillbaka den informationen som ger dem bättre stöd för att liksom hur de ska jobba i framtiden så.

Jacobson, Adam - STA: Då är det inte bara produkten som ni analyserar, utan också hur den förhåller sig gentemot kunden. Det är ju lite som du sa här beslutet att ni också analyserar åt dem via verktyget.

X : Precis, det gör vi.

Jacobson, Adam - STA: Nästa fråga handlar om AI-baserade verktyg. Skulle man kunna kalla detta AI-baserat, den här funktionen som du sa i slutet? Eller är det liksom en algoritm baserad på kod?

X : Vi har vissa delar som är AI-baserade som vi håller på och faktiskt och utvecklar nu och det är. Det handlar ju om att använda AI för att förbättra allt det här med funktionalitet till exempel. Vi har deras data och så lägger vi en maskin och motor på den och sen så används den sen när de ställer frågor mot den och får bättre svar. Men när det gäller den här feedbacken: Det finns säkert förbättringar man kan göra där med hjälp av AI, men det är ju det är ju rena fakta alltså, så det är ju inte att vi bidrar med slutsatser utan vi presenterar fakta lite mer åt dem. Så här ser verkligheten ut, så här funkar det, så här presterar de. Och så får de dra sina slutsatser och det skulle man ju så klart kunna hjälpa dem med det liksom, men det är ingenting som är i nuläget.

Jacobson, Adam - STA: Men då är det feedback från era kunders kunder.

X : Ja.

Jacobson, Adam - STA: Hur ser det ut när ni liksom analyserar? Era kunders feedback gentemot er för jag antar att ni har ganska många kunder liksom.

X : Ja, exakt. I vår systemmiljö är det inte så mycket användarinteraktion, utan det är mer system till system. Så vi har inte så många användare som interagerar direkt med vårt system, men för de som är inne försöker vi naturligtvis att underlätta deras arbete. Vi har utforskat olika lösningar för att till exempel generera innehåll baserat på den befintliga datan och skapa ännu mer innehåll. Men vi har också tittat på att implementera assistenter för att stödja dem i att utföra olika uppgifter, vilket kan vara en av våra prioriteringar. Inom ett av våra mål strävar vi efter att erbjuda en produkt som är användarvänlig, även om den hanterar komplexa processer som att generera dataskåp och liknande. Därför ser vi att assistenter är ett jättebra verktyg för att göra användarupplevelsen så smidig som möjligt.

Jacobson, Adam - STA: Då är vår fråga idag om ni har övervägt att använda assistenter som GitHub Copilot?

X : Nej, det har vi faktiskt inte gjort. Vi fokuserar mer på att använda systemet för att guida användarna och minimera antalet klick samt göra det tydligt för dem hur de ska navigera i systemet. Även om vi strävar efter att göra det så användarvänligt som möjligt, har vi ännu inte utforskat användningen av assistenter för analysändamål.

Jacobson, Adam - STA: Nästa fråga, har ni någonsin använt en generativ AI som till exempel ChatGPT?

X : Det är något vi har undersökt noggrant och det är för att skapa innehåll. Om man tar en produkt till exempel som man ska sälja online så har den vissa attribut som beskriver den. Och det vi har tittat på där är ju hur man kan anpassa en produkt genom olika produktbeskrivningar för olika målgrupper. Så om någon av er besöker en e-handelssajt och tittar på en telefon då skulle man kunna ta fram texter och budskap som är riktade specifikt till er, åldersgruppen ni tillhör och er tekniska kunskapsnivå, medan om en äldre person skulle titta på samma produkt så kanske man vill ha ett annat budskap som är lättare att förstå, med mer grundläggande information om funktioner och användningsområden.

X : Så det är verkligen att anpassa budskapet för olika målgrupper för att maximera försäljningen. Idag tar det dock mycket tid att skriva olika texter för olika målgrupper, så här kan AI spela en viktig roll. Det är faktiskt något som är väldigt intressant och...

Jacobson, Adam - STA: Det påminner om intervjun vi hade igår där de berättade om hur de skapade en chattbot med hjälp av Chutney och sedan använde Chutney för att analysera köpbesluten som chattboten gjorde. Så det var lite ironiskt, liksom. Gör ni något liknande det du beskrev? Att...

Jacobson, Adam - STA: Att generera data åt era kunders kunder.

X : Ja.

Jacobson, Adam - STA: Använder ni en AI själva eller använder era kunder. Är det analysarbetet själva generativa av?

X : Nej, det är deras... de gör ju... de ser ju...

X : Det man kan titta på här är ju konvertering, alltså hur mycket mer säljer vi när vi har olika produktbeskrivningar istället för en generisk beskrivning. Men det är ingen AI-analys på resultatet av det genererade innehållet. Däremot kan man titta på försäljningsciffror före och efter för att se...

X : Om vi säljer mer av den produkten. Så den typen av analys görs, men det är ingen AI-analys på det genererade innehållet. Man kan göra A/B-tester med produkterna för att se om de säljer...

Jacobson, Adam - STA: Men gör ni det i dagens läge?

Jacobson, Adam - STA: Nej nej.

X : Nej, vi gör inte det, men våra kunder gör det så, men vi gör inte det åt dem.

Jacobson, Adam - STA: Nej.

Tony: OK.

Jacobson, Adam - STA: Är det bästa ni?

Tony: Då var ju närstående. Frågan var ju hur ni använder AI för att analysera datan, men det är lite mer av era kunder som gör det. Så om vi går vidare till nästa fråga så nämnde du ju tidigare att ni jobbade med business intelligence-verktyg för att presentera datan. Är det någon AI-komponent med i den här visuella analysen?

X : Nej, vi har inte det. Jag skulle säga att BI och AI är lite besläktade, men om man jämför dem så är BI mer exakt, du vet exakt vilken data du har och vad du vill veta. Med AI går det snabbare och är mer flexibelt...

X : Men vi arbetar mycket med försäljningsanalyser och lagersaldo. Men där skulle man kunna använda AI för att...

X : Ge förslag på produkter som ska ingå i en kampanj baserat på vissa kriterier. Idag tittar man på lagersaldo och omsättning, men med AI skulle man kunna vara mer kreativ och få förslag på produkter utifrån olika

X Vad vad man ska sälja för en viss liksom kampanj till exempel eller att göra.

X : Och ge den lite olika kriterier och bara ge mig ett förslag, så att...

Jacobson, Adam - STA: Du, förlåt.

X : Nej, men det här med att kombinera AI och BI, det är inget jag har sett en...

Jacobson, Adam - STA: Ja.

Tony: Nej.

Jacobson, Adam - STA: Det som arkitekt, jag tänker säga något alls, eBay.

X : Ja, jag vet inte. Jag gör inte det, men jag sitter ju med våra kunder och hjälper dem. Jag hjälper bland annat nu så sitter jag och hjälper 1/1 av våra kunder. De håller på och ska ha BI-rapporter och vi ser till att de har rätt data som behövs för det. Så jag sitter inte och gör några BI-analyser på det sättet.

Jacobson, Adam - STA: Nästa fråga blir då om du ser dig själv som AI som besläktad till en viss gräns. Men om du skulle se lite mer renodlad AI, vilka fördelar skulle du då säga med det? I en analys behöver inte vara BI, men...

Jacobson, Adam - STA: I analysen av data, liksom.

Jacobson, Adam - STA: Hanteras de stora datamängder.

X : Jag ser fördelar med AI, det är alltså, men så här, om man jämför med BI så är BI ju tungrovt, det går långsamt. Det är en nattliga körningar och du måste veta exakt vad du ska ha, så det är väldigt precist, medan med AI så går det ju snabbare, du kan vara mycket mer flexibel så att kunna leva så där ser jag när det gäller just att jobba med stora datamängder och...

X : Det går rätt lätt att få fram resultat och testa, så det är ju mycket mer lättviktigt. Sen är ju...

X : Nackdelen som jag ser, det är ju att...

Jacobson, Adam - STA: Ja.

X : För någonting liksom eller ge mig en kampanj utifrån de som har lägst försäljning så ska vi ha en reamakpanj med dem, liksom. Det är ju väldigt exakt om du frågar en AI samma sak, du vet inte exakt vad du får, så det är väl lite där skillnaden är liksom.

Tony: OK, hur ser du på utmaningarna och riskerna med reglementerad AI inom dessa analysprocesser?

X : Utmaningarna är ju...

X : Det är just det här med att veta om det är rätt, alltså om jag får ett förslag, det är ju...

X : Om vi tittar på den generativa AI som vi har haft, så genererar den upp texter, liksom. Jag kan gå in och verifiera och göra stickprov, men du vet ju inte att allting blir bra, liksom, och det är ju samma sak om du får ett resultat. Hur vet du att det är rätt, och speciellt om man automatiserar?

Tony: Tack.

X : Ja, det kan ju bli väldigt, väldigt fel och vi har exempel på partners som vi har samarbetat med som har lagt ner vissa initiativ för att det har varit alldeles för oprecist, så det är väl...

X : Det tror jag är väl den stora utmaningen i det här läget.

Jacobson, Adam - STA: Nästa fråga handlar om säkerhet i relation till detta. Jag antar att det är också en del av säkerhetsbekymren, då är något annat du vill peka på kopplat till säkerhet.

X : Ja, säkerhet, det är ju...

X : Eller om man säger så här, om någonting går fel, vem ser felet alltså?

X : Det är ju det klassiska med självkörande bilar. Rent tekniskt så går det att göra det där IAI som kör varsin bil och de har ett regelverk som säger hur de ska bete sig, men om det sker en olycka, vems är felet? Är det personen som sitter i bilen? Är det biltillverkaren? Är det så där är ju...

X : Den delen är ju lite...

Jacobson, Adam - STA: Påstås kusten.

Tony: Nu vet jag inte om vi ska ha med den etiska aspekten.

Jacobson, Adam - STA: Jo, men det tycker jag vi kan fråga.

Tony: Ja, OK, hur ser dina...

X : Nu är det ju också väldigt intressant för att när det gäller det finns ju olika delar inom AI. Det här är ansiktigenkänning och det är ju det där som de gör i Kina. De liksom kartlägger alla människor i hela deras liv liksom. Nu är det ju en stat som gör det, men om någon gör det utan att man vet om det samtidigt så samlar ju in väldigt mycket data. Vi har ju...

X : Jag gillar, även om de inte är självkörande idag, så bilar samlar på sig väldigt mycket data som skickas tillbaka. Hur vet man?

X : Hur vet man hur den datan används? Det är ju samlar ju väldigt mycket. Vi har ju också det här...

X : Folk har ju släppt in det i hemmet. Vi har sådana här olika assistenter, Alexa, Cortana och Siri och så, som också samlar massa personlig data. Var används den datan till, hur tar den vägen? Det är ju väldigt mycket.

X : Och det? Det är svårt att reglera, jag liksom...

X : Sätt exempel på liksom någon...

X : Handdammsugare som någon av dem hade kollat, som hade någon uppkoppling och då visade det sig att den skickar data till Kina, liksom. Så det är ju där är väldigt den delen är så man har svårt att kontrollera vad datan tar vägen, hur den används, och det är ju...

X : Så svårt som konsument att skydda sig mot, men det är ju...

X : Och det? Ja, där är det stora etiska liksom.

Jacobson, Adam - STA: Kan du peka på exempel inom er verksamhet och era kunders...

Jacobson, Adam - STA: Kopplat till etik då och kanske ja, analys.

X : Ja, vi har ju vissa...

X : I vissa för vissa kunder så har ju vi kunder och ibland är de och då gör ju vi. Vi kan ju göra som så riktade budskap mot en känd kund och kunden loggar in och så vet vi ju vem var den kunden har köpt innan och då kan man ge rekommendationer på...

X : På vad den borde både köpa och det är ju det är ju sånt som är...

X : Det är ju avtalat via vi äger DP så att kunderna har ju signat en weaver där de godkänner att deras data används. Men frågan är om de verkligen är medvetna om vad det innebär. Du går med i en kundklubb och du får 20% på ditt köp och du tycker är jättebra så har du den här kryssrutan jag godkänner jag är visst och så har de inte läst vad det innebär liksom. Det är ju det finstilla och det gör ju att man kan ge riktade och det är ju...

X : Allt, och där handlar det ju om att för våra kunder handlar det ju om att tjäna mer pengar och sälja mer grejer, så att men jag vet inte om kunderna som skriver på egentligen är medvetna om det själva.

Jacobson, Adam - STA: Men det är de är det mer algoritm baserat på alltså riktig kod eller är det mer AI baserat?

Jacobson, Adam - STA: Just där sista du nämner.

X : Alltså det här med rekommendationer är ju, det är ju...

X : Att ta fram rekommendationerna det är ju nu använder inte vi AI till det. Vi har inte kommit dit ännu, men det är ju absolut något som vi har i vår water att kunna ta fram produktrekommendationer utifrån. Det tror jag kan bli rätt så häftigt för idag när man ger rekommendationer så är det ju...

X : Det är ju, det är ju bara algoritm baserat och det är liksom du har köpt det här och då föreslår den, då ska du köpa det här också. För det köpte de andra liksom, så det är ju väldigt mycket sånt och det. Det blev väldigt fel. Om du köpte ett par byxor så får du reklam för ännu mer byxor men då har du redan köpt byxor liksom så att jag tycker det det det kan slå väldigt fel, men då skulle man ju kunna använda AI för lite mer coola rekommendationer för det liksom, men det är ingenting som vi gör i nuläget men det är ju absolut en...

X : Nej, nej, jag har inte stött på det, men samtidigt så är det ju så, men en AI är ju inte, den är ju inte bättre än den träningen som den har fått liksom, så det är ju det här är ju en begränsning liksom. Sen är det ju såklart problematiskt att om du har mina frågor du en människa så nu finns det ju såklart. Vissa vissa män som aldrig skulle erkänna att de inte kunde något liksom att hitta på ett svar men vanlig en vanlig person skulle ju säga liksom att nej, men det jag vet inte det får jag kolla upp och återkom.

X : Liksom mellan en AI skulle ju kunna hitta på någonting och...

Tony: Exakt.

X : Har man inte källkritik så kan det ju bli att man tar det för liksom som en sanning, så där är det ju också så vet man att man interagerar. Det finns ju väldigt mycket å i assistent och där folk interagerar. De tror att det är en människa där de pratar med liksom om de hittar på svar eller ger direkt felaktiga svar.

X : Det kan ju vara liksom lura lura vem som helst liksom om man inte om man inte vet att det är en robot och den inte liksom så det kan ju bli en väldigt konstig situation liksom så att.

Jacobson, Adam - STA: Ja.

X : Precis för det är oftast det. Och vad ska man lura någon så ska man bara vara självsäker och säga att ja, kan det här är det här är rätt liksom? Det vet, jag sa att så där, där är ju absolut 1, 1, 1 problem som man.

X : Som man liksom kan hamna i det.

Jacobson, Adam - STA: Vi pratade igår också.

Tony: Ja.

Jacobson, Adam - STA: Det som det har varit inne på detta idag, men ett företag planerar då att...

Jacobson, Adam - STA: ...göra ett business decision som baseras på data som är felaktig.

Jacobson, Adam - STA: Vems fel är det?

Jacobson, Adam - STA: Omöjligt att svara på, men.

Tony: Vad vi kan.

X : Ja, men det det, det blir ju precis, det blir ju det och det, det är ju ett av.

X : Problemen som man måste måste liksom adressera med med det här liksom, och det är ju.

X : Som jag nämnde innan som det här med självkörande bilar. Det är ju en av anledningarna till att de inte har släppt ut dem för att man kan inte lösa ansvarsfrågan liksom. Och det är ju.

X : Det där är ju och jag menar AE är ju osäker så jag tror det är väl också en anledning till att.

X : Men i vissa fall inte vill förlita sig på det på det helt och hållet liksom, men.

X : Det är.

X : Den behöver man behöver verifiera det liksom. Det är så är det liksom och.

X : Webb webb så är fallet. Ja, det är det väl.

Appendix E. Interview 3

Interview 3 (R3), Company X representative, Length of interview: 40 minutes, X = respondent.

Tony: Yes, då kör vi. Kan du ange ditt namn och din nuvarande position och hur länge du har jobbat i den rollen?

X: Jag är en data engineer. Men jag jobbar främst med front-end inom data engineering. Mycket med dataanalys generellt. Och det har jag jobbat med i ungefär ett år.

Tony: Grymt. Adam, vill du köra nästa?

Adam: Skulle du kunna dela med dig av några exempel på ditt arbete där du då jobbar med dataanalys?

X: Ja, jag har arbetat med datanys genom att ta data hämtningar från API:er och gjort om det till statistik i olika analysplattformar. Samt har jag också använt business data och marknadsföringsdata och jobbat med dataanalys inom det.

Tony: Okej, intressant. Hur skulle du säga att du jobbar med kvantitativ dataanalys?

X: Kvantitativ dataanalys, jag jobbar väldigt mycket med nummer generellt. Väldigt mycket med olika försäljningsnummer och produkter. Hur mycket de har sålt, summan de har sålt för, mycket kring det. Mikrokopier kring det.

Tony: Och hur ser det ut på den kvalitativa delen då, där du jobbar med data som inte är någon annan?

X: På den kvalitativa delen kan man jobba med diverse grejer, men ett exempel kan vara, som jag har gjort då, är metadata. Jag har jobbat med metadata och kollat hur saker och ting sker generellt inom en databas. Och gjort en datanals på det eller i en verksamhet gjort en datanals på det

Adam: Alright, Nästa fråga handlar om i vilken utsträckning du använder AI-baserade verktyg i din analys då Jag tänkte först börja med om du använder en AI-assistent och det skulle kunna vara GitHub Pilot då som är för kodning i sig nu när du använder en AI-assistent för analysen i sig

X: Ja, Jag använder ingen AI som github-copilot, men generellt när det kommer till dataanalys då tar jag absolut hjälp av AI för att göra saker och ting bättre.

Adam: Kan du ange vilka olika verktyg som du använder?

X: Ja, jag använder mycket av chat-gpt för att förenkla för mig när det kommer till dataanalys generellt.

Adam: Och då är det huvudsakligen checkability du använder?

X: Det är huvudsakligen checkability.

Tony: Vad konkret använder du den i ditt arbete för?

X: Väldigt mycket med när det kommer till verklighet, exempelvis förklara verklighet, men också när det kommer till datakorrelering, när det kommer till hur jag ska Hur ska jag få fram en viss typ av mätning inom data. Så kan det hjälpa till med ibland.

Adam: Så du använder det för att underlätta din egen analys och inte nödvändigtvis analysera åt dig?

X: Nej, precis. Det är mitt eget tankesätt.

Adam: Okej. Använder du då, om du är konsult och använder era kunders data in i GPT?

X: Nej, jag skulle aldrig använda någon sorts data från varken vår verksamhet eller någon annans verksamhet inne i GPT framförallt så det är först och främst det är regelbrott det är därför jag aldrig ska göra det.

Adam: Och då är det liksom på personliga grunder eller är det ditt företag?

X: Det är företaget. Ibland hade det varit underlättat om man hade kunnat använda den här datan i GPT. Vissa saker kanske hade gått snabbare. Men i och med att det finns så mycket säkerhetsrisk med det så kan vi inte göra det.

Adam: Jag förstår. Frågan om big data, och du var inne lite på det när du pratade om kvantitativ analys. Jag kan tänka mig att försäljningssiffrorna till exempel är väldigt mycket att analysera. Men som du tidigare sagt så har du inte använt AI på samma sätt, en AI-assistent till big data. Men om du hypotetiskt i framtiden säger hur AI skulle kunna hjälpa sådana typer av arbete. Det vill säga väldigt, väldigt stora mängder data.

X: Ja, det gör det absolut. Jag tror att AI kan på ett väldigt bra sätt, i framtiden skulle jag kunna tänka mig, kategorisera in big data och strukturera det väldigt bra så att man kan analysera det till att börja med. Analyseringsbiten, det vet jag inte om de kommer att fixa det lika väl. Men struktureringen av datan tror jag absolut att man skulle kunna fixa.

Adam: Vad ser du för nackdelar med ett sådant verktyg?

X: Egentligen ett sådant nackdel med ett sådant verktyg är väl att Om det inte är privat och om man inte har koll på det tillräckligt bra så kan potentiell data läcka ut. Det får ju inte hända.

Adam: Vill du fortsätta, Tony?

Tony: Absolut. Hur mycket har du använt dig av business intelligence-verktyg såsom Power BI eller liknande i ditt analysarbete?

X: Det är huvudsakligen Power BI. Man kan gärna säga att Qlik används också för den delen, men huvudsakligen så använder jag Power BI.

Tony: Finns det AI inslag i Power BI?

X: Nej, alltså egentligen inte. När jag just strukturerar datan innan jag ställer in den i Power BI då kan jag ibland ta hjälp av AI såsom då ChatGPT. Eller så, när datan väl är inne och den är strukturerad då kanske jag behöver tips om olika mätvärden. AI just nu är någonting att bolla med, som en kollega att bolla med nästan, fast det kanske är lite snabbare svar.

Tony: Okej, absolut. Ser du några möjligheter inom just business intelligence tools att AI kan underlätta eller förbättra vissa arbetsprocesser?

X: Det tror jag definitivt. Jag tror att AI kan jag tror att det kommer att ta väldigt lång tid innan AI kommer att kunna göra det helt på egen hand. Vi kan komma till det sen. Men jag tror absolut att AI kan bli en större del av det. Och hjälpa till mer och mer med dataanalys och göra det mer effektivt.

Tony: All right. Det fanns. Vi har redan varit inne på många av de här delarna men om du pratade innan om fördel med AI till big data så kanske denna frågan är fördel utifrån dina erfarenheter liksom konkret det som du har såklart kan du svara på mycket av detta men är det något du skulle vilja tillägga om du kollar konkret på dina arbets det du gör om dagarna eller något annat

X: Du kan tänka på något som AI kan underlätta med, men det är mer hjälp, mer förståelse och att man inte kan lita på att AI gör det ordentligt än, tycker jag. Att AI inte är tillräckligt bra än. Men absolut, det hade varit väldigt skönt om man hade kunnat förenkla vissa delar som kanske skulle ta lite längre tid ibland med AI. Sådana delar kan anses lite repetitiva. Sådant hade varit väldigt skönt om man kunde ha kommande AI till.

Adam: Vi har en fråga om AI relaterat till säkerhet inom analys och det har du också varit inne på med dataintegriteten. Är det något annat du vill peka på där kopplat till säkerhet och AI?

X: Egentligen är det väl bara samma sak som innan. Just nu kan man inte använda egen data, företagsdata, kund data generellt, i AI-verktyg som inte är privatiserade. För det kommer bara inte att gå. Det är inte lagligt heller för den delen. Ingen kommer kanske att se det på ett lite längre tag, men om det blir en breach så kan det bli väldigt illa.

Adam: Då använder ju de här GPT och liknande, de använder ju datan för att träna sina egna robotar och det är liksom men planera liksom att vi har ett scenario där företagen tränar sina liksom till att utföra helt självmant och inte använder data från liksom användarna skulle det liksom vara mer okej då?

X: Jo men det går absolut bra det fungerar det kan absolut fungera Men samtidigt så, jag för mig att man får använda företagsdata rätt så brett. Men så finns det också viss företagsdata som är lite mer konfidentiell. Vilket kan bli ett problem. Samtidigt som också att om ett företag skulle använda... Sin företagsdata som kundernas data, kan det bli problem om kunden vill ha bort sin data från det företaget, för det är ju en regel som vi måste följa också. Då kan det bli problem för då är AI redan inkorporerad med det och kan bli svårt att ta bort.

Adam: Hur ser ditt företag på att man får använda ChatGPT och andra AI-verktyg? Har de någon specifik policy på det?

X: Policyn är just att vi får inte använda data från kunder. Vi får inte använda data från företaget. Vi får använda dummy-data, egengjorda data. Som inte spelar någon roll helt enkelt, som inte betyder någonting egentligen. Men det uppmanat är det kanske inte hundra procent, det är 50-50, vissa gillar det, andra gillar det mindre. Men det är absolut helt okej att använda det. En sak med det AI som man måste tänka på är att den har inte alltid rätt.

Tony: Exakt. Vi har en fråga angående detta, om du vill ta den Adam?

Adam: Absolut, AI-hallucination kallas begreppet det vill säga att AI i sig liksom förmedlar ett budskap som låter 100% självsäkert, men som kan vara helt fel och då tänker jag liksom, vad säger du vad är problemet med det?

X: Jo, men hela saken med datan är att datan ska ju vara riktig. Datan ska ju betyda någonting. Den ska ju ge korrekt information. Om den inte gör det, då är det ju inget håll. Det är lika gärna att slänga rapporterna. Då är det bortkastat arbete. Bortkastad tid. Du tror absolut att det kan bli bättre. Men samtidigt det som jag ville peka på tidigare det var just att AI kommer kunna hjälpa oss mer och mer men det kommer alltid finnas en risk med att AI inte alls kommer kunna validera data och

information lite bra som människor kan för en AI just nu Kommer inte kunna på väldigt länge och om någonsin förstå hela, alltså helheten.

Adam: Om ett företag som ni exempelvis gör ett affärsbeslut baserat på AI-data Och dataanalysen med AI och det sker något problem, det uppstår en AI-hallucination. Vem anser du ska bära ansvaret för det?

X: Den som egentligen, det beror på vem som byggde det till att börja med. Godkände det för den delen också. Det är väl en kombination. Egentligen är det väl någon chef som egentligen har godkänt detta. Men om en chef inte hade någon aning om det, då blir det lite mer komplext. Då blir det ju lite mer själva den anställdas fel. Medan chefen har ju också ansvar, men det skulle nog aldrig hända riktigt. Det är svårt att säga, det är svårt att svara på.

Adam: Precis, jag förstår det. Det är inte liksom ett scenario. Det kan säkert hända någon gång. Det är så pass dumt att göra det och ha så mycket tillit på just nu. Man måste oavsett vad man validerar. Det är alltid en grej som man måste göra.

Adam:Om vi har ett scenario där ni konsulterar åt ett företag och sen så använder ni en AI-mjukvara för att analysera deras data så ni blir mellanhand mellan mjukvaran och företaget ni konsulterar åt om du förstår vad jag menar.Och ni gör ett beslut i analysen som er kund då investerar väldigt mycket pengar i då har man ju ett skede där ni eventuellt skäller på Ai:n medans kunden skäller på er medans de också kanske borde ta lite ansvar i att tänka innan man investerar så pass mycket pengar om du förstår scenariot

X: Jag förstår snart, men det var på vårt initiativ att använda AI eller var kunden också med på det?

Adam: Nej, det var på ert initiativ.

X: Ja men egentligen oavsett är det vårt fel, för att man måste som konsult och som analytiker och data engineer så måste man ha koll, man måste ta ansvar. Man är ju ens kompetens, ägarnas kapital och ens resurs. Man måste det helt enkelt, det är ens eget fel i så fall. Det är huvudsakligen ert fel i detta fallet, eller mellanhandlets fel, jag förstår. Absolut, det är konsultens fel. Det är rätt vad man känner.

Tony: Då tänkte vi egentligen bara fråga om det finns någonting annat i diskussionen som du skulle vilja lägga till som du tycker vi har missat? Någon diskussion som för oss inom en verksamhet kanske kopplat till AI?

X: Ja, kan vi bara tillägga det? Absolut, jag tror att AI kommer användas mer och mer. Användas som hjälpmedel mer och mer, det ser vi just nu det är en det är absolut någonting som kommer att bli mer populärt hela tiden men samtidigt så har jag svårt att se att AI skulle ta över för att det behövs generell AI intelligens för det och inte bara narrow AI intelligens Som då en generativ AI är exempelvis.

Adam: Du har pratat om ditt företag som jobbar på deras ton. Det är kopplat till en generativ AI och deras syn på det. Men har de en vision på fem års sikt? Till just AI. Är det något som pratas om konferenser och det likt? AI pratas absolut konstant. Men det gör det ju väldigt många verksamheter generellt skulle jag vilja säga.

X: Konsultserier. Men det är också väldigt mycket på mikronivå. I och med att man hittar sina egna arbetssätt också. Man har väldigt mycket frihet. Det är också så att man kan vara väldigt effektiv och lära sig från varandra. Men AI är absolut något som är populärt bland Folk i den kulturbranschen generellt.

Adam: Har du några kollegor som använder verktyg som inte du använder?

X: Inte vad jag vet. Kanske Microsofts AI. Eller kanske Googles AI som jag inte tycker kan lika bra. Amazons AI kanske också. Det beror på. Det är väldigt brett. Många använder olika verktyg. Det beror också på vad man jobbar med. Det är väldigt brett, helt enkelt.

Adam: Men generativ AI, skulle du säga att det är en generell trend, kanske?

X: Det är absolut en trend. Det är det verkligen. Jag ser det väl mer på ungdomar generellt och jag hör mer på ungdomar generellt som använder. Gärna att vi var det och skämt mycket i deras arbete men det är absolut aldrig så att det är 100% eller att de gör det för mycket skulle inte jag säga heller.

Appendix F. Interview 4

Interview 4 (R4), Skatteverket representative, Length of interview: 45 minutes, X = respondent.

Tony: Så om vi börjar med första frågan skulle du kunna ange ditt namn och din nuvarande position och hur länge du jobbat i rollen?

X: Jag jobbar på Skatteverket som typ data scientist och jag har jobbat på Skatteverket i lite mer än år. Jag började på Skatteverket, så att jag har varit på Skatteverket ganska ganska lång tid.

Tony: Kan du dela med och det är några exempel på ditt jobb när du jobbar med dataanalys.

X: Ja, det kan vi väl göra en mängd saker egentligen. Man kan väl säga att. Jag börjar egentligen min min karriär, på skatteverket då som som mer av en handläggare inom mervärdesskatt och arbetsgivaravgifter. Men kom ganska snabbt in på det vi inom Skatteverket kallar då för. Urval eller eller riskhanteringen om och urval för oss handlar det egentligen om att. Generellt kan man säga att alla de handlingar eller deklaringar eller ansökningar som vi tar emot behöver vi hantera på ett annat sätt och kopplat till de här ansökningarna, handlingarna med mera så finns det ju också olika typer av risker. Risker att personer gör att denna felaktiga avdrag eller har fyllt i ansökningarna eller angivet felaktiga uppgifter.

X: Här gör vi olika typer av. Vad ska vi säga filter, regler och annat för att identifiera då deklaringar, handlingar med en högre risk för fel?

X: Och jag kom ganska snabbt i kontakt med den delen av Skatteverket och den delen av vår verksamhet. Om man säger det någon mot vår kontrollverksamhet eller egentligen det som. Identifierar deklaringar eller handlingar med högre risk för fel. Så i den så har jag jobbat egentligen med data och analys av data under ganska lång tid. Men det var väl då jag först kom i kontakt med det och sen har det egentligen fortsatt i olika typer av konstellationer.

X: Jag jobbade ju mycket med de här uppgifterna i ett antal år och sen blev det att man. Centraliserade mycket av den här funktionen på Skatteverket och då följer med till dåvarande huvudkontoret och där blev jag placerad på det som har analysenheten på på huvudkontoret då. Och där blev det också en mängd andra typer av dataanalyser och bearbetningar kopplat till allt från effektutvärderingar till andra typer av mera aggregerade riskanalysen mera möjliga så att data och analys av data skulle ändå säga att jag har jobbat med i stort sett hela mitt yrkesverksamma liv så så lite mer än år skulle jag säga att jobba med olika typer.

Tony: OK. Du nämner det ju att jag har jobbat mycket med själva datan och analysen av olika ärenden skulle kunna ge konkreta exempel på ärenden där det varit mer av en kvantitativ karaktär på datan du analyserat?

Här är det mycket av en mix av kvantitativ och kvalitativ analys. Mycket av de kvalitativa delarna sker i början av arbetet medan det går över mer åt kvantitativ analys i senare delen av ett utvecklingsprojekt (givetvis beroende på vad det är för typ av problem)

X: Jag tappar det lite grann, men var det inne på kvalitativt kontra kvantitativt?

Tony: Exakt.

X: Ja. Ja, men i mångt i mångt och mycket så kan man väl säga det? Det är lite grann beroende på vad man lägger i begreppen, men Skatteverket jobbar ju med en. Inom det som är det som är benämnde som urval då och det vi jobbar med det inom vår kontrollverksamhet så använder vi ju. En mix av olika tekniker och metoder för att identifiera då deklaringen. Handlingarna saknar med högre risk för fel är en av de metoderna är ju egentligen eller en del av metoden är kanske mer erfarenhet drivna.

Tony: OK.

X: Och och i den aspekten kan man väl säga att det är till viss del ligger mer åt det kvalitativa hållet, men där vi kanske bedriver mycket av av. Ja, man ser att vi sätter samman olika typer av erfarenhet, drivna modeller, alltså utifrån expert eller domän kunskap inom myndigheten. Så det kan till exempel handla om att vi genomför olika typer av workshops vi diskuterar. Vi vill ju det även där få en om man säger det man pratar genom kvalitativa studier. Den här mättnadskänslan då egentligen att att vi vi genomför ett antal workshops och seminarier. Vi får fram olika typer av intressanta.

X: Indikatorer eller faktorer då som vi tror kan påverka den här risken för fel och sen försöker vi egentligen kanske operationalisera den kunskapen som vi samlar in då från från våra handläggare i första hand. Och det är det som gör utgör de har erfarenhet xdrive modellerna. och andra Å andra sidan så kan man ju säga att vi har också den kvantitativa delen där det är egentligen den här insamlingen av intressanta faktorer. Kanske bara gör utgör ett första steg där vi också då kanske har tillgång till historisk genomförda kontroller inom ett visst område där vi också söker att bygga, utveckla eller träna då en någon typ av statistisk modell med en AI modell om ni vill.

X: Och där kan man ju säga att där kan ju de här workshoparna vara ett sätt att att få intressanta faktorer som vi vill undersöka, det vill säga input till att göra den kvantitativa analysen. och sen så operationaliserar vi vi testade med olika typer av statistiska modeller och tränar olika typer av modeller för att då få mer om man säger

Tony: OK

X: Och sen, så operationaliserar vi. Vi testade med olika typer av Statistiska modeller och tränar olika typer av modeller för att då få mer om man säger. Mer ett kvantitativt underlag för att statistiskt bedöma vilka faktorer som har den största påverkan och vilka tillför värde, vilka tillfällen inte väder. Men vi jobbar ju ibland kan man säga. Ibland jobbar vi med enskilda metoder. Ibland jobbar vi med en mix av olika metoder och tekniker. Ja, men det är väl lite lite brett, så lite olika exemplifieringar.

Jacobson, Adam - STA: Nästa fråga handlar om AI baserade verktyg och det var ju inne lite på det. Du nämnde en maskininlärnings applikation, Har ni en extern som ni också använder eller är allting som ni skapar på eget hand så att säga?

X: Om man, om man tittar brett då på hela Skatteverket om man tittar på på ai området så har vi ju en en mängd olika tillämpningar av ai i det som idag kallas för AI. Det som idag ligger under paraplybegrepp som har hjärna någonstans. Det är en mängd olika tekniker och metoder som helhet som ligger under det, men att.

Här kan man ju säga att vi primärt utvecklar mycket själva kopplat till det som är det som ligger inom det som är riskvärderingar. Området kontrollen som jag berättade lite grann om tidigare. Tittar man sen på delar det vi har leverantörsled lösningar och egentligen mer färdiga programvaror som vi arbetar i. Då är det ju mer till exempel våran chatbots som byggs på en leverantörsled lösning. En leverantörs plattform där vi egentligen står för träningsdata och inputen. Men vi tränar på en plattform helt enkelt. Och sen har vi egentligen andra delar där vi också till stor del jobbar med. Det är ju inte så att vi utvecklar allting själva utan vi jobbar också väldigt mycket, i många fall med med öppen källkod så att vi tar in öppen källkod så vi tränar inte från grunden i vissa lägen. Vi jobbar också med språkmodeller eller större eller mindre språkmodeller. Det finns en del från från kungliga biblioteket som vi använder i vår produktion i dag till exempel.

Bland annat då så har vi ett AI system där vi klassificerar inkommande e post till Skatteverket, det vill säga att försöker ämnet bestämma och fördela ut det till rätt kompetens då för att svara på de frågor vi får in. Det här använder vi en större språkmodell från kungliga biblioteket för att underlätta vid träning. Så på det sättet kan man ju säga att att i vissa lägen utvecklar vi helt från grunden själva och tränar från grunden i andra lägen så bygger vi en del på öppen källkod och i tredje eller i tredje hand men II det tredje graden så sa vi också en del inköpta produkter eller leverantörsled lösningen.

Jacobson, Adam - STA: Använder ni GPT som öppen källkod?

X: GPT skulle jag säga jag skulle inte kanske lägga det i facket öppen källkod men men om vi säger att att om vi tittar på GPT eller GT som som GPT skulle jag säga, jag skulle inte kanske lägga det o facket öppen källkod men men om vi säger att att om vi tittar på GPT eller GT , som möjligt för oss att använda som svensk myndighet så skulle jag säga att vi är i ett utforskande stadie just nu. Vi använder inte generellt och vi använder inte generellt i hela vår verksamhet. Men vi har ett antal pilotfall igång där vi undersöker möjligheterna till att använda både egentligen den typen av externa generat.

Vi var i modeller, men vi tittar också på andra delar där vi undersöker möjligheterna att det finns ju en hel del utveckling inom det här området nu. Bland annat så finns ju en

del modeller som mistral med mera som som är möjliga att kanske ta in och hostar lokalt på Skatteverket. Och på det sättet kan man ju säga, då kan vi dra nytta av en en större språkmodell utan att exponera vår egen data utanför vår eget. Skyddade nätverk så att säga eller det som är påskyndade miljö och det är väl en av de viktiga bitarna är just nu.

Jacobson, Adam - STA: Det är ett dilemma att liksom analysera användare AB för att analysera riktig data eller använda AI för att underrätta sin analys, men inputen är inte riktig data. Hur ser ni på det? Hur används till huvudsakligen om du förstår min fråga.

X: Nej inte riktigt.

Jacobson, Adam - STA: Nej, alltså. Jag tänkte liksom vi hade en intervju igår där liksom de vågade använda riktig data in i GPT för att i och med att de tränar på den datan och det är integritetskränkande och så vidare. Men jag tänker liksom om ni har AI baserade verktyg inom er organisation så kommer inte det, så tränar ju ni dem utifrån egen data liksom.

X: nej Man kan väl säga så här då att i den mån så kan man ju säga att. Vi har ju diskuterat mycket. Vi har till exempel en en översättningstjänst internt inom Skatteverket. Varför har vi det? Varför använder vi inte Google translate istället? Ja, men det beror ju på att vi tillämpning vi användningen så exponerar du också den typen av datainformation som du vill översätta och vi har väldigt många behov då, både när det gäller kanske chattgrupper t om man säger att ja, men sammanfattar den här texten åt mig eller översatt för den delen också så kommer man ganska snabbt till ett läge att ja, men vad är det för texter vi vill översätta som myndighet? Jag ofta så har det att göra med var min.

X utövning, Slutövning det innehåller sekretessbelagda uppgifter skulle kunna innehålla i väldigt många fall personuppgifter och andra delar och där vill vi ju index exponera den typen av data utanför vårt eget nät och det gör att vi är mycket, mycket mer intresserad av kanske att ta in saker som vi kan hålla i med våra väggar än att exponera för mycket information på utsidan. Däremot om man säger att det är det mer allmänna dokument eller säga att jag skulle skriva en en AI strategi skulle jag då kunna bolla eller sammanfatta eller diskutera till exempel med.

Att GT, ja, där skulle jag kanske ha mycket större friheter så att det är lite grann. Det som vi också utforskar i de här användningsfallet som vi tittar på i, vilka vilken omfattning kan vi använda oss av externa tjänster? I vilken omfattning behöver vi interna tjänster? Ja, vad fungerar bäst? Och det beror ju som sagt mycket från från fall till fall eller tillräckligt till.

Jacobson, Adam - STA: Nästa fråga handlar om AI baserade verktyg kopplat till big data, vilket om jag förstått rätt har du nämnt tidigare. Samt, Har du en tydlig bild på skillnaden mellan vad big data och är?

X: Nej ja, alltså, jag vet inte big data, detta begreppet tycker jag är är ibland svårt att förhålla oss till, för det är också ett sådant där begrepp som är ganska brett och används i en mängd olika betydelser. Så så det är väl vad om jag börjar med en motfråga vad vad det innebär big data i ett kontext eller som ni ser det?

Jacobson, Adam - STA: Nej, men precis det är väl inte det finns kanske en gråzon där? Antingen är det data eller big data. Det finns liksom inget mellanting. Det är lite intressant att säga det, för vi hade en ett annat intervjuobjekt som sa likadant liksom. Att dom inte läser stor vikt vid själva begreppet byggdelar, utan alltid data, mer eller mindre.

X: Ja, jag hörde någon. Jag hörde någon definition för några år sen som som konstaterade att big data är det när du inte har beräkningskraft som är tillräcklig för att hantera det, då blir det stordata, annars är det bara data.

Jacobson, Adam - STA: OK, men har ni något exempel på liknande data som inte kan kan beräknas så att säga?

X: Nej, det skulle jag väl inte säga i nuläget, men men tittar man tillbaka lite grann så har vi ju haft utmaningar vid vi tittade på möjligheterna för jag tror att det var år sedan vi nämnde det här med med maskinöversättning, små problem och som vi har tittat på och där kan man ju säga att där ja vi vi hade vi hade då en miljö där vi hade en MGPU processor som kunde jobba med. Och den stod och gick nästan på % i dygn och vi har fortfarande inte färdig med träningen av de översättningsfel vi håller på med då så så där gick vi lite grann i taket med. Men idag jobbar vi lite i en annan miljö där vi har ett lite större kluster som vi kan jobba med träning i och vi har också valt till vissa lägen, kanske andra vägar som det gäller maskinöversättning då och så har vi ju jobbat.

Med källkod och öppna modeller egentligen. De utvecklade av ett universitet i Helsingfors, har jobbat mycket med med översättning. Översättnings delar, så det är vad vi använder i vår miljö idag.

Tony: Använder ni använder er av business Intelligence verktyg såsom power BI under analys processer.

X: Det är du tänker analysprocessen på vägen fram eller utvecklingsprocessen då kopplat till en AI lösning eller tänker du generellt inom verksamheten?

Tony: Generellt inom verksamheten

X: Man kan väl säga att att Skatteverket idag använder verktyg för det som är BI och och rapportering. Dels är det en en lösning från IBM som heter cognos som vi använder och sen så använder vi också oss av Qlik och QlikSense, och där är olika programvaror för att egentligen visa mycket av den data och den information som har använts ,och i Skatteverkets datalager och det används ju utifrån allting från.

X: Analys så att säga, vi har olika typer av analyser till uppföljning av vår verksamhet för ledning, styrning.

Tony: OK och har dessa verktygen något slags av AI implementerat?

X: Jag tror det. Jag är lite ute på lite tunnare is här men men jag vet att det har förts diskussioner om att det finns vissa funktioner inom till exempel Qlik och jag vet att IBM har också en del. Jag vet att jag har varit på en del demos där man visat att det kommer. Jag kommer inte ihåg namnet, men där hade man även en röststyrd funktionalitet för att visa siffror då eller vrda data på ett visst sätt, det vill säga att man skulle kunna ställa frågor till databasen genom ett mer. Men det är ingenting som jag vet att vi har aktiverat idag och det är ingenting vi använder idag. Det är möjligtvis det närmaste vi kommer en och tror jag i Qlik så fall att du kan få förslag och rekommendationer om att ja, men här kanske börja använda den här typen

av diagram eller det skulle kunna visa data på det här sättet? Det tror jag.

Tony: OK.

Jacobson, Adam - STA: Nu ska man fråga lite mer generella, Vad ser du för fördelar med A inom dataanalys?

X: Det, det är ju lite grann. Det är en bred fråga och det är lite grann beroende på hur man ser det. Ska jag svara mer utifrån mig som utvecklare av så att säga eller ger alltså dataanalytiker och hur jag använder AI för att göra min dataanalys.

Jacobson, Adam - STA: Ja ja analysen.

X: Eller är det mer för vi jobbar ju. Vi jobbar ju väldigt mycket med utveckling av olika typer av AI lösningar av olika AI system för att lösa vissa behov och sånt där. Men tittar man på liksom hur jag använder AI som dataanalytiker?

Tony: Ja.

X: Mm, ja dels, så skulle jag väl säga att vi vi har och vi undersöker möjligheten att jobba med olika typer av co pilots för att hjälpa mig att skriva kod bättre mer effektivt, så det är ju ett sätt att det som dataanalytiker kan ta tillgodogöra mig så att säga och skriva bättre sökningar och så vidare mot data.

X: Delvis det jag var inne på med med våra bi verktyg finns ju möjligheten där.

Tony: ja

X: Sen kan man väl säga det är väl här någonting man kommer in i liksom olika typer av analysmetoder och tekniker. Men det är klart, AI eller inte, ibland en svår gränsdragning. Men jag kan ju använda olika typer av av absorbine learning verktyg som truster analyser eller sådana bitar för att visa hjälpa mig att hitta strukturer i data eller på något sätt göra.

Jacobson, Adam - STA: Nej, precis.

X: Transformera Ostrukturerad data till till mer strukturerad som gör att jag kan liksom ta mig ta mig an problemet eller ta mig an problemet snabbare. Så i den månen, så ja.

Jacobson, Adam - STA: Till exempel.

Tony: ja

Jacobson, Adam - STA: Till exempel copilot, sade att ni använder till en viss utsträckning, då är det en extern mjukvara då.

X: Ja, det är ju på det sättet väldigt både ja och nej skulle jag säga det. Vi tittar faktiskt på nu är att återigen som jag sa vi, vi vill ju hålla tillämpningarna inne på Skatteverket och vi tittar faktiskt mer på öppen källkod än att vi tittar på inköpt programvara från en leverantör.

X: Så även här tror jag att nu kommer jag inte ihåg, men jag tror att den sista är en av de sista modellerna som vi håller på att utvärdera och titta på. Idag är ju som det är wizard coverpro om vad direktionen vissa coacher som vi tittar på att nyttja då.

Jacobson, Adam - STA: Då tränar man i den.

X: Just nu så kan man ju säga att i första läget så tar vi ju in den färdiga modellen och vi kan använda den egentligen auto som den är. Vi förpackar den kan man säga i Skatteverkets miljö så att vi lägger ett skal runt den här så att den passar i vår IT infrastruktur, men annars så ligger modellen till en oförändrat det vi kan göra om vi vill och om vi har möjlighet till det. Det gör att också lägga till den o domän träning, det vill säga att vi anpassar den här utifrån Skatteverkets sätt att skriva.

Och då skulle vi också behöva ha tillgång till den datan som krävs för att göra

den domänen träning, det vill säga det är ytterligare ett steg. Men det är ju en möjlighet, helt klart.

Jacobson, Adam - STA: Väldigt positiv träning.

Tony: Okej, risker och utmaningar ser du med användningen av ai inom data analys?

X: Det är också en en ganska Bred fråga, jag ska bara se så jag håller reda på vart vi är någonstans. Vilken fråga är du på nu på?

Jacobson, Adam - STA:

X: Sjuan ja. Nej, men det är väl det är väl även där så att. Kan jag använda AI för att så att säga ge mig förslag eller rekommendationer baserat på data till sammansättning eller annat så. Är ju den någonting som som argumenterar eller boostar mig som dataanalytiker då i den valen. Men, men det är som sagt den är svår att svara på för att vi vi kommer ju mer inser man tänker på det är ju i termer av Vi utvecklar AI på det sättet och sen så kan man säga att vi använder AI. Här blir det ju använder AI utifrån en dataanalytiker slag.

X: det är inte det är inte där jag normalt

Jacobson, Adam - STA: Men vi kan väl kan fortsätta då?

X: Det är inte, det är inte där jag normalt är de här säger så så att jag har lite svårt att se liksom om det är liksom concerns eller utmaningar någonstans kopplat till användandet av ai inom dataanalys. Då blir det ju mer ett allmänt svar där man någonstans behöver man ju alltid värdera eller prova det man får ut från ett AI system att sätta det i kontexten och vi behöver också som alltid ha en en viss källkritik så att säga.

Jacobson, Adam - STA: Nej, precis.

X: Inte är svår. Men det här beror ju också på. Var finns källan till det här? Vem är det som har utvecklat och tagit fram systemet har jag har jag har jag förtroende för för utvecklande om det är en annan myndighet eller om det är elleverantörer.

Tony: Ja.

Jacobson, Adam - STA: Precis så nästkommande frågor beror ju både säkerhet och etik. Men det har vi också varit inne på med tanke på det här med allmänna handlingar och inte allmänna handlingar syntest delägda handlingar.

Tony: ja

Jacobson, Adam - STA: Säkerhet, jag tänker kanske med cybersäkerhet kopplat till det eller någonting. Det blir pekar på det där.

X: Nej, men det är väl det är väl där också egentligen den värnar vi ju mycket om. Det som är vårt skalskydd så att säga och att att vi behöver tänka på det gånger på vilken information vi vi kan bör och ska exponera utanför skalskyddet. Så vi tittar ju i mångt och mycket här på att göra saker internt inom vårt skalskydd sen sen. Sen får vi helt enkelt diskutera och och se hur, Finns det möjligheter att även jobba på andra sätt kan vi garantera samma säkerhet om vi tittar på molnlösningar eller hybridlösningar på olika sätt. .

Jacobson, Adam - STA: Okej, vi har ju varit inne på det successivt nu men fråga nummer handlar om AI hallucinering, vilket är ett problem om en AI baserad verktyg levererar felaktig data som enligt dem själva ar korrekt. Hur ser du på det? Behöver ni förlita ni er på att leverantör har gjort kvalitetssäkring?

Jacobson, Adam - STA: Eller känner ni att ni också vill ta ett eget ansvar i det här?

X: Om vi tar exemplet har vi använt den leverantörs lösning och vi pratar om det som är vår chattbot skatt idag så kan man ju säga att att här har vi valt att inte använda oss av generativa i den tillämpningen.

X: Och det handlar ju om att vi behöver som svensk myndighet från att stå för de svar vi lämnar och det är väldigt strikt vad vi vad vi svarar på, vilka frågor vi svarar och så vidare. Så på det sättet behöver vi ha en ganska stor kontroll och det är inte leverantören utan det är vi själva som kontrollerar det som är träningsdata och det som är svaren och de svar som frigges kopplat från frågorna. Så på det sättet så äger vi den frågan själva och vi kommer aldrig kunna peka på någon annan Tittar vi sen bredare och funderar liksom på det här med generativa AI och och hallucinering så att säga, så är ju det. Det är ju ett problem och en utmaning med dagens tillämpningar och det finns ju en mängd teknik och metoder för att begränsa de skadeverkningarna så att säga eller begränsa de möjligheterna för i systemet att svara. I vissa fall så kan man jobba med olika typer av provningstekniker för att undvika att modellen svarar trots att den inte har ett svar. I andra fall så kan man jobba med det här med virtual argumentation eller rank för att undvika den typen av av formuleringar.

X: Så det finns. Jag tror att det finns olika tekniker metoder för att begränsa just den här biten och det är ju sånt som vi vi tittar ju på undersöker. De har nu de här metoderna kopplat till att vi också då ser på möjligheten att använda jämna relativa AI bland annat inom. Dels inom. Chatt generellt och kanske inom en del av våra tillämpningar eller våra interna stöd till till handläggare på Skatteverket så tittar vi på undersöker dels olika externa modeller, men vi har också undersöker hur väl vi kan begränsa eller eller skära av den här möjligheten att helt enkelt fabulera.

Jacobson, Adam - STA: Kvalitetssäkring som du pratar om det du metoder ligger liksom den bördan på er eller då leverantören om om ni då skulle ha en leverantör.

X: Ja, jag skulle nog säga att. Jag tror att vi har svårt som som myndighet att komma undan vårt ansvar gentemot medborgare och företag.

Tony: ja

X: Ja, jag tror att vi skulle ha en väldigt stor utmaning i att avtala bort en sådan skyldighet. Delar kan vi säkert få in i ett avtal med en leverantör, men jag tror att det är. Jag tror hellre att vi ska ha ganska stor egen kontroll och egen utvärdering av det eller möjligheten att ställa så pass tajta krav så att det är liksom att det blir tydligt.

Tony: okej

Jacobson, Adam - STA: en Men precis. Jag tänkte på det här med översättning av språk också. Ibland finns det ju ganska svårt att hitta en direkt översättning om det rör sig om väldigt många språk så då går det ju alltså. Även om år så kommer det inte kunna gå. Och förlitar sig helt på en AI alla fall i det scenariot.

X: Nej, jag menar någonstans det vi gör om vi om vi tar maskinöversättningar som ett annat exempel, så det vi har gjort i de Skatteverket nu, det är ju att vi har. Vi har valt att avgränsa användningen eller tillämpningen av det, så vi säger egentligen att vi använder vår maskin. Översättningstjänst för att förstå inkommande kommunikation till Skatteverket. Och det här kan ju handla om till exempel e-post på andra språk. Det skulle kunna handla om att vi genomför en kontroll, en revision och vi får in kanske kontrollmaterial alltså av.

X: Och tal eller kanske bokföringsmaterial som som avser liksom kommer från andra länder eller på andra språk eller någonting sådant i den typen av verksamhet så är vi beroende av att få kanske en snabb översättning. I många fall så har vi kanske sett att. Ja, men säger att vi har dokument normalt så skulle jag börja skicka de här dokumenten till en auktoriserad översättningsfirma och sen får jag de tillbaka och se att OK, det var alltså bara dokument som är intressanta bilar när mängden.

Då har vi lagt både ganska lång kalendertid och vi har lagt ganska mycket pengar på att få översatt de andra nittonhundra nittioåtta sidorna som egentligen är helt irrelevanta genom maskinöversättningar så så vinner vi egentligen tid och vi blir mer relevanta i vilket material vi skickar då vidare till en auktoriserad översättning. Så vi har valt egentligen att att säga att.

Vi har vi har till och med egentligen avgränsat så att användarna inte möjlighet att översätta från svenska till andra språk, så den funktionaliteten är dold för användande då och det enda vi tillåter att översätta från svenska till engelska. För där. Känner vi att de flesta har inom Skatteverket, men tillräckligt god förmåga att så att säga och rätta till och korrigera eventuella misstag som som AI översättningen gör.

Medan skulle jag till exempel ge mig på att översätta någonting till finska? Ja, då skulle inte jag ha en chans att att så att säga kvalitetssäkra den översättning som jag precis säger.

Jacobson, Adam - STA: De här finns det inte en risk att en eller försvinner i mängden om det är ett spår till exempel?

X: Ja, då tänk då är det ju mer egentligen om man säger att tänk det är mer utifrån att att du översatt dokument för dokument så det du får egentligen som som handläggare då i en sådan situation. Det gör att du får en ganska snabb uppfattning om. Vad är det här för typ av dokument?

Jacobson, Adam - STA: OK.

X: Är det är det ett avtal är det. Är en förbundsstat test eller handlar det om ett vigselbevis beroende på vilken del Skatteverkets verksamhet man finns inom? Men du får en uppfattning om dokumentets innehåll och på det sättet kan du bedöma. Är det här relevant för den utredning och den hantering jag håller på med. Och då behöver vi egentligen även om det finns fel översätta ord i det så kommer inte i situationen att du ska fatta ett beslut givet den maskinen översätta informationen, utan det är ett sätt egentligen att sovra bland ett större informationsmaterial. Sen kommer du ändå att skicka iväg de relevanta dokumenten för den applicerade översättningen. För är det ett avtal man går igenom, då behöver du den mer exakta översättning.

Jacobson, Adam - STA: Förstår, jag förstår.

Tony: OK.

Jacobson, Adam - STA: Vi pratade lite om ägandeskap och ansvarstagande och i så som jag tolkar det så så behöver ni ta ett övervägande ansvar. Med tanke på att ni är en myndighet om vi har ett scenario där. Konsulter använder en leverantör för analys. Och sedan gör konsulternas kund ett business decision som visat sig vara fel. Baserat på analysen är det aktörer, det vill säga leverantören av mjukvaran. Och sedan använder analytikerna av mjukvaran för att hjälpa sin kund. Och sen så investerar kunden x antal pengar som visar sig vara fel. På grund av att mjukvaran var fel.

X: Ja, i den typen av förhållanden, om man säger så, så det är klart att det är lite kanske konstruerat, men det är klart att. Skulle vi driva någon slags talan eller någonting sådant så skulle vi ju driva den mot den som har leverantörer som har avtalet med oss så att säga, det vill säga konsulten i det här fallet. Och sen när du vill skapa upp till konsulterna, ta det vidare. Men men jag tror ju i termer av konsulten så så tror jag nog att det skulle vara väldigt, väldigt svårt att driva den typen av talan för många av de här modellerna. Och kommer vi på generativa AI och nyttjandet av dem så är det ju ganska tydligt i väldigt många licensavtal att att det du kommer ha väldigt svårt att driva en sådant fall skulle jag nog säga.

Jacobson, Adam - STA: Vem anser du ska ta störst ansvar i det scenariot? Risken är att leverantören eventuellt gör nedskärningar i kvalitetssäkringen. För så länge dom säljer sin produkt så har ju dom gjort en vinst. För så länge dom säljer sin produkt så har ju de gjort en vinst.

X: Det det, det är ju svårt. Det är svårt på det sättet att. Som jag försöker bara omsätta det i vilka typer av beslut, men om vi om vi säger att vi i vår kärnverksamhet, det vill säga beskattning av privatpersoner eller någonting sådant skulle använda AI och även om vi utvecklar själva eller vi säger att vi skulle ta in någon typ av leverantörer. Konsult som utvecklar en modell för oss så skulle ju vi som myndighet ha så att säga ansvaret för den modellen då egenskap, vår modell och vi behöver också införa tillräckligt god säkerhet för att verifiera att modellen gör de modellen ska göra.

X: Och det kan vi aldrig friskriva oss från. Sen kan vi i nästa led möjligtvis eventuellt driva en tävlan mot så att säga konsulten som i sin tur kan driva en talan mot leverantören. Ansvarsförhållandet är tydlig gentemot om jag tänker mig utifrån att vi har en en. Kunder eller en privatperson som skickade in sin deklaration. Vi kommer aldrig att hävda att nej, nej, nej, det är inte vårt fel. Det är det leverantörerna borta som du ska snacka med. Men vi kommer aldrig att kunna friskriva oss på det sättet. Aldrig.

X: Lång utläggning, men ja.

Tony: Var sista fråga om du vill lägga till nåt i diskussionen som du känner att vi har missat eller vore intressant att diskutera vidare på?

X: Nej, jag tror väl inte, jag tror väl inte det. Men jag tror att vi har varit inne på på de flesta frågorna. Som sagt, det är väl egentligen bara att att. Ger lite grann kontexten också att att mycket av de svaren och det vi har pratat om bli utifrån texten kanske mer att att.

Det är en en stor möjlighet. Vi jobbar med ganska mycket egen AI utveckling så att det blir ju mer generella svar som ni märker då i vissa av frågeställningarna. Och det kanske inte blir alltid i rollen som att jag jobbar till exempel som vi utvecklar och använder av i mitt arbete som vi utvecklare och eller använder AI mitt troll som som analytiker av data utan här blir det ju mer kanske att jag har svarat ganska brett utifrån vad Skatteverket arbetar med i termer av av AI och tillämpningar av AI.

Appendix G. Interview 5

Interview 5 (R5), Qlik representative, Length of interview: 40 minutes, X = respondent.

07:49 X.: So basically so my name is ----. And my current role at Qlik is a senior data scientist and I'm working as a part of the Edma team, which is the enterprise data management and analytics team so. It has been recent. It was called Global Analytics team before and now it has recently been changed to EDMA which is like enterprise data management and analytics which is like a combination of since we recently acquired a company called Talent. So Qlik acquire talent like last year. So since then we've had a lot of reorganisation going on within the company and we have been renamed from Global Analytics to EDMA.

08:33 X.: And I've been a cliques for the past three years. I joined in January of 2021. Since then, I've been a CD scientist working on various projects, internal projects, and working with a bunch of stakeholders. So my role is not customer facing role. It's mostly like internal stakeholders that we have and we're going to. We usually work on data science projects that cater to the needs of our internal stakeholders like sales, increasing their revenue.

09:05 X.: Sales forecasting and churn prediction models like how can we reduce churn within the company like within the for our customers. So those are the different model prediction models that I've built so far as a part of my role here at Qlik prior to that, I worked at at two other companies. My first job was at Delhi MC and my second job was at LendingTree, which is a loan borrower company and at Delhi MCI was a data science consultant.

09:35 X.: Mostly a customer facing role I was working with a lot of big data. Oh, like in in terms of big data vision workshops, conducting like workshops in a lot of different client facing companies and mostly like working on data science projects for them. So that's my little bit of like a brief background about myself. And yeah, do you want to go ahead and, you know, continue with the other questions or how do you want to do?

10:07 Tony D.: We can we have an interview guide. So we're just gonna ask you questions and you can just.

10:13 X.: Sure. Yeah, go ahead. Yeah. Yeah, yeah.

10:16 Tony D.: So we can start with. Could you share some examples of your work and how you work with data analysis? In your everyday work.

10:28 X.: Basically, all of our projects like although I'm a data scientist, all of our projects basically start with a data analysis work of our I mean, of all of our data that we have, right. So most of our data guarantees sitting in Snowflake. So it's a snowflake is a cloud data warehouse and the some of the projects, the examples of projects that I can give you is the sales forecasting model. So we generated it's it's an end to end machine learning tool.

10:59 X.: For opportunity level sales forecasting. Which enables our sales teams here at Qlik to make decisions on what would be our net new licence revenue in the future for the next quarter. What would be our net net new licence revenue and it gives us also it gives us the ability to predict whether an opportunity would be closed one or closed lost and trying to estimate you know how long a deal would take to close. So in for this particular project the first step that we usually do is exploratory data analysis.

11:32 X.: So once we do the data extraction from Snowflake, the next immediate step is to understand what are the relationships and patterns that we are finding in the input variables of the data. So for example, if I'm building the binary classification model of, you know whether an opportunity will be a closed one or a closed loss, which is like falling in just two different buckets. I'm trying to

find out what would be the input variables associated with it, like what is the age of the opportunity, like trying to understand what are the relationships correlations between?

12:04 X.: Different variables put together and also I tried to build a lot of visualisation because pictures speak 1000 words right? Like we need to understand how these visualisations actually help us in determining the I mean bringing us to the final model prediction. So for the first step that we always do is exploratory data analysis and I mostly work in Python And the tools that I use within Python packages are like.

12:35 X.: Pandas. Matplotlib in order to, you know, plot our. You know the the visualisation. Seaborn is another one. So all the various functions for that we could use within Python, so that's what is used for our data analysis. And yeah, the sales forecasting is one example. The other example that I can give you is about the churn and renewal prediction model and to predict SaaS customers, what I mean, how can we predict which are the SaaS customers who are most likely to churn?

13:09 X.: Within the next few months or the next quarter. So for that we have to 1st select in the data analysis. If if I want to talk from a data analysis perspective, we need to 1st filter out you know form formulate a sequel query within within Snowflake. So there are a bunch of tables, say for example five different tables within Snowflake that we want to extract the data from and filter out on only SaaS customers. We are focusing only on SaaS customers, not the entire customer base so.

13:42 X.: We use various filters where clauses within the you know sequel query. When we're formulating it, and then finally bring that, you know data down to only SaaS customers. And then once we have that data set, the final sample data set, we try to do the exploratory EDA analysis. As I said, exploratory data analysis and that's the process how we, you know, go about in doing things.

14:07 Tony D.: OK, so if I understood it right that you work that data degradation was snowflake and working with those those that data is mainly data. Yeah you give can you give us examples when you work with qualitative data analysis if you do so.

14:28 X.: Yeah. So in terms of qualitative data analysis, uh, it's it's mainly so you're talking about like non numerical data's, right.

14:36 Tony D.: Exactly. Exactly. Yeah.

14:38 X.: So for non numerical variables means we usually use like. So I create a lot of. Duration fields and I mean we we we have a concept of feature engineering within EDA analysis. So we create new fields from the existing ones, we derive new fields from the existing ones. So for example if I want to give you an example of the qualitative data, right. We have. I'm just looking at my data sets over here. One good example that I could give you is OK so.

15:20 X.: Having the description of an opportunity like what is the opportunity actually describing itself like what is a forecast category name and what are the opportunity stages that this particular you know? You know this particular licence is falling under. So suppose for example, we are predicting for a a licence ID and we are predicting the final target for the licence ID, whether they will churn or not churn some of the qualitative variables that we use is the age bucket like what is the bucket, is it like less than 120 days of an age bucket? Does this particular licence have or is it like more than 120 days or 90 to 100 days or something like that? Those are the different buckets that it falls under, right?

16:07 X.: So those are some of the non numerical examples that I could give you off the top of my head and there are a bunch of other date fields like what is the usage month like in like what is the subscription contract date? When did it actually start right? What is the date when it started or a date or month. You can also say we can derive month from a date right easily so it it could be a

month for the subscription when the subscription contract actually start for that particular licence.

16:37 X.: So those are the certain things that I could give you examples of. Yeah, qualitative data.

16:44 Adam J.: So like for example Netflix, whenever I enter subscription I always get an e-mail where I.

16:52 X.: Right, yeah.

16:53 Adam J.: State my my reason for leaving. So consider like so those are presets.

16:59 X.: So those are like some of the unstructured data that you're talking about, right, like some of the?

17:04 Adam J.: Or if you consider them qualitative or quantitative.

17:06 X.: Yep. Yep, Yep, Yep.

17:08 Adam J.: Which one?

17:10 X.: I'm talking about qualitative.

17:13 Adam J.: OK, but like I I'm I'm thinking that given that they're preset. They'd still be easily easier to. Analyse given unstructured compared to unstructured.

17:28 X.: Yes, that's right. So quantitative data, you already have a numerical you know you can easily find a distribution for it, right? We can create scatter plots and it's a continuous variable. Basically, you can quantify it. So the name itself the same but qualitative data analysis, it can be a series of categories like you can have a categorical variable with multiple cardinalities, right? It could be a high cardinality or a low cardinality, but it's these are the variables that could be.

18:00 X.: Used for our model purposes and we can. However, we want to use is. We can encode, encode that variable or whatever it is, but non numerical data such as what you're pointing out is one of it. One of it is the comment that Netflix, I mean Netflix actually gives you or you give Netflix. So yeah, that's also an example for it. Yeah.

18:22 Adam J.: Perfect. Thank you so much. We'd like to continue with the AI based perspective. So simply yes or no. Have you ever used an AI AI based tools? It could be an assistant or generative AI. When conducting your analysis.

18:35 X.: Yes, I have. I have used. I haven't used githubilot but I have used databrix AI pilot.

18:43 Adam J.: OK. Could you describe it more in?

18:47 X.: Yeah, so I use it ml apps platform called databricks wherein all of our code is ingested and we synchronise that with GitHub repositories. So it in the Databricks platform itself, in the environment itself. When you're writing the code in in terms of a notebook.

19:05 X.: Once you start writing a cell, there is an AI assistant within the. It's embedded within the. You know the core like within the notebook itself, saying that you can start typing or you can start generating with AI. So once you give a question like a pseudo code or something like that, it automatically generates the code for you. So I have used it a couple of times. I haven't used it extensively, but you can toggle the assistant once you open up a notebook within the Databricks platform.

19:39 Adam J.: Cool. And would you consider that to be an assistant or generative? Those could be tough to define at times.

19:47 X.: Yeah, I would use it as an assistant, mostly AI assistant, so I haven't used the GitHub pilot, but yeah, this databrix one is like similar to that.

19:58 Tony D.: Yeah.

19:58 X.: Yeah.

19:59 Adam J.: Have you ever used something similar to chachiba tea or Bart for example?

20:05 X.: For work, yeah.

20:07 Adam J.: Yeah.

20:08 X.: Yeah. Yep. Yeah. For for work? Yes. I have used ChatGPT. I've used antopic cloud AI as well.

20:16 Adam J.: And what do you use it for? Do you like help with with? Defining the code snippets. Or do you use it for the analysis itself?

20:26 X.: I use it for uh formatting my sequel code and uh, once I have a sequel code which is throwing me an error. Umm or something like that. It it? It can easily format it for me and with the right indentation and all of that. So I've used it for that previously and also to give me examples of say for example I define a use case right for it, it will actually build that entire step for me. Like if I say hey build a what are the step to step process to build a binary classification model?

20:59 X.: For renewal and churn and it's going to show me all the steps with the sample data set. So those are some of the snippets that I could use. I mean it's it. I don't usually use the exact snippets that it gives, but I use the idea of it. So I tried. I tried take this essence of it and then I tried to apply it into my it apply it into my work.

21:19 Adam J.: So yeah, do you see issues with inputting real data in the GPT for example?

21:27 X.: Oh, I have never done that, to be honest. Oh, I haven't not. I've never inputted any real data. I did input because real data is like giving away your company proprietary information. So I've never done that. So I I think that is kind of they need to actually restrict it, but they have not restricted a Qlik or to be honest with you, even though I am connected to VPN and everything, they haven't restricted it. But I do not give away any of our personal information. I just asked for the snippets of code by giving the use case itself. So so that none of our personal.

22:02 X.: Information is sorry, none of the company information is revealed over there.

22:06 Adam J.: OK. So you mentioned some tools, does some of them use your inputs to train itself? Or like those the the software that you're using in Sonic for example, so not obviously.

22:24 X.: Software you mean? Are you talking about the data bricks one or are you talking about the AI tool that I was talking to you about?

22:30 Adam J.: Yeah, so so like. Obviously Houston GPT it you'd you'd be able to train those GPS, which isn't necessarily beneficial for your company, but like when using it in Qlik, it could be beneficial for you to train.

22:46 X.: True. True. Oh, you mean if it's embedded within? Umm, our uh Qlik tools, right? Yeah. Yeah, yeah, yeah. So that has not been done yet because there is a if. Have you heard of Qlikstage? That is like an AI tool which is like going to be embedded within our Qlik Qliks and's platform itself. So that is something that I haven't, you know, got my hands on yet. But I'm planning on using it, leveraging it in the near future. But I haven't used it yet within the Qlik.

23:19 X.: You know platform itself. Yeah.

23:21 Adam J.: Perfect. I see. Thank you. So would like to continue with the AI assistance in regards to big data and you mentioned big data a bit earlier. I believe it was.

23:31 X.: Yeah, in one of my, yeah, if I never used big data, there's no big data and Qlik to be honest, there is no, we have very. So the Max amount of data that we ingest into our, you know into our models is like 5 gigabytes or 6 gigabytes. So that is nowhere closer to big data. I don't consider that as big data, so.

23:53 X.: My thoughts like I've used big data way back in the past, right, like six or seven years ago for my projects at Delhi MC and yeah, like AI assistance for big data. Obviously for complex data sets, it's going to be very

helpful because the algorithm that it is going to develop and all of that will be like it can analyse obviously massive amounts of data, right?

24:18 X.: Without any human intervention, and we don't have to do anything manually. So I feel like some of the advantages of it would be like, as I said, predictive analysis that I am doing for my for some of my projects. Suppose for example, the churn data is really huge, it's like one terabyte or two terabytes and it has like multiple columns. I am using hundreds and hundreds of columns for my.

24:44 X.: Model development. So definitely that time AI is going to be very helpful for handling that amount of data. And also we use natural language processing. For you know, some of our projects we use like case data like there are some data, there are some comments data that customers actually put in like if they like a particular feature of a tool. So in Qlik sense if they don't like a feature they actually put a comment over their CSAT comment saying that it's a customer satisfaction survey, right.

25:16 X.: So they kind of write some kind of comments on it.

25:20 Adam J.: OK. I see.

25:23 X.: So those comments can be used like since they're all unstructured data. AI can be AI powered, NLP can be really helpful in that way.

25:32 Adam J.: So you make the distinction between machine learning in regular data analytics compares to big data. OK.

25:40 X.: Yeah. Yep. Yep, Yep, I do that. Yeah.

25:42 Adam J.: Would you like to continue, Tony?

25:45 Tony D.: Yes, just one side question before I continue, what would you define as big data? I'm just wondering personally.

25:55 X.: I would define big data as I mean obviously the data that can be cannot be easily processed with a without using any parallel processing techniques. So for example in earlier days we had Hadoop. Right now everyone is using Spark right? But in memory utilisation and in memory processing of data so spark is really powerful and if you don't see the use of need of spark using spark.

26:25 X.: A multiprocessing or sorry parallel processing system like Hadoop. Then I feel like big data is, you know, not necessary.

26:34 Tony D.: OK, OK.

26:35 X.: Yeah.

26:36 Tony D.: OK. So next question, based on your experience, what are the benefits you see with the use of AI in data analysis as a whole?

26:45 Adam J.: Should we ask about power BI or BI tools?

26:49 Tony D.: Yeah, yeah.

26:50 Adam J.: That's.

26:50 Tony D.: Yeah, sorry. So we can.

26:52 X.: I will be what's the next question, sorry.

26:54 Adam J.: So just just or no, have you ever used? BI tools in your analysis.

27:04 X.: I used only Qlik sense so far and I use Tableau.

27:09 Tony D.: OK.

27:09 X.: Does that count? I use Tableau before. I've never used power BI, but I used to have a tableau licence. Prior to joining Qlik, I used to have that and also after joining Qlik also I had it for a brief time just to compare our tool with that tool because there are competitor right. So since there since Power BI is also competitor but I've never used it, I use Tableau a lot and.

27:35 X.: Yeah, I would like to see it more AI generated. As I mentioned, Qlik stages at that particular I mean Qlik stages, something that I'm looking forward to and I want to see what are the benefits of AI assistant for using Qlik I mean data analysis at Qlik.

27:52 Adam J.: Hmm.

27:53 X.: Yeah.

27:56 Tony D.: And so now to my question. So based on experience, what are the benefits you see with the use of AI in data analysis?

28:08 X.: Oh, mainly I think accuracy. It's an improvement in accuracy, right? Your, like all the AI systems, can apply like some kind of analytical models, and they shouldn't make making logic. They're a little bit consistent. It will reduce the risk of like a human error or the bias in data analysis. So it's a more it's more accurate and it's more consistent, I feel and.

28:36 X.: Obviously, improvement of predictive capabilities. Like you know, the identifying the patterns and historical data that you have, the AI models can be used to make predictions about future outcomes. So that's also an important advantage I think and also.

28:55 X.: Like like a good collaboration between in human and machine, we have a human and we have machine right? Rather than replacing all the human data analysis analysts, we can use AI to augment and enhance the capabilities. Instead of, you know, just replacing them, allowing allowing the data analyst to work more and efficiently and effectively by using the tools. I think that's what I feel should be important.

29:23 Adam J.: Pinpoint any concerns or challenges, maybe in terms of safety or ethical concerns.

29:31 X.: Ethical concerns? Obviously there.

29:36 Adam J.: It could be could be a broad perspective or in terms of your organisation as well, so it could be.

29:40 X.: For the organisation. Right? Oh, OK. I haven't seen it yet personally, but I feel like it could pose problems. Maybe for privacy and our data rights that we have.

29:52 Adam J.: Hmm.

29:53 X.: Do we need to, like, strike a balance between the benefits of, like, you know, a data-driven insights with protection of our our privacy as well data privacy as well. So that is one thing and also maybe. There could be some unjust. AI decisions that could be made like you know. Unfair or unjust AI based decisions it can. It can kind of perpetuate societal inequalities also. So we don't know what it's going to come. But that is that could be one of the threats or concerns that I would have.

30:30 Adam J.: How do you feel about AI confabulation or hallucination, which means that? The AI will give you, instead of admitting that it may not necessarily know the correct answer, it will give you what they would think would be the right answer, which might be wrong. So do you see that?

30:52 X.: Yeah, my thoughts on that is like. I think that it could harm it, it could be harmful, right? It is a potential of being harmful. But do these can be. We have we have some kind of techniques right now in order to rectify that like you can use the usage of racks like context based approaches, augmented generation, generative AI. So those things can be. Clearly ingested into the data and each company, they have their own rag based models. Then that could be helpful.

31:32 Tony D.: Yeah. So if we go back to tomorrow, more of the concern perspective when it comes to AI and data analysis, assume that a company has you can take Qlik makes a business. The business decision based on AI data.

31:50 X.: Mm hmm.

31:50 Tony D.: But it is wrong. What is issues do you see with that? Where is like? Where is the responsibility? Is it from the? AI, like the third party that you're incorporating in Qlik. If you use a third party? Or is it?

32:09 X.: Oh, if you use like an external. AI tool right for data analysis. OK, I feel like it will generate. I mean, it's not going to generate like optimal business outcomes like our company. This could put our company at a very at a very bad position like a competitive disadvantage or like it could even threaten like other

customers even using it right for the long term viability. And there could be like a problem of like an erosion of trust, the loss of trust that can be it. I mean once the trust is lost, you can never get it back, right.

32:52 X.: So there will be distance with that one as well, and it could. It could hurt our reputation like image that Qlik has in the market right now Qlik has a really good name in the market. In the Gartner chart and all of that, right.

33:07 Tony D.: Hey.

33:07 X.: So I think that is going to the reputation will be damaged because of this and also the ethical concerns that we have.

33:15 Adam J.: Hmm. So in terms of quality. Work for the third sacking tonight. Quality cheques of the of the software. So in case you would use an AI vendor for your software.

33:32 Adam J.: To consider them to be responsible for quality checking. Or would you do some as well?

33:39 X.: We also have to do it like you before even moving ahead with anything, right? We need to do APOC strategy with a with a valid sample data set that with this which is like related to Qlik not not a random sample set that you just download online or from any website we need needs to be Qlik's data that you need to do a POC with and have the proper contract signed with the with the vendor that the data is not going to be, you know leaked or something like that. It is properly protected.

34:10 X.: And after that we can. You know more ahead with the tool if you like it or not.

34:17 Adam J.: Hmm hmm.

34:18 X.: If it, if it passes all your quality cheques or not.

34:21 Adam J.: Hmm. And if you would have a customer who invests a lot of money based on an AI based analysis which turns out to be wrong, should they take accountability as well? Or is it mostly you and the vendor in that case?

34:39 X.: I think it should be mostly the mostly US and the vendor. The Qlik and the vendor should be taking the responsibility, yeah.

34:48 Adam J.: OK. That case. Would like to just ask you if you would like to add something else to our discussion. That we may have missed out on.

35:02 X.: Yeah. So for the thesis that you're working on, is it like based on a topic, what's what's your, what's your is it like a broad topic or is it like a specific topic for your thesis? I'm just curious, yeah.

35:15 Adam J.: But the recent question, yeah, it's how data analysts perceive and experience AI usage in data and.

35:21 X.: Oh, got it. Got it. So you're doing a survey for with everybody else also? I mean, not just.

35:27 Adam J.: Yeah, qualitative survey. So we've had five interviews counting today and one more tomorrow. So as I mentioned earlier, it's some are data analytics, some are data scientist, some are working mostly with the code.

35:43 X.: Oh, I should have been in the data scientist one I feel.

35:47 Adam J.: Yeah, exactly, exactly. And I'll be at. Gin would be more considered also a data scientist, right.

35:53 X.: Right, right, right, right. So the role is very broad right now. All of our all of the. So you know it's not just a data scientist. There is like an AI scientist that is a prompt engineer. There are like so many roles that are coming up. So yeah, it's like branched out massively.

36:13 Tony D.: Exactly. So the main point of our research is to capture all of the data stakeholders using AI within their work processes and. Collect like a when collecting the data, we'll get a picture of how they perceive an experience. This usage of AI, so that's why the questions are pretty broad and not so specific.

36:39 Adam J.: So yeah, yeah. Yeah. So in light site we we might have, you know,

looked for a more specific topic, but we found it very interesting. So we tried to gather the impression of many people instead of different roles.

36:55 X.: Mm hmm. OK, that's awesome. Yeah. Well, I'm glad I could help you guys. At least a little bit in your research.

37:02 Adam J.: Yeah, it's been very. Thank you so much.

37:03 X.: Good luck.

37:04 Adam J.: So yeah, so there's nothing else you'd like to add to our discussion, any topic that we may have.

37:13 X.: No, not really. I think you guys pretty much covered like all the ethical concerns and data privacy issues and whatever you wanted to ask me, right, so that's that's pretty much.

37:24 Adam J.: In that case, we we would like to thank you so much.

37:27 Tony D.: Yeah, we do. We appreciate it. We can send you the transcription of the interview if you would like to do it.

Appendix H. Interview 6

Interview 6 (R6), BXASE representative, Length of interview: 45 minutes, X = respondent.

10:56 Tony : Skulle du kunna ange ditt namn och din nuvarande roll? Och hur länge du jobbat inom den här positionen?

11:08 X : Hej, jag heter ---. Jag har senaste över 10 år. Det blivit nu jobbat som frilansande konsult inom e handel och online marketing i roll av, dels management Consulting på en strategisk nivå men även som projektledare. Inte regnbåge och Gotland rent praktiskt tagit över och roller hos hos kunder inom det här området och innan dess så jobbar det som.

11:40 X : Olika chefspositioner och e handels och online marketing företag.

11:45 Tony : Härligt. Skulle du kunna dela med dig av några exempel på ditt arbete när du jobbar med dataanalys?

11:56 X : Ja alltså, jag som sagt har jobbat med online marketing väldigt länge och online marketing är något en en bransch som har förändrats väldigt mycket de senaste. 15, 20 åren egentligen som har jobbat med det och den har alltid varit väldigt datadriven att man mäter massa mätetal på på performance, hur mycket man betalar för olika former av reklamkampanjer, vad folk klickar på, vad de inte klickar på, du hur många som sett den. Har någon så många som klickar på annonsen om hur många som hamnar på 1, 1, 1 landningssida så att säga på en webbsajt där de tittar på en produkt.

12:29 Tony : ja

12:42 X : Hur många som sedan väljer en storlek på produkten, hur många som? Beslutat sig för att köpa hamna på betalsida och hur många som lyckas avsluta en betalning och sen även då så jag följer våra kunder göra under under ett års tid och det man säger att man jobbade på. fram till 4 5 år sedan egentligen det var helt enkelt att man samlade om man fick all den här datan från olika format och sen så sätter man in det ett stort excelark och sen så börjar man tittade på hur man kan optimera det här Fram till 4 5 år sedan egentligen. Det var helt enkelt att man samlade om man fick all den här datan från olika format och sen så sätter man in det ett stort excelark och sen så börjar man tittade på hur man kan optimera det här.

13:20 X : Vilka typer av reklamkampanjer är? Effektiva i termer utav att de ger bäst avkastning tillsammans stoppar in 1000 kronor i budgeten. Man lägger det på 3 olika annonser och sen så ser man vilka 5 som är de bästa och de vill man då fortsätta lägga på de 5 dåliga de de. Slutar man ge budget till så att säga och och det här är ju på ett sätt dels tidigare då kvantitativ dataanalys som du har kopplat också till till kvalitativ data lista för att man måste.

13:57 X : Sätta sig ner och fundera på varför fungerar vissa saker bättre än andra? Hur kan man förklara att att en produkt är i lönsam och en annan inte är det? Och då som sagt vad sitter man och tittar på på här? Här måste ut ett titta på det här och då liksom bara ibland och kvantitativ och och kvalitativ dataanalys och mycket utav de här optimeringen har skett skett för.

14:13 Tony : ja

14:29 X : Band. Det som har skett de senaste åren inom online marketing, det är att AI har fått en större roll. Och det har skett på det sättet att dom här stora marknadsföringsmejl. Nätverken som erbjuder annons Källström, står det då Google och Facebook? De har själv tagit över mycket av det här optimerings arbetet, men de berättar inte och optimeringen handlar då anser att de vill.

15:00 Tony : ja

15:02 X : De vill optimera.

15:03 Tony : ja

15:06 X : vinna Dina intäkter på reklam? Och sen gör de då genom att matcha annonsörer med människor som vill köpa någonting på ett sätt som är så effektivt som möjligt och det de har gjort de senaste åren är att de har släppt mindre och mindre data till över hur enskilda annonser förklarar, fungerar och har för sig mer eller mindre.

15:31 X : Stängt in den här. Annons styrningen i en, i en, i en i en Black box som dom bara anser insyn till som dom själva optimerar. Och det innebär att jag sättet man jobbar så på som som håller marketing, men du har förändrats så att den här kvantitativa analysen har mer eller mindre försvunnit. Den gör dem, vilket innebär att man har sämre möjligheter att påverka optimera också, utan det blir som sagt att man får.

16:03 X : Bygg en kampanj på det sättet att man får sätta ett mål och sen så är det upp till Google eller Facebook att leverera det och sen finns det inte mycket mer man kan göra.

16:14 Jacobson, A.-S.: Säljer jag inte dem den då? Den här blackboxen säljer jag inte om den här blackboxen alls då.

16:16 X : Förlåt. Alltså, det är de det de säljer är ju är ju en en möjlighet att man köper reklamplats så att om jag tidigare om om jag tidigare var plötsligt att jag fick en budget på på hundratusen.

16:21 Tony : ja

16:31 X : Och sen så kunde då sätta OK från dom här hundratusen så vill jag göra reklam där och där och där. Jag vill använda de här reklamerna och jag vill använda det här budskapet och de här bilderna. Och sen fick man någon form utav resultat tillbaka och så fick man då. Fick man då performance utav det här och sen så kontinuerligt så höll man då hela tiden på optimera och titta på. Okej, det här funkar. Det här funkar ju inte så att det funkar på nu mer eller mindre. Det är att. Har sett och ren hundratusen till Google och sen säger man att jag vill sälja för så här mycket. Att jag vill betala per försäljning och då vill jag betala. 100 kronor för varje såld. Byxa så att jag faller och byxor. Och.

17:26 X : Och på ett sätt så är det är det är det bekvämt därför att du inte göra mycket mer än så och sen så tar de hand om eller den här matchningen mellan annonsörer och och. Och människor som vill köpa saker, men du betalar också sen i snitt 100 kronor per per per per per byxa i reklam utgiften där du tidigare kunder riskerar det att om du var dålig på den här marketing så fick du betala hundrafemtio kronor.

17:45 Tony : ja

17:52 X : Men vad du är duktig och kunna optimera på ett bra sätt så kunde du trycka ner koll för den till 45 till så att säga så att företag konkurrerade med varandra genom att man var duktigare eller mindre duktig på när marketing det går mer eller mindre inte idag därför att hela hela den här annonsering processen och vilka vilka saker man använder II marknadsföringen har automatiserats.

18:17 Jacobson, A.-S.: Så vad blev det edge då är det analysen som är edgen.

18:22 X : Alltså, det som är analys är det de här fördelen. Det är det som sitter de ju på alla. mm Besöks inte och det vill säga de vet de tittar ju på de vet var människor söker, de har monopol på dem, på Google search eller på Youtube eller på Facebook och mata.

18:39 Tony : Ja.

18:43 X : Men sen så har ju de datan också. De har ju jättemycket data som dels dom har själva sen som får köper de in. Och en enskild marknadsfönster, man kan inte kan inte. Konkurrera med den de datamängden och de har och att de sitter på datan och de de har liksom mer mindre monopol på på reklamtyta och.

19:04 Tony : ja

19:06 X : mm Sitta i en väldigt väldigt bra situation där och kan därmed på ett så här. Har en ett extremt övertag gentemot annonsörerna. För dom dom säljer och det innebär liksom att. Vill man göra om den marketing? Ja, då måste man gå till Google och Facebook, liksom dom tillsammans här. Charcot mer än 80 90% av marknaden så att säga och det finns ingen väg runt förbi dem. Så att det det det är hur hur har det här förändrat det? Det som har det som däremot man har möjlighet att förändra det är ju på på på sin egen.

19:47 X : vid sitt Med sitt erbjudande och då kommer vi in på det här med vad min omgivning så att säga kring kring det hela och det är just för att det som man. Som handlare annonsör själv kan hantera det till exempel oftast prissättningen på en vara. Och. När man prissätter det vara så så så kan man göra det på. Och på på. Ja men titta på den totalkostnaderna år, men tilliten är lite marginal och sen så har man prissättning som som som sköter det. Men det man vill göra är att man vill hitta ett sätt att optimera prissättningen. Och det vill säga att sajten kostar femhundra, men. Vad händer om priset? Jag erbjuder fyrahundraåttio kronor kommer att sälja mer då.

20:45 X : Kanske då, men vad händer? Blir det mycket mer för att de motiverar prissänkningen eller om man om man höjer priset till femhundraåttio kommer säkert sälja lite mindre men motivera den prisökningen än den totala intäkten blir blir större. Och. Det här med försöka hitta den optimala prissättning kallas för för för liksom det priset eller stridighet. Som man vill optimera för och ett sätt att testa då utvärdera det historiskt. Det är att jobba med med maskinlärning för att göra analyser på historiskt data. Och det är utvecklat en metod för nöjda man. Ta in kunddata som har med prissättning att göra, det vill säga vilket pris man erbjuder, vilken rabatt man erbjuder.

21:44 X : Vilka associerade kostnader är till exempel lite säga att jag säljer byxorna för femhundra så har jag en en produktkostnad på hundrafemtio kronor själv och hur mycket man betalar för för för för marketing för att hitta hitta ett sätt att optimera det här på. Och då är idén att man med maskinlärning går in och tittar på historisk data för att hitta hitta sådana här priser. S elasticitetsmodul kurvor för antingen enskilda produkter eller för olika kategorier eller för varumärken. Och.

22:21 Jacobson, A.-S.: Vad kommer den datan ifrån? Alltså är det bara er är det är det bara er data ni kollar på då ni kollar ni på all möjlig data för att bedöma det?

22:24 X : Flat. Man tittar på kunddata. Så att det kunde kunde kundens egen.

22:34 Jacobson, A.-S.: Men liksom kollar man branschen eller bara kunden då liksom.

22:38 X : Jag tittar bara kunden.

22:39 Jacobson, A.-S.: OK OK. Skulle du beskriva mjukvaran är det liksom mer en assistent eller skulle säga att den är generativ? Assistent kanske är lite mer än chattbot i medans generativ. Gör det liksom självmant.

22:53 X : det är Egentligen ingen mjukvara, utan det är en algoritm och det är en algoritm som. Egentligen.

23:06 Tony : ja

23:08 X : I grunden gör man, antar jag? Mm. Det han tror att jag orkar regressionsanalyser man gör hur den har måtten fungerar, det ska vara ärlig så vet jag inte riktigt vet. För det är liksom det är en det är en helt enkelt revisionsarbetet, så tittar man på sannolikhet utifrån vissa priser man erbjuder och sannolikt utfall då över. Vad den har köpt processen?

23:37 Jacobson, A.-S.: Jag tänker liksom II kan ju hjälpa mig väldigt, väldigt stora

datamängder men om man bara ska analysera en produkt inom en kundbas.

23:39 Tony : ja

23:47 Jacobson, A.-S.: Så kanske inte. Det är liksom miljoners miljoner data, eller det kanske det. Jag tänker liksom behövt alltså, vad är skillnaden med hur det har sett ut för 5 år sedan? Numera liksom.

23:58 X : Så det som är. Det som är skillnaden, det är just att man. Kan. det är alltså maskinlärning består ju av 2 bitar den ena är supervise där man liksom sätter in jag tydliga regler för det här och sen så kan man då med hjälp utav maskinlärning göra olika tester och testa olika priser och mäta olika priserna Det är alltså maskinlärning består ju av 2 bitar, den ena är supervise där man liksom sätter in jag tydliga regler för det här och sen så kan man då med hjälp utav maskinlärning göra olika tester och testa olika priser och mäta olika priserna.

24:29 X : Ja, den där är att man just hittar den här hittar den här optimala prissättningen på ett på ett på ett sätt som som svårt att göra på ett manuellt sätt därför att man inte hinner räkna på det sättet.

24:35 Tony : ja

24:42 X : alltså det är liksom den den den vattenkraften där men sen så finns det ju det här 1 1 1 1 annan aspekt av det som kallas för answers och det är att man helt enkelt skickar vi en massa data II systemet och sen så söker du ändå hitta olika olika Alltså, det är liksom den den den vattenkraften där, men sen så finns det ju det här 1, 1, 1, 1. Annan aspekt av det som kallas för answers och det är att man helt enkelt skickar vi en massa data II systemet och sen så söker du ändå hitta olika olika.

25:02 X : runt i samband Tunnband som man själv inte var medveten om från början. Och det kan till exempel vara hur man hur man. Klassificerar kunder på ett visst sätt eller kundgrupper. Att ofta när man gör den här typen av analyser så slår man in liksom ner kunden. Då kan du titta på. Det på på kända fakta som vilken ålder kunderna har, vilket kön de har det man kvinna och vad de bör liksom tydliga sådana attribut du kan titta på på det så kallt så behöver hon attribut, det vill säga vilka typ, hur vilka köpvanor de här.

25:38 Tony : ja

25:46 X : Om de köper mycket. Men sällan och de köper. Om dom hela tiden. Vi har den billigaste varan och lättare och lättare rabatter. Eller om det inte spelar någon roll. en i och Och och och göra den här typen utav analys som manuellt. Och liksom det blir ett enormt direktiv arbete och man liksom berättar verkligheten. Noll en i en höstack dör. Med med med med diplom nu, så kan man så att säga sätta in de här parametrarna och och få ut väldigt snabbt en analys av av datorn.

26:20 Tony : ja

26:24 Jacobson, A.-S.: Men då Tranås den på befintlig data inom kundbasen och man vill ta marknadsandelar i ett nytt land. Då är inte det samma data, kanske som det ändå har varit. Så jag tänker, liksom?

26:37 X : Nej, precis. Precis datan man tittar på i första hand så tränar algoritmen. Det är historiskt data. Och och. Det man ju så att man jobbar på att man gör först analys utav historiens data ser man hittar något samband hittar man några samband sen så så kan man testa dem. På på på på ny dator och sälj lajv att man då tar det till till butiken och testa.

27:01 Tony : ja

27:04 X : kom vi fram till att för den här kundgruppen som köper på det här sättet Kom vi fram till att för den här kundgruppen som köper på det här sättet. Så är dom inte lika priskänsliga och då gör vi antagandet att utifrån analysen att vi kan ta 10% mer i pris.

27:09 Tony : ja ja

27:16 X : När vi då får en ny sån här kund från den kundgruppen så erbjuder högre priset och sen mäter vi då efter att tillräckligt många har. Har köpt och

sådär. Så att.

27:26 Jacobson, A.-S.: Det känns som vi har en täckt mycket av 2 och 3 nu huvudsaken här frågorna. Jag täckt hela spektrat och för sig, men vill du säga någonting om vi då skulle du klassa lite som big data. Vi pratar lite om data volymen. Har du någon tydlig gräns för vad det innebär?

27:46 X : Alltså big data var någonting. Man pratar väldigt mycket om. För 10 år sedan byggdes vid nyåret som man liksom skulle driva alltihopa. Och sen säger man att det finns en sak som byggde, det finns bara dit, det vill säga det måste finnas data i strejk kan inte säga någonting för för det måste alltid finnas en kontext. det måste alltid finnas 1 1 1 1 problem som ska lösas eller precis och Det måste alltid finnas 1 1 1 1 problem som ska lösas eller precis och. Och. Reset. Det som avgörande för maskinlärning är just mängden data man har, men hur stor den mängden måste vara IIII.

28:34 X : Beroende från fall till fall, så därför skulle jag säga att det liksom inte finns någonting som alltså det där begreppet big data är inte relevant för för mitt mitt område.

28:43 Jacobson, A.-S.: Nej, men vi har haft olika hållningar kring husbygget också. Många som är inne på samma spår som det är någon sa att bygga det där är liksom det data som inte som man inte kapabel till analysera. Det är därmed är därför det är 2 bygg så att säga så det är olika hållningar. Vill du förstår ni?

29:03 Tony : Yes. Har du använt dig av bevis Intelligence verktyg? I ditt arbete, vilket känns som? Vad satt på det då pratat om tidigare i intervjun så är det väl mycket som utgår ifrån business Intelligence verktyg. skolan För analysen.

29:22 X : år Ja alltså sättet jag har mest varit med Excel och sen med med programmering som har bearbetat datan och byggde oftast gränssnitt II pyton.

29:34 Tony : ja OK.

29:38 X : Ja. I för att ni ställer en fråga kring power BI och. Det har jag inte.

29:47 Jacobson, A.-S.: tabell Kablarna kanske?

29:48 Tony : Ja alltså.

29:49 X : Ja precis så att det blir precis det blå har jag har. Jag har jobbat med för han heter det andra. Med det mest visualiseringsverktyg så att säga och och det är någonting jag tänkte på när som jag fortfarande inte sett mig som jag tänkte på när chatty slog igenom till att börja. Det är ju att alla de här verktygen, framför allt Excel och Microsoft verktygen är ju ganska oanvändbar ovanliga.

30:02 Tony : OK.

30:21 X : Kamut av att man vet vad man vill göra och behöver göra, men exakt vilka tecken, kommandon och tyvärr igångsatt. Är det får man lära sig och sen glömmer man då sen så får man liksom när man behöver det så. Om man går in och gillar serien och så vidare. Det hade kunnat göra det. Gränssnitt hade kunnat göra mycket mer användarvänligt med hjälp utav naturliga processen till exempel som som var väldigt starkt på. Till exempel att du har du har datan framför dig och sen så liksom före OK, jag vill ha den här. Datan presenteras i en tidsaxel med det här på exakt den här by axeln och jag vill ha det nedslaget från belönad och jämföra flera år efter varandra.

31:08 Jacobson, A.-S.: Så arket är en allt putt istället för input och allt kan man kanske säga det.

31:13 X : Ja precis precis och det. Det är liksom det. Det jag sa nu. Det var liksom ett ett 10, 10, 15 sekunder så prompt.

31:14 Jacobson, A.-S.: Det.

31:20 X : Ja och jag vet, jag vet liksom jag vill ha när ska jag göra det i Excel eller Outlook eller tablå? Så kommer det ta mig 3 timmar och få den avbruten.

31:37 Jacobson, A.-S.: vi fortsätter vi har redan varit inne på föräldrarna såklart eller något annat du vill peka på kring just fördelar med Vi fortsätter, vi har redan varit inne på föräldrarna såklart eller något annat vill peka på kring just fördelar med A inom dataanalys.

31:49 X : Ja, det är ju det här med när man kan, det är ju den här, för den är ju framför allt snabbheten för att man direkt kan gå från en tanke till att få fram data och göra göra undersökningen. Och att prestationen blir blir snabbare och tydligare därför att man slipper lägga tid på. Där hantverket. Men jag ser inte på det sättet att AI förändrar jan själva analysen därför att det man får fortfarande sitta och tänka och fundera på varför det är på visst sätt. Till exempel om jag ser i försäljningen att förra året så var så jättemycket röda byxor och det här året säljer väldigt mycket färre röda byxor. Vad beror det på? Och där tror jag inte datan i sig är så mycket till hjälp utan jag trodde mer handlar då om en.

32:42 Tony : Ja.

32:45 X : Allmän generell trend eller något psykologiskt har gett oss människor som gör att. Netgears till ett annat, så köper man köper man en färg. och vad det beror på är i otydligt så att säga och där tror jag fortfarande man är långt ifrån att kunna använda AI och förstå sådana saker Och vad det beror på är i otydligt så att säga och där tror jag fortfarande man är långt ifrån att kunna använda AI och förstå sådana saker. Ja.

33:07 X : Så att precis det ja 2 fördelar framför allting som blir snabbheten och och enklare. Jag har ett handtag och på ett sätt så ska vi ha lika mycket av konsultarbete så sitter man och jobbar med det här hantverket och liksom 20% analyser och 80% män håller på att fila med olika prestationer och kan man skifta det?

33:15 Tony : ja

33:30 X : Så kommer man frigöra mycket mer tid till att man verkligen sitter och och och tänker och skulpterar och Ann en använder mindre tid till att. Till att klippa och klistra som egentligen är ett av av data.

33:46 Jacobson, A.-.S.: Vill du påstå att han är?

33:48 Tony : i Ges vilka utmaningar och risker ser du med att använda ett av AI?

34:07 X : I. Man kan alltså utmaningar i och med att. Har inte med själva AI att göra det för att det viktiga är själva analysdelen och den måste man fortsätta göra själv. Man ha, i kan inte göra analysen åt den del en del de här, de här systemet. Vi har nu med även med kopite. Funkar inte på det sättet? Risker.

34:36 X : Kravell med så generellt. Med. ja i AI system som är så att säga? Som som inte bara gör analys utan även cancer så här ta. Fatta beslut beroende på på vad man har lagt in och sen, men det är ju en del utav en. Reda diskussion liksom inte som kopplat till mitt. Området direkt. Så att på det sättet så. Ja nej, jag ser inga större, större större ryskor där så att säga vad vad det gäller användandet utav a inom inom inom mitt område.

35:20 Jacobson, A.-.S.: det kan vara risk att ni om du till exempel använder truth BT och användare är kunddata till det att era kunder känna en bristande förtroende eller inte koketteri specifikt men liksom att ni använder Det kan vara risk att ni om du till exempel använder Truth, BT och användare är kunddata till det att era kunder känna en bristande förtroende eller inte koketteri specifikt, men liksom att ni använder.

35:38 Jacobson, A.-.S.: Deras data inom andra program liksom.

35:41 X : Ja. Men sa inte vad vi använder data utan att, utan att. Kunde nog sitt tillstånd till en och ofta gör jag då kunna se tillstånd till igenom när de betalar köparen vara så ger man ju alltid recept till till tomten. Conditions II ett köp.

36:02 Jacobson, A.-.S.: Men då kanske de tänker att det töms gentemot, det vill säga företaget som köper han inte. 3, 4 andra a leverantörer liksom precis.

36:07 X : Ja. Ja, det är alltså det den biten löser man nog bäst och det är det jag kommer ihåg. Presentera som liksom lösningen är att man anonymiserar datan. På det sättet att man tar bort den mest, liksom. Namn adress. Behöver man inte för att göra den här typen utav analyser. Det man behöver födelseår och om det är en man kvinna, det finns det ju risken att att den här risken är ju den här risken är liksom kopplas generellt till till till data. Det är liksom ingenting som är

speciellt för för AI och det här. Risken finns ju att man med hjälp utav. Den här datorn ändå kan identifiera vem exakt den här personen är. Beroende på på hur mycket data man har tillgång till. Och.

37:18 X : Ja den den problem som ju finns och att man man man gör med just och har liksom en väldigt tydligt tydligt det liksom tar hänsyn till lagstiftningen och att man har rätt information IIII.

37:23 Tony : ja ja

37:34 X : Användarvillkor, men det är inte, det är inte kopplat till. AI sig utan det är liksom den den generellt. Den här datan har varit företagen samlat in och sedan måste jag så länge haft internet att säga.

37:45 Jacobson, A.-S.: Precis. Jag tänker att det är skillnad. Skillnaden är väl kanske då att AI kräver data för att maskininlärningen och helst då riktig data så löser det dilemmat så måste man ju på något sätt hitta riktigt data och det kan man då göra genom ett direkt.

37:48 Tony : ja

38:06 Jacobson, A.-S.: så Direkt kommunikation eller då indirekt kommunikation när man använder sina kunders kunder data liksom. Så det kanske är den aspekten? Där AI är skillnad gentemot tidigare. Alltså incitamentet av maskininlärning.

38:24 X : Väntar dig vad du menar med att det var skillnad skulle vara.

38:28 Jacobson, A.-S.: Men innan så. Så i sortimentet maskininlärning kräver riktigt data för att bli bättre. Mm och det behövdes inte på samma sätt tidigare.

38:43 X : Nej, precis, men samtidigt man samlade man har samlat in. Det är samma data man samlat in.

38:47 Jacobson, A.-S.: Ja.

38:48 X : Och det är samma datum men som någonsin tidigare och det är samma dator man fortfarande samlar in.

38:53 Jacobson, A.-S.: Och då har det också varit kunders kunder, liksom i samma utsträckning som tidigare.

38:54 X : Nej. Alltså det man det? Tidigare. Ja ta igen. Vilken data man samlar in man samlar in så att säga. Extern data och hur man lagrar det och vad man sparar så vidare. Det är det vet jag fast inte så mycket om det skulle. Jag ska nog inte uttala mig om. Hur som helst för den här modellen så är det liksom kunddata som som kunden som kunden har så att säga. Var över de som har köpt på på på på webbsidan och.

39:40 Jacobson, A.-S.: Nej, men precis och det borde de ju. Det har ju mycket och känt. Då har det liksom via knappt mycket och liksom inne. Alltså inne i huvudet, liksom det. Det är liksom associationen liksom.

39:49 X : Ja ja. Precis de har de har de har godkänts och sen om det här med att man använder till maskininlärning. Analys. om Och det är någonting de är medvetna om eller om det är någonting som borde göras medveten om och om de blir medveten om de blir förändra beteendet.

40:06 Jacobson, A.-S.: just nu Ja, det är lite det pekar på liksom.

40:08 X : Det är ja precis.

40:10 Tony : ja

40:15 X : Ja jag.

40:17 Jacobson, A.-S.: Så länge man anonym anonymiserar det, liksom så kanske.

40:21 X : Ja precis alltså som sagt vad man man anonymiserade så långt långt det går. Men det har jag sådana gränser så att säga men se huvudpoängen är nog i alla fall att. Man bör nog. Precis som företag bör man då bör man med det liksom kunder om detta så att de är medvetna om det.

40:45 Jacobson, A.-S.: Vi fortsätter, det står lite han å jag hade instruktion. Är det något begrepp jag talas om?

40:53 X : Ja.

40:53 Jacobson, A.-S.: som kretsar Kretsar innebär det att en AI levererat svar

som låter som att den kan till 100% säkerhet, men det är inte kan. Men det ibland så ser man också begreppet, liksom bara som en generell kritik mot arg. Så det är också ett dilemma liksom om de om de om man inte vet svaret så kanske det är bättre att de säger det i stället för att ge falskt data. Hur mot vacker ni det liksom hos er är det någonting du stött på.

41:21 X : om hur På något sätt. Det här begreppet AI konsumtion. AI svaret AI presenterar ju den som den utifrån matematisk analys är högst sannolikt med högst sannolikhet är korrekt, men det är ju fortfarande bara en sannolikhet. och det är möjligt att man kanske den sannolikheten ska presenteras med något värde liksom att det är 99 eller 75 och så vidare

41:53 Tony : ja

41:55 X : Och det är möjligt att man kanske den sannolikheten ska presenteras med något värde, liksom att det är 99 eller 75 och så vidare. Men. Det, det är liksom ett sådant den presenterat fel svar och den vet om att det är fel svar, för då kan ju inte. Då presenterar ju inte det fel. Svar så att säga. Den illa den den ologiska. Den kan inte vara 100% säkert, men jag vet eller att den är 100% säkert är fel. Den har ju det är en sannolikhet och. Och det är kanske möjligt att att det är något riktvärde där så borde presenteras att med givet informationen vi har fått och nu har gjort. Kan vi 99% säkert säga att det här är rätt svar.

42:33 X : Och sen kan man då ta hänsyn till det? Det är möjligt att det är något som borde till, men. det är Är det de då måste man vara vara medveten om hur det här systemet fungerar så att det här är inga. Ingen absolut kunskap som som som dom skickar ut utan det är en sannolikhet.

42:55 Jacobson, A.-.S.: Nej, men precis vad vi står ju till exempel som jag använt den där. Den är väldigt bra på kod till exempel, för den är väldigt tränad på det. Men konkret fakta kan den ibland vara helt fel på, så det är ju också om ni har en mjukvara som är avsedd specifikt före i kommers. Så kan man väl tänka att. Bedöma trovärdighet utifrån det också.

43:18 X : Ja.

43:20 Jacobson, A.-.S.: hon Och inte bara kör kör lite liksom för det som du gör utan.

43:25 X : Nej, precis att det man har, man har precis, men det det man får en högre sannolikhet. Men det är ju i grunden fortfarande en sannolikhet.

43:33 Jacobson, A.-.S.: Precis, vill du fortsätta tony?

43:37 Tony : Yes. Jag har gått igenom nummer 11 adam, eller skulle säga att den är.

43:45 Jacobson, A.-.S.: Nej, 11, tänkte vi där söder att. Till protetisk scenario egentligen som vi pratade om lite innan att vi har en älleverantör som levererar en mjukvara till en konsult och sen hjälper konsulten en kund och så visar det sig att kunden gör ett business decision som är fel. Baserat på data alltså. Analysen var fel från början på grund av analysen. Hur ser du på ansvarstagandet gentemot de 3 parterna?

44:16 X : Han har tydligt kommunicerat att den analys, speciellt bra i. Och det blir fel resultat, då får man skylla sig själv för då måste man veta att vad är för någonting? Att det är ett sannolikhet och att det kan bli fel så att säga.

44:35 X : Men om man inte har om man som låtsas att det är en riktig person som har gjort den här och kommer fram till den slutsatsen och sen tänker bara fuskat med hjälp utav vad jag sett flest exempel så är det ju är det ju bättre bedrägeri så att säga och det det straffbart att säga så att om om kunden så att säga medveten om att Turkiet det här är en analys bakom en sajt jag ska göra på rätt sätt och jag sen går och gör det och det går käpprätt åt skogen, då får den det. Då får den kunde skylla sig själv.

45:06 X : Om den är om den är välmotiverad, därför att. Precis det här du har byggt någonting liksom. Du har informationen från färg och.

45:13 Jacobson, A.-.S.: Hemma.

45:16 X : Mycket mer.

45:17 Jacobson, A.-S.: Nej, men det perspektivet har faktiskt inte hört tidigare. Det var väldigt intressant.

45:18 Tony : ja Nej.

45:22 Jacobson, A.-S.: För det var vissa som pekar på, liksom att. Ja kunna skylla på konsulten konsulterna skylla på leverantören och så blir det inget av det liksom. Det är också ett scenario. Det är någon som människor att consultant ska ta allt ansvar, liksom i och med att de de tjänar ju som det är sagt tid och pengar på att använda i analysen så då har de också kanske ett ansvar i kvalitetssäkring och liksom att.

45:32 X : Ja.

45:48 Jacobson, A.-S.: Så det är en avvägning där.

45:48 X : Ja om de förser om de om de kommunicerar det till till till kunden så tror jag att det här är mjukvara som delvis är framtagen av av AI.

45:52 Tony : ja

45:58 X : Och. Det beror alltså på liksom vad man har. Precis som sagt var är man är man tydlig med det medvetna om det så så då då får man som lyssnar som kunde skydda sig själv men.

46:16 Jacobson, A.-S.: Men anser du att kunderna ska betala mindre? Då för de de de anställda ju konsulten för att de vill få en bra analys.

46:17 X : Det.

46:25 Jacobson, A.-S.: Så om den då AI baserad då? Borde ju liksom.

46:31 X : Alltså, det är ju. Det är ju om man om man. Nej, det tycker jag inte därför att de ändå kommer inte att vara det behöver jag inte kommentera. Men har man kommit överens med en kund till exempel så här att det här liksom att jag går till ett ett företag jag går till till till Zalando så att säga jag vill göra den nyss på er data med hjälp utav maskininlärning och ta sedan fram 10 rekommendationer.

46:57 X : Och de väljer att köra på dem om du om de sedan går och köper rätt åt skogen så har ju det ändå liksom. Jag varit tydlig med att jag har gjort arbetet och till och med processen bakom och därmed också tydligt att det här fungerar inte nödvändigtvis utan det är liksom det är sannolikhet.

47:13 Tony : ja

47:15 Jacobson, A.-S.: Men om det är en fortlöpande relation så att säga där konsulten från dag till dag byter till en AI baserad i stället.

47:15 X : Alltså.

47:24 Jacobson, A.-S.: Då?

47:25 X : Allt allting handlar om vad man har kommunicerat och vad man har kommit överens om. som Om kunden veta om att det är AI baserat och en medveten om riskerna så får kunden ta dom riskerna och det man betalar för egentligen en arbetsinsats som som konsulten gör. Men om om konsulten, inte. En del av detta och därmed ger skenet av att man som det det personer bakom som har gjort det här så att säga. Är det min uppfattning, bedrägeri och sen bara så länge det är tydligt och transparent så får man fatta beslut och ta risker utifrån det.

48:03 Jacobson, A.-S.: Sa konsulten. Om de sparar 2 timmar på en AI analys kontra tidigare ska de lägga de 2 timmarna på kvalitetssäkring då. Om du förstår min poäng?

48:19 X : OK om så att kunden har kommit överens om 8 timmars arbete.

48:23 Jacobson, A.-S.: Precis.

48:24 X : Du menar då att 2 timmar till det här sparar man in genom att använda däri.

48:27 Jacobson, A.-S.: Men behåll.

48:32 X : Om man. Ja, du ser du. Jag tycker du att det är viktigt att vara tydlig till kunderna säger liksom att jag har gjort. Det här arbetet, men jag har också använt arg och det är inblandat i den här i den här tiden.

48:44 Tony : ja

48:48 X : Och. Om man skulle säga har gjort 8 timmar tidigare har tagit betalt för 8 timmar och sen så är det helt plötsligt. Skiftar man då till att jobba 6 timmar bara och låta a utan de 2 timmar då menar jag att då måste man meddela det till till kunden så att de får ta och säga liksom att kunden får säga det och jag kunde. Men jag behövde betala lika mycket eller mindre eller mer till och med.

49:11 X : Så att det är.

49:11 Jacobson, A.-.S.: Namnet protest? Det var det vi det var lite efter liksom.

49:15 X : Ja att man är att alla alla vet vad som gäller och man är tydliga med det.

49:15 Jacobson, A.-.S.: Ja.

49:20 Tony : ja

49:24 Jacobson, A.-.S.: Till det facila tony eller egentligen sista frågan nu.

49:27 Tony : Liksom. Så det är det något mer du skulle vilja tillägga till. Den här diskussionen någonting som vi kanske har glömt att ta med.

49:43 X : Det är. Ja alltså, det är det håller en hög kvalitet då det exploderar då för hösten med med chatty. Jag har jobbat lite med. Sådana här saker för länge sen och de var ju så pass mycket. Fungera så pass mycket sämre. den här Så den har jag typ chapitre kom i förra hösten i november sen och blev väldigt imponerad utav hur den kunde samla all den informationen. Och och ge svar som var som väl var väl bara om man tänker sig vad vad vi vilka, vilket genomslag. Oj kommer få. Jag tror vissa vissa områden kommer att få jättemycket genomslag och det är just det här med generativ Harry. Att det går att spara jättemycket tid och kraft på att.

50:38 Jacobson, A.-.S.: Fotoshop på.

50:40 X : för Precis liksom jag tror till och med snart komplett se filmer som som är genom till varje liksom när det blir något av. hur hybridlösning att 4 skådespelare spelar in 4 5 scener tillsammans under en veckas tid och sen så gör gärna till genom tillvara i resten av filmen och en film kan bli klar på 4 veckor det tidigare tog 2 år att spela in en film Hybridlösning att 4 skådespelare spelar in 4 5 scener tillsammans under en veckas tid och sen så gör gärna till genom tillvara i resten av filmen och en film kan bli klar på 4 veckor. Det tidigare tog 2 år att spela in en film. Sådana saker. Vad gäller analys så tror jag redan vi har nått.

51:08 Tony : ja

51:13 X : Så datacenters analys liksom jag tror det är redan nått taket på något sätt om de kan göra för att. Ge de här systemen. Är inte kreativa på det sättet. Dom tänker nya saker att de får stå och liksom väldigt komplexa saker det de gör är att de kan räkna mycket snabbare. Kan kan så att säga, hitta, hitta, hitta, hitta, hitta samband men. Det är inte så att vi så att 1, 1 1, 1, 1 AI system kan snabbt lära sig hur man spelar schack på ett bra sätt. Men de kan inte uppfinna ett nytt schackspel. Är därför den typen av kreativitet. Vet vi fortfarande inte fungerar och den finns bara hos hos människor så länge.

51:56 Jacobson, A.-.S.: Nej, men det är intressant.

51:56 X : att det Det gäller liksom att veta det för mig, liksom inte att just liksom vilka områden kommer att påverka. Vilka kommer inte påverka så mycket?

52:06 Jacobson, A.-.S.: När vi hade 1, 1, 1 litet projekt han. Han var ju inte rolig för att e skulle jag sätta hans jobb.

52:07 X : Och.

52:14 Jacobson, A.-.S.: Och då var han design test medans dom. Andra pekar på att det kommer alltid krävas en humaninsulin upp. Det kommer alltid krävas riktig intelligens och inte bara difficile, så det är intressant. Men inte ens som liksom 50 år.

52:29 X : Precis liksom.

52:31 Jacobson, A.-.S.: Tror du att? Alltså taket kanske är lite lite högre än 50 år men inte på samma sätt som andra delar.

52:39 X : Det, det är ju det är ju.

52:41 Tony : ja

52:42 X : gör man Hur man pratar om det här? det man pratar om Att det finns en generalitet Alliance AG som liksom är det här ett eget medvetande och så vidare. Jag tror vi fortfarande är väldigt långt därifrån och jag tror att dom har dom har framsteg som har kommit inom mercutio framför allt. ja Inte kommer kommer leda till det för några, liksom sätt att. De här systemen kan är jättebra på att jobba med med med data, liksom med de enligt de är tankade mer eller mindre all data på internet. Och det man har gjort hittills är den vagn, den där man mindre mindre kan göra. En åtgärd han nästa steg att att göra uppfinningar och upptäckter och liksom.

53:23 X : är det återigen Ja vad vi vet ju fortfarande väldigt lite om hur vår hjärna fungerar. Och. Jag tror inte det bara är. Saknar kraft så att säga utan att det finns ett det det. Ja, jag tror vi är långt därifrån fortfarande.

53:51 Jacobson, A.-S.: Grymt då är sista frågan om du skulle vilja ha transcription. Av antalet ja dammet stoppa inspelningen också.

54:03 X : Tack. Absolut nej. Jag behöver förändra en. Lita på att ni inte?

54:09 Jacobson, A.-S.: Vill du ha med ditt namn och så i uppsatsen?

54:13 X : är

54:14 Jacobson, A.-S.: hon

54:14 X : Järna med den som intervjuar och få.

54:15 Jacobson, A.-S.: Ja.

54:18 X : Men hur hur tänker ni gå vidare nu med de här resultaten?

54:19 Jacobson, A.-S.: Runt. Mm. Ja, det kommande veckan kommer vi sitta ner med alla transcription is och väljer ut 5 6 teman ungefär och sen lite hur de har sett på det.

54:34 Tony : Exakt.

54:36 Jacobson, A.-S.: mm Exempelvis vi pratat om. Det här med ansvarstagande? Den här frågan vi hade mot slutet?

54:45 X : Ja.

54:46 Jacobson, A.-S.: Den har ni svarat väldigt olika på allihopa så den är väldigt intressant men uppsatsen är klar.

54:50 X : OK ja.

54:54 Jacobson, A.-S.: 10 maj någonstans där. Så du kan läsa den om du vill.

54:57 X : OK. Okej, kan ta ett ta ett x kollar på dem. Ja.

55:04 Jacobson, A.-S.: Använder stort tack. Det här var väldigt. Det gav oss väldigt mycket.

55:09 Tony : Verkligen.

Appendix I. AI statement

We used two AI tools while writing this bachelor's thesis. As non-native English speakers, we primarily utilized ChatGPT to improve our text, ensuring that we met the academic standard. This involved using synonyms for words that were considered repeated within the same chapter and rephrasing parts of certain sentences to improve the flow of the text. All sentences were written by us before using ChatGPT and we were careful to point out that no new content should be added to the text. We also used ChatGPT to help create a consistent template for all our interview transcripts since they were recorded on different platforms. The second tool used was Tactic which we used to transcribe all our interviews.

Since we used the ChatGPT tool for language improvement it has been used sporadically throughout the whole thesis. As we mentioned earlier, all the ideas and creation of the bachelor thesis has been done by us and the AI tools have strictly only been used to edit and improve our original ideas and content.