Introduction

Artificial Intelligence (AI) and Machine Learning (ML) have extended their reach far beyond the realm of technological innovation, now infiltrating diverse sectors of human activity. Of particular note is their permeation into academic research, which has seen a profound transformation through the application of these technologies. The inherent potential of AI and ML is immense, not only introducing innovative methods for enquiry but also opening up vistas for novel discoveries. One striking innovation is the Generative Pretrained Transformer 4 (GPT-4), an advanced Large Language Model (LLM) developed by OpenAI. This highly sophisticated language model offers a robust mechanism for processing and analysing substantial quantities of textual data, making it an invaluable asset in contemporary research.

The focus of this study is to explore and apply GPT-4 in the field of qualitative research, specifically within the realm of organisational behaviour in IT departments. The study is conducted from a single researcher's perspective. The research takes place in the IT department of Vejle Municipality in Denmark. Despite having a well-defined IT investment process, the department faces challenges related to employee compliance. These challenges have resulted in operational inefficiencies and inconsistencies, which this research aims to investigate. The objective of the study is to uncover the root causes of non-compliance using the power of GPT-4 to gain a deep and insightful understanding of the complexities involved. A unique aspect of this research is the use of GPT-4 to generate affinity diagrams. These visual representations of thematic connections capitalise on the LLM's superior capacity to process, synthesise, and identify interconnected themes within extensive text data.

The rapid evolution of digital technology continues to intersect with various aspects of human existence, leading to significant changes in traditional practises and the establishment of new paradigms. This transformative trend is particularly evident in academia, where advancements in AI and ML present unprecedented possibilities. This study aims to investigate one such possibility— the integration of generative language models, represented by GPT-4, into qualitative research methodologies.

This research scrutinises the IT department of Vejle Municipality, a local government unit dedicated to delivering key services to a dynamic community comprising over 120,000 residents (Vejle Kommune 2022). Despite a sturdy IT investment process in place, the department is plagued by challenges tied to

employee compliance, leading to operational inefficiencies. This investigation seeks to delve into the causes behind these compliance problems and put forth strategies for enhancing employee compliance.

GPT-4 is interwoven into this research in two specific ways. Firstly, it functions as a supporting tool within traditional qualitative research, offering instantaneous assistance and insights, behaving as a constant academic helper. Secondly, I venture into a more probing approach by harnessing the capabilities of this leading-edge technology to accelerate and enrich the qualitative analysis process. The goal is to automate the creation of an affinity diagram from interview data, thereby teasing out novel insights from the massive dataset. For this, I resort to a bicameral dialogue approach, emulating a bilateral conversation akin to a parliamentary discourse, prompting the model's self-dialogue. Nevertheless, despite the innovative bicameral dialogue approach, the resulting diagrams fell short of the initial expectations. Consequently, it is advisable to conduct further research to explore the constraints and potential uses of GPT-4 in the realm of qualitative research.

Through a meticulous analysis of the interaction between technology and human behaviour, this study seeks to offer remedies for the challenges experienced by Vejle Municipality, while concurrently illuminating new paths for academic inquiry. The present research contributes to the comprehension of a unique socio-technical system within Vejle Municipality and investigates the potential of LLMs as groundbreaking tools in academic research. This investigation is positioned at the thrilling crossroads of technological innovation and human complexity, aspiring to enrich both spheres with novel insights and strategies.

GPT-4's extraordinary capability to process and produce insightful analyses from extensive text input has been harnessed in this study. With the 8K model capable of handling up to 8,000 tokens and the 32K model remarkably processing up to 32,000 tokens, where each token represents approximately 0.75 words in English text (OpenAI Pricing), GPT-4 stands out as a potent tool for data analysis.

Hence, this exploration has a dual purpose. Firstly, it clarifies the socio-technical system within Vejle Municipality and the specific dynamics impacting the IT investment process. Secondly, it exposes the potential of AI and GPT-4 as pioneering tools in academic research, providing an original perspective on qualitative studies. Consequently, this investigation is steered by two interconnected research questions:

1. How does GPT-4's capability to generate affinity diagrams from interview data compare to that of a human researcher, and what insights can this provide for qualitative research methodologies?
2. What are the key factors contributing to employees' non-compliance with the IT investment process in Vejle Municipality, and how can these factors be

addressed to improve adherence?

These research questions aim to comprehensively evaluate GPT-4 as an innovative tool for qualitative research while providing deeper insights into the dynamics of the IT investment process in Vejle Municipality.

Relevance of Research

This research holds significant value from both a pragmatic and scholarly perspective. From a practical viewpoint, the results of this study aim to improve the efficiency and adherence to guidelines within the IT investment process in Vejle Municipality. The aim here is to provide tangible recommendations and insights to better navigate the complexities inherent in municipal IT investment decisions. By doing so, it will help facilitate a more sustainable, robust, and accountable IT ecosystem within the context of Danish municipalities.

From an academic standpoint, the study explores the potential of GPT-4 as a novel tool in qualitative research. Despite the increasing recognition of AI's transformative potential across various fields, there is a lack of scholarly research specifically focused on the application of GPT-4 in qualitative analysis at present. This research will contribute significantly to the ongoing discourse about the integration of AI, specifically LLMs, into academic research practises.

Furthermore, to the best of the author's knowledge, there is no comprehensive study that explicitly investigates IT investment processes within the specific context of Danish municipalities, making this research both timely and necessary.

The fusion of these perspectives underscores the profound relevance of the research, with immediate practical implications while potentially shaping the future course of academic research methodologies. It paves the way for a broadened scope of investigation into the role LLMs like GPT-4 could play in qualitative research, raising intriguing questions and setting the stage for future explorations in this emergent and promising field of study.

Overview of Structure

This thesis consists of two parts, each addressing a specific research question. I will work my way backwards, with the first part focusing on a traditional qualitative investigation of the IT Investment Process in Vejle Municipality. The second part evaluates the utility of GPT-4 in the context of the aforementioned qualitative research, specifically focusing on prompting the model to generate affinity diagrams from interview data.

The first section begins with an introduction to the case background, followed by a literature review providing the theoretical context. The research methodology involves semi-structured interviews and field notes for data collection, with affinity diagrams used for data analysis. The results, discussion, and recommendations form the subsequent segments of this part of the study.

The second section focuses on GPT-4, starting with a technical background for those unfamiliar with the model. This is followed by a methodology section, detailing data collection through user-model interfaces, details about the iterative prompting approach used, and establishes a set of criteria for assessing output quality, forming the basis of data analysis.

Case Study: Vejle Municipality's IT Investment Process

In this case study, I will present a comprehensive qualitative analysis of the IT investment process in Vejle Municipality from the perspective of a single researcher. The investigation will commence with a detailed case background, providing insights into the organisational structure of Vejle Municipality, the composition and function of its IT department, and the mechanics of its IT investment process, among other details.

Next, I will conduct a literature review to encapsulate the existing body of knowledge that forms the basis for this study. By exploring previous scholarship, I can position my current research within a broader academic context and shed light on the novel aspects of this investigation.

After the literature review, I will present the collected data, providing detailed information on its relevance and potential implications for Vejle Municipality's IT investment process. I will meticulously examine this evidence to ensure it supports the study's objectives and assumptions.

The chapter will conclude with a comprehensive discussion, synthesising my findings and highlighting their significance within the framework of my broader research goals. This concluding section will effectively tie together the various elements of the chapter, offering a cohesive analysis of the IT investment process within Vejle Municipality.

Case Background

Residing within Vejle Municipality's administrative building, the IT department has historically faced challenges in effective communication, particularly with external stakeholders and individuals in decentralised roles. Recognising the

need to bridge this gap, the department strategically incorporated four IT architects into its ranks over recent years, each bringing a distinct blend of skills and expertise.

These IT architects, integral to the municipality's IT ecosystem, strive to reconcile municipal employees' expectations with the capabilities and resources of the IT department. They deftly navigate technical and non-technical perspectives within the municipality, balancing the need for clear communication with the complex understanding required of their role. A blend of "hard" and "soft" skills allows them to empathise with the diverse needs of users throughout the municipality while staying conversant with current and emergent technologies.

Among these architects are Jonas and Mathilde, who together provide a harmonious blend of technical and human-focused insights. Jonas excels in the intricate nuances of IT architecture, while Mathilde advocates for a more empathetic, human-centreed approach to IT. Their collaborative efforts ensure the fluid integration of new technologies and foster a sense of shared understanding and ownership in IT investment decisions across the organisation.

Jonas and Mathilde's invaluable guidance and profound knowledge have significantly shaped the course of the present study. Their generosity in sharing their insights is reflected in the multiple references within the forthcoming text.

In response to the challenges in procuring new technologies, Vejle Municipality has implemented an IT investment process, facilitated through an online webform accessible via the intranet. The webform serves a dual purpose: keeping the IT department aware of which technologies the organisation works with and connecting appropriate IT personnel with the investment project for smooth implementation and potential impact evaluation.

The user interface of the webform presents a clear directive to all employees of the municipality:

> "If you need to address a new IT task, update or enhance an existing system, have an idea for an IT system, or purchase a new IT system, the City Council requires you to utilise the IT investment process. By following this process, we collectively ensure that the system or function can be utilised while meeting our IT requirements. Additionally, when you follow the investment process, you will receive guidance to other departments if there are specific areas that require attention, such as information security."

A list of examples provided clarifies the instances in which the IT investment process should be invoked. These scenarios include exploring existing solutions, making changes to current systems, seeking IT support for a process, purchasing new equipment requiring network access, or implementing a new IT

solution.

To engage in this process, employees are required to complete an electronic form on the designated intranet page. After submitting the form, an IT department architect typically contacts the submitter within 48 hours, initiating a dialogue involving the employee, the supplier, and an IT department representative. This dialogue ultimately leads to feedback that includes an overall IT assessment of the system. If required, additional references to other departments within Vejle Municipality, such as Competition Suspension, are provided.

However, despite these comprehensive instructions and the broad applicability of the webform, its usage among employees remains limited. Significant efforts have been made to simplify the webform, making it more accessible by removing intricate questions that non-IT personnel found challenging. The primary goal was to increase the form's usage and prevent staff from directly contacting IT architects or other members of the IT department, thereby streamlining the IT investment process. However, these efforts have not resulted in the expected increase in webform usage.

In light of these observations, this study aims to delve into the reasons behind employees' inconsistent adherence to the IT investment process and propose strategies to improve compliance. The ensuing report will thoroughly investigate the complexities of this case before using the rich qualitative data generated for experiments in the field of NLP in the penultimate chapter.

Literature Review

This literature review will delve into three key areas pertinent to this study: IT Investment Governance, Ethnomethodology, and Affinity Diagrams. The analysis of these topics aims to build a theoretical underpinning for understanding the complexities of IT investment in the public sector and how ethnomethodological research and affinity diagrams can help derive valuable insights into this intricate process. Let's begin by unpacking the first concept: Information Technology Investment Governance (ITIG).

IT Investment Governance

IT governance is a vital component for both private and public sector organisations, playing a central role in aligning IT investments with institutional goals and maximising associated benefits (Weill, 2004). Encompassing decision-making, planning, and controlling IT investments, it sits at the heart of effective governance.

One influential concept for understanding IT governance is the resource-based theory (RBV), proposed by Wernerfelt (1984). RBV suggests that organisations can secure a sustainable advantage by optimally employing their resources and capabilities, those possessing valuable, rare, imperfectly imitable, and non-substitutable (VRIN) attributes (Wernerfelt, 1984; Barney, 1991). RBV underlines the strategic necessity of resource allocation for maximum value.

When considering RBV in public sector contexts, such as Vejle Municipality, the idea of a "competitive advantage" is recontextualised. Unlike private firms, public institutions aim to deliver the best possible public value rather than outperform market competitors. Here, the efficient allocation and utilisation of resources equate to high-quality service delivery. In the realm of IT investment, strategic decisions should maximise public benefit through increased efficiency and effectiveness.

Building upon the RBV, Bharadwaj (2000) positions IT resources as a distinct strategic category. organisations, he argues, possessing superior IT capabilities, can enhance their operational efficiency and service quality (Bharadwaj, 2000). This perspective fits seamlessly within the RBV, reinforcing the critical need for effective governance of IT resources and investments.

Ali et al. (2015) developed a novel construct, known as IT investment governance (ITIG), also premised on the RBV. ITIG gauges an organisation's competence in effectively governing its IT investments, thereby enabling the efficient allocation and utilisation of IT resources. The construct comprises four elements: IT investment value governance, IT investment value monitoring, IT investment appraisals, and IT investment project management.

The relevance and efficacy of the ITIG construct were demonstrated by Ali et al. (2015), highlighting a significant positive correlation with organisational performance. This critical finding implies organisations with robust IT investment governance are poised to extract greater value from their IT investments and align these more closely with their strategic goals. Consequently, effective IT investment governance enables organisations to amplify operational efficiency and public value, thus reinforcing the core principles of the RBV.

While Ali et al. (2015) made notable strides, certain gaps still persist, particularly in understanding the behavioural elements influencing effective IT investment governance, such as employee compliance with IT governance processes. Nonetheless, the RBV and ITIG construct offer a robust theoretical foundation to scrutinise and enhance IT investment governance within the public sector, paving the way for future research and improvements.

Ethnomethodology

The field of social science research has seen a significant surge in the application of Ethnomethodology (EM) as a theoretical approach. EM explores how individuals establish and uphold social order through their everyday interactions, focusing on practical methods and techniques they utilise to comprehend their social realm. As observed by Crabtree (2012), EM research primarily centres on how communities or social settings employ quotidian practises and knowledge to formulate the regulations governing their social existence (p. 316).

EM's wide-ranging application spans diverse fields such as sociology, anthropology, communication studies, with more recent integration into human-computer interaction and interaction design studies. It is instrumental in examining how individuals engage with technology in daily life and how technological advancements impact and shape social practises.

Crabtree, Rouncefield, and Tolmie (2012) assert the value of EM in deciphering technology utilisation across various settings. They suggest that the methodology's focus on minutiae of day-to-day practises can aid researchers in recognising practical challenges encountered by individuals while using technology, and how design can alleviate these issues. Moreover, they posit that EM's emphasis on the social context of technological use can provide insights into how technology integrates into broader social structures and influences social relations.

Interviewing: A Word of Caution

Interviews, according to Crabtree, Rouncefield, and Tolmie (2012), should be approached with prudence. They propose conducting interviews within the authentic flow of work as circumstances allow. Highlighting the discrepancy between what people claim to do and their actual actions, they note that while not being intentionally misleading, interviewees tend to gloss over their work. To overcome this, researchers are advised to focus on the precise actions and methods of work, avoiding an interview format driven by a set of pre-formulated questions detached from the actual work process.

The Importance of Field Notes

The creation of field notes is a crucial component of fieldwork, state Crabtree, Rouncefield, and Tolmie (2012). Field notes allow researchers to document their observations, capturing what they witness, hear, and are informed of. This not only provides a record of the research but also assists researchers in tracking and organising their thoughts. The active process of note-taking encourages the researcher's attention to the work as it occurs, thereby fostering a better understanding of the task at hand.

The researchers' notebook, Crabtree, Rouncefield, and Tolmie (2012) suggest, should not be a mere collection of disjointed remarks. Instead, it should serve as a structure for the researcher's thoughts, crafting a coherent narrative of the setting's work. Diagrams of the work ecology are recommended to frame inquiries and represent the work environment to others. Detailing the environment, the inhabitants, their roles or responsibilities, and the artefacts employed for work can lead to a more comprehensive understanding of the setting's work and the methods employed by members to organise it.

Formal Organisation of Work and Its Flow

Understanding how the work is 'formally organised' across a division of labour and individually, is crucial, according to Crabtree, Rouncefield, and Tolmie (2012). This comprises the plans, procedures, processes, and routines invoked by the members to account for their work organisation. Researchers should also concentrate on the work's flow, noting how it moves across individuals, activities, and culminates in an endpoint.

Discrete Sequences of Interactional Work, Cooperation, and Collaboration

Crabtree, Rouncefield, and Tolmie (2012) encourage researchers to enrich their depiction of work's flow by concentrating on discrete sequences of interactional work entailed in completing specific activities. The focus should be on what is done, who does it, and the process of work completion. Attention should also be given to the cooperation and collaboration between individuals during discrete sequences of interactional work. This would involve observing and documenting who interacts with whom, their discussion points, joint activities, task transactions and handovers, and subsequent responses.

Affinity Diagrams

Affinity diagrams, also known as the KJ method or affinity charting, were first developed by Japanese anthropologist Jiro Kawakita (as cited in Scupin 1997). Affinity diagrams are used to synthesise and categorise large amounts of qualitative data, such as observations, interviews, and field notes, into meaningful and easily understandable themes and patterns (Hanington and Martin 2019).

The affinity diagramming process typically starts with the raw data being transformed into discrete statements or observations. These statements are then grouped based on their similarities and relationships (Hanington and Martin 2019). The groups are subsequently labelled with descriptive headings, which capture the essence of their content. This iterative process allows for the

identification of patterns, themes, and relationships among the collected data, thus providing insights and guidance for further analysis and design Holtzblatt and Beyer (2016).

In the context of ethnographic UX studies, affinity diagrams serve as a valuable tool for making sense of the complex and often messy data that emerges from immersive fieldwork Holtzblatt and Beyer (2016). By organising and categorising data in a structured manner, researchers can identify user needs, behaviours, and pain points, which can inform design decisions and enhance the overall user experience Hanington and Martin (2019).

Furthermore, affinity diagrams facilitate collaboration and interdisciplinary communication among research team members Holtzblatt and Beyer (2016). By engaging in the process of grouping and labelling data, researchers from different backgrounds and expertise can contribute to a shared understanding of the user experience, leading to more innovative and effective solutions.

Methodology

This section delineates the research methodology undertaken in this study, with a detailed examination of both data collection and analysis processes. The data collection encompasses two salient approaches: semi-structured interviews and field note documentation. Following this, the data analysis strategy is explicated, focusing on affinity diagrams and the employment of GPT models as a peripheral method. These methodologies provide a synergistic approach, marrying qualitative and quantitative aspects to ensure an all-encompassing understanding of the research context. This thorough exposition of the processes involved serves to elucidate the rigour and thoughtfulness in the methodological design of the study. Furthermore, it underscores the comprehensive trajectory from data procurement to analysis, illuminating the rigorous approach taken to ensure the validity and reliability of the research findings.

Data Collection: Interviews & Field Notes

In this study, the process of data collection was established upon two critical components: semi-structured interviews and field notes documentation.

Semi-Structured Interviews: This data collection methodology was a primary channel to gain insight into the users' experiences and perspectives. As suggested by Crabtree, Rouncefield, and Tolmie (2012), the interview setting was informal to encourage genuine and in-depth discussion. The interviews were designed to be "decontextualised" due to the nature of the study, where the subject matter (new IT investments) was not part of the daily routine for most employees.

A selection of twelve employees across all six administrative divisions of Vejle's municipality was interviewed based on a list provided by an internal contact at the municipality. To maintain confidentiality, each interviewee was assigned a pseudonym and an identification number.

The interviews were conducted face-to-face, barring one instance where a video call was arranged to accommodate the participant's circumstances. Handwritten notes were taken during the conversations to ensure a relaxed atmosphere, and the questions were based loosely on an interview guide (Appendix A), allowing flexibility for natural discourse.

Following each interview, the brief notes were used to create detailed transcriptions, which were then developed into coherent prose summaries. It is important to note that the interviews were conducted in Danish, and the translation into English occurred at the summary phase. The complete interview summaries and notes are available for reference in Appendix B.

Field Notes: As a complementary method, field notes played an essential role in the data collection process. Drawing from the significance emphasised by Crabtree, Rouncefield, and Tolmie (2012), field notes were used to document observations, thoughts, and insights about the research setting.

A physical notebook was consistently maintained for documenting fieldwork observations, particularly during interviews, to avoid the intrusion of digital devices. Additionally, a cloud-synced notes application served as a supplemental tool for situations where typing was preferable to handwriting.

The process of taking field notes enabled a comprehensive understanding of Vejle's work environment, shedding light on specific dynamics such as task handovers and transactions between individuals. Additionally, the field notes acted as a reflective tool, aiding in evolving an understanding of the work setting over time.

In sum, the combination of semi-structured interviews and field notes documentation provided a comprehensive and valuable resource. The detailed accounts served as a solid foundation for the analysis and interpretation of the data, effectively informing the subsequent phases of this research study.

Data Analysis: Affinity Diagrams

Following the process of collecting interview summaries and field notes, an affinity diagramming technique was employed to synthesise and analyse the data. This method provided an effective way to organise the large volume of collected data, identify patterns, and uncover underlying themes.

The affinity diagramming was conducted virtually using a computer programme. Each interview summary was reviewed one at a time, and data segments were captured on virtual sticky notes. Each interviewee was assigned a unique colour for their sticky notes, thereby enabling easy identification of their respective insights. However, due to the limited palette, the colour differentiation ended once the range was exhausted, and the remaining sticky notes were kept in grey.

As the process unfolded, these sticky notes started forming natural clusters of related information based on the similarities and patterns that emerged from the data. This process allowed for the visual grouping of insights according to their inherent relationships, providing a high-level view of the data landscape.

Additionally, lines were drawn between these clusters to illustrate the connections and dependencies between different groups of data, making the relationships explicit. This process enabled a more profound understanding of the data and facilitated the identification of overarching themes and patterns.

Throughout this process, the affinity diagram was consistently reviewed and iteratively adjusted to ensure that the emerging patterns and relationships accurately reflected the data. The diagram served as an evolving visual representation of the collected data, allowing the research findings to be presented in a comprehensive, organised, and accessible manner.

In conclusion, the affinity diagramming method was instrumental in providing a structured approach to data analysis. It facilitated a holistic view of the data collected through semi-structured interviews and field notes, allowing for a nuanced understanding of the research context, which in turn informed the interpretation of the research findings.

Data Analysis: GPT Models

While the data analysis of the present chapter primarily relied on creating a traditional affinity diagram, the LLMs GPT-4 and its predecessor GPT-3.5 served as valuable supplementary tools, aiding the author in the research process. Their inclusion offered a unique opportunity to enhance the breadth of information accessed beyond what was personally known or available in the immediate context of the study. The LLMs were primarily utilised for their conversational abilities, providing an interactive platform to generate diverse perspectives on the research topic. Drawing from an extensive text corpus, LLMs offered insightful responses to the author's queries, often sparking new lines of thought, novel questions, and unexplored research avenues.

Notably, the LLMs did not merely function as an easily accessible knowledge resource, but they also catalysed critical thinking and intellectual curiosity. They served as a sounding board for the author's ideas, providing a space for in-depth

exploration of varying perspectives, encouraging a deeper contemplation of the research. Furthermore, their extensive knowledge base, available on-demand, supported the author in making real-time decisions on research matters.

While LLMs can serve as a source of knowledge, it is always worth noting that these models can also output false information, so all information attained from the models was double-checked using traditional information searching methods.

Presentation of Interview Data

The data gathered for this study comprises semi-structured interviews with twelve employees of Vejle Municipality, spanning various roles across all six administrative divisions. The data collection process involved taking rough notes by hand during the interviews, followed by more refined digital notes afterwards. From these refined notes, I created two affinity diagrams: one by hand and the other with the assistance of GPT-4 through a carefully crafted prompt.

The analysis in this report primarily utilises my hand-made affinity diagram, as it encompasses not only the insights gleaned from my notes but also the nuances of my experiences and observations within Vejle Municipality.

Out of the interviewees, the majority were familiar with the IT investment process: six indicated familiarity (including one who had helped create the process in her role as a digitalisation consultant), five were unfamiliar, and the final participant, who was also unfamiliar, deemed the process irrelevant to her work and was therefore excluded from the analysis.

Perceptions of an Impersonal and Bureaucratic IT Investment Process

My first interview was with Freja, a seasoned web manager from the Policy, Analysis & Communication department. During our discussion, Freja expressed strong opinions on the IT investment process, perceiving it as excessively bureaucratic and impersonal. With twenty years of experience working closely with the IT department, she felt confident in her ability to communicate directly with them when necessary, bypassing the formal process. Freja advocated for a more relationship-based approach, emphasising the importance of positive reinforcement over punishment.

Freja's perspective laid the groundwork for my research, as the issue of the IT investment process being seen as impersonal and bureaucratic recurred in subsequent interviews. Her experience and insights into the challenges of collaborating with the IT department were invaluable, and her close connection to the department provided a unique viewpoint that illuminated broader issues surrounding the process.

Lise, an administrative assistant in the Daycare department, acts as a vital bridge between the Administration Building and daycare workers, who have distinct needs due to their more practical roles. Lise expresses concern and dismay at the prospect of using a form to contact the IT department. She values her personal relationship with Jonas from IT, as he understands the unique context of daycare workers. Lise worries that submitting a form may result in losing "unspoken considerations" arising from the social and educational differences between the two groups.

To alleviate these concerns, Lise suggests adding an option on the form to indicate if the case involves a target group struggling to adapt to technological changes. Another appealing solution she proposes is including a preferred IT department contact person on the form, ensuring that someone familiar with the specific context is involved.

Benefits of Using the Process: Potential Talking Points

Claus, a development manager at Vejle's Centre for Special Education for Youth and Adults (CSV), finds the IT investment process helpful because it connects them with the right people, such as an architect who oversees the case, ensuring it is remembered and acted upon. Using the IT investment process also provides a clear line of communication when working decentralised, as is the case with CSV. It allows them to follow a case in the system and see which IT department members are involved and what they are doing. Claus appreciates the ability to intervene if he feels his needs has been misunderstood.

Potential for Strengthened Cooperation Between IT and the Rest of the Organisation

Emma, who has been working in the municipality for thirteen years and took over her current position as an Implementation Consultant in the Department of Welfare six months ago, told about welfare's own work and how they have a similar form to the IT investment process for welfare investments. She suggested that it would be more efficient if their applicants could automatically involve IT in projects related to IT.

Another example of potential strengthened cooperation comes from Anne, a professional consultant for the Salary department and primary contact point for KMD, the municipality's payroll system. Anne shares a similar attitude with the IT department, wishing that employees in the organisation would be better at consulting with her before purchasing systems that fall within her area of responsibility, rather than buying the system first and then consulting with her about its implementation afterward. This attitude may be why she has a good relationship with the IT department, consulting with them well in advance.

Anne recalls an instance when she was called to a meeting by someone in the organisation, along with Jonas from the IT department, thinking that she 'might' have a stake in the implementation of their new system. She described the meeting as, "It was a video meeting, but I could sense that Jonas and I were looking at each other on the screen, thinking, 'is this really happening right now.'" This was because the third person had already purchased the system without asking and wanted it up and running in three weeks, which she found very unreasonable. She tells that such situations have at times created a tremendous workload in their wake. Thus, although Anne and Jonas work in different professional areas, they share some frustrations. A central frustration is that employees' expectations for the timeline of implementing new things do not match their own.

During the interview, I suggested the idea of merging the systems, and Anne reacted positively to this idea, agreeing that implementing a central system that directs communication based on certain guidelines could help ensure that employees with new investments in the pipeline do not have to guess who should be involved and when in the process they should be involved.

Time Requirements for Implementation of New Solutions

As touched on earlier, Anne from the Salary department has grappled with some employees in the organisation who were eager to introduce new solutions, with expectations regarding the timeframe for these implementations that, in her view, were unrealistic and lacked sensitivity to the accompanying workload. At present, there are no formal guidelines laying out the expected duration for the

implementation of new solutions. This void likely stems from the inherently diverse nature of IT investments, making the creation of universal guidelines a challenging task.

The Name Perceived as Opaque and Unattractive

Despite not being familiar with the IT investment process prior to our interview, Emma recalls hearing about it from Jonas, one of the IT architects, at a meeting. Emma assumed the process was primarily relevant to IT and economy personnel due to its name. She agreed that a more tangible name, such as "Indkøbsguiden" (Danish for "The Purchase Guide"), might make it more accessible to the grassroots level of the organisation. However, according to Jonas, a similar system in another municipality also faced issues with lack of use, indicating a more attractive name cannot stand alone.

IT Investment Process Not Widely Promoted

Interviewees such as Anne have a view of the IT investment process not being widely promoted or talked about by the personnel in the IT department. She has been used to just contacting Nicklas or one of IT's other architects when she needed something.

Broad Guidelines for When Reporting is Required

Some interviewees such as Mikkel from the Vejle's central library's IT department shared a feeling of being unsure as to when it actually was necessary to use the IT investment process, as the guidelines laid out on the Vejle's intranet cover a very broad area of investments. This is especially problematic for such central institutions that have the resources to implement new investments themselves without involving central IT, with an underlying feeling of participants wanting to be able to avoid using the process if it is unnecessary, despite the narrative at central IT being, if you are unsure, still use it to be safe.

Central vs. Decentralised

Søren, responsible for purchasing IT equipment for schools, explained that, while employees of the municipality are met by the municipal intranet as the first thing when they log onto their computers, employees in schools have to specfically browse to the intranet's address and log in to get access, making an extra barrier to entry to the IT investment process for them. Søren is nevertheless hopeful that, while school employees currently don't use the IT investment process, that it would be possible to get them to use it with the appropriate training, despite the fact that IT investments aren't a part of the everyday work of school employees. Søren went on to say that he thinks it is important that passionate school

teachers have the adequate freedom to purchase unique IT equipment to practise their unique styles of teaching, as he seems to view this as an important part of a healthy school system.

Peter, a care assistant who recently retired but still works at the residential care facility one day a week to help with their IT, raised concerns about the feasibility of relying on support from Vejle's central IT department in decentralised areas. He wonders what will happen to the facility when he fully retires, as they will have to rely on support from central IT, which may not be viable considering the facility's location and the fact that they operate year-round, providing constant care to residents.

Frustrations With Restrictions From IT

During the interviews, it became apparent that some interviewees, such as Claus, were frustrated with the limitations imposed by the IT department. Claus expressed his desire for more flexibility and autonomy in updating programmes on his computer without having to consult with central IT. Additionally, Claus hoped for more collaboration between Søren from school IT and central IT, as these areas are becoming increasingly interconnected.

Another interviewee, Peter, expressed frustration with the lack of flexibility in smartphone models available. He believed that investing in the phones recommended by the supplier of their systems would be more beneficial, as they come with a guarantee of receiving updates for up to eight years after purchase. However, IT disagreed with his reasoning, claiming that he was influenced by the supplier. Peter's argument was that the phones offered by the IT department had to be replaced after a shorter lifespan due to not receiving necessary security updates from the manufacturer.

Importance of IT Architects as Communicators

Many interviewees, such as Søren from school IT, acknowledged the importance of IT architects as communicators. Søren notes that the arrival of IT architects in the organisation has improved communication and collaboration between central IT and decentralised areas, as well as between IT and the rest of the organisation.

Claus went so far as to call the IT architects a "godsend," giving IT a more service-oriented outward face. Freja touched on that IT, through the years, has attained a greater range of skills than were available before, both in technical expertise but also particularly in the area of soft skills such as communication, nodding to the hiring of IT architects within the past five to ten years.

The architects, including Mathilde, have reached out more to decentralised areas in an attempt to build better relationships. Both Søren and Mikkel mentioned Mathilde's joint meetings, describing them as very helpful in the process of making IT's work more accessible throughout the organisation.

Current Wins for the IT Investment Process

While the interview data yielded a fair deal of critique to the process, it is not all bad. The web manager Freja expressed that, in her experience, that she thinks people have become more aware of the need to avoid double investments through the years. Søren from school IT described the simplification the IT investment process underwent some years ago as a considerable improvement, taking more consideration for decentralised employees.

Discussion

Upon analysing the data from the interviews conducted within Vejle Municipality, it becomes clear that navigating the IT investment process is a complex and challenging task. These interviews spanned a diverse range of roles within the Municipality, painting a comprehensive picture of how the organisation perceives and interacts with the IT investment process.

A key revelation from these findings is the perception of the IT investment process as bureaucratic and inaccessible. This perception affects the process's acceptance and effectiveness. Many employees, especially those whose roles are not directly linked to IT, express a hesitancy towards engaging with it. This sentiment highlights a noticeable divide between the central IT department and the remaining sectors of the Municipality, which has exhibited signs of improvement in recent years but remains in need of further attention.

Moving forward, I will draw from these insights to suggest practical and focused recommendations for improving the IT investment process within Vejle Municipality. These suggestions concentrate on refining communication methods, making the process more relatable, and fostering enhanced collaboration within the organisation.

In line with the recommendations, I plan to delineate potential directions for further investigation. I propose extending the breadth of this research to incorporate insights from the field of management science, findings from various other municipalities, and the application of ethnomethodology to enhance my comprehension of the IT investment process.

Through this exhaustive exploration of Vejle Municipality's IT investment process, my objective is to catalyse enhancements that will benefit not only the Municipality, but also broaden the understanding of IT investment processes in

analogous situations.

Recommendations for Vejle Municipality

Based on the data analysis and understanding of the current IT investment process at Vejle Municipality, it's clear there are areas of opportunity for improvement. The recommendations for the Municipality centre around five major themes: (1) enhancing communication and visibility of the IT investment process, (2) setting an example in reporting responsibilities and implementation timelines, (3) simplifying the introductory text and revising guidelines, (4) humanising the IT investment process, and (5) unifying reporting across the organisation and easing restrictions where applicable.

The suggestions emphasise a more user-friendly approach that bridges the gap between the central IT department and the decentralised units, making the process more accessible and efficient. These steps are intended to foster an environment of collaboration, understanding, and efficient execution of the IT investment process across the Municipality. Let's delve into the specifics of each recommendation.

Enhance Communication and Visibility of the IT Investment Process

To better communicate the IT investment process, it should be positioned as a support service rather than a bureaucratic obstacle. Adding an option to indicate preferred contact persons in the IT department could help alleviate concerns about the process's impersonal nature. Moreover, prioritising improved communication and collaboration between central IT and decentralised areas, and between IT and the rest of the organisation, is essential. IT architects have made progress in this area, but continued efforts are crucial for effectively integrating the IT investment process into Vejle Municipality's daily operations.

Establish Clearer Guidelines on Reporting Responsibilities

IT architects play a vital role in communicating the process through personal interactions and other means. As Anne, an interviewee, highlighted, sharing the process with the right people without bothering irrelevant employees is essential. While the intranet contains ample information on when investments must be reported, it lacks clarity on who is specifically responsible for reporting investments to IT. This ambiguity can lead to situations where no one takes responsibility for contacting IT.

Clearer guidelines for reporting responsibilities could help remedy this problem and make it easier for architects to find and communicate with the appropriate individuals within the municipality.

Address implementation time and workload concerns

As reported in the interview with Anne from the Salary department, instances have occurred where employees approach IT and other stakeholders with unrealistic expectations regarding the timeframe for solution implementation. Such situations can lead to disappointment and/or impose an excessively heavy workload to meet the desired schedule. It is therefore crucial to focus on the time requirements of projects. If feasible, guidelines should be established to provide employees with a realistic estimate of the time it will take to implement their IT investments.

Alternatively, or in addition, the IT investment process form could incorporate an option for indicating a preferred implementation date for the new IT investment. This provision would allow the IT architect to manage expectations right from the first meeting, helping to realign any unrealistic timelines. Consequently, the IT architect would be less likely to be taken aback by an impractical timeframe presented during the meeting. This strategy could also mitigate the element of surprise and reduce the likelihood of a burdensome workload being imposed at short notice.

Set an Example

When employees contact IT architects for assistance, they should receive help with their IT investment. Simultaneously, IT architects should create a record of the new IT investment in the appropriate database, demonstrating mutual respect for the process. After logging the investment, IT architects could inform the employee and politely request that they use the IT investment process for future investments, ensuring to explain the process's purpose.

Simplifying the Introductory Text

IT should consider the presentation of the IT investment process on the intranet page further. The extensive technical text on the page should be relegated to less prominent areas, as first impressions are crucial in persuading employees to engage with the process. Instead, a more concise, simplified version should be presented, emphasising the rationale for the process's existence and the advantages of utilising it, ideally employing more engaging communication methods.

For instance, a video that explains the process, its purpose, and its benefits could be showcased on the intranet page, making it a valuable resource to which IT architects can direct employees for further information.

By making employees aware of the rationale behind the IT investment process and its potential benefits, they are more likely to be receptive to changing their behaviour. Simultaneously, understanding the logic underpinning the IT investment process provides a more intuitive way of explaining and grasping its relevance. This approach could help address the image problem and confusion arising from the current guidelines, as mentioned in the previous section.

Consider Revising Guidelines for When Investments Should be Reported

For IT-savvy employees outside of the IT department, such as Mikkel and Freja, the guidelines regarding when IT investments need to be reported can appear overly broad, seemingly intended for employees whose daily work and competencies do not focus on IT. Consequently, these IT-savvy employees may feel that they are exempt from using the IT investment process or perceive that employees with their competencies have not been taken into account when formulating the guidelines.

Although the current guidelines serve as a comprehensive catch-all, which can be suitable when the complexity of a situation makes it difficult to establish precise rules, the IT department should be cognisant of the image problem and the confusion generated by prominently displaying these guidelines on the intranet page for the IT investment process. This prominence can overwhelm readers or appear condescending.

Humanise the IT Investment Process

Interview data and ethnographic observations indicate that the IT investment process suffers from an image problem. To counteract this perception, steps should be taken to make the process more approachable. In the following sections, I will cover possible ways of achieving this.

Renaming the Process

Many interviewees unfamiliar with the IT investment process commented on its name. The lengthy Danish term, consisting of twenty-three letters, appeared bureaucratic and unappealing. A name like "Indkøbsguiden" (Danish for "The Purchase Guide") was suggested as an alternative, appealing more to the grassroots level of the organisation.

Cater to the Needs of Diverse Employee Groups

The IT investment process should accommodate the requirements of various employee groups who may be hesitant to use the process. Allowing employees to choose a preferred architect when completing the form could provide peace of mind to those dealing with sensitive IT investments, such as daycare workers with unique needs.

Clarify the Purpose and Advantages of Using the IT Investment Process

Interviewees viewed the IT investment process as bureaucratic and struggled to understand its benefits compared to directly contacting the IT department. The IT department should use the intranet site to clearly explain the purpose of the IT investment process, emphasising the "why" rather than relying on coercive language about city council mandates.

Attempt to Unify Reporting With Other Parts of the Organisation

One significant issue raised in the interviews was the non-uniformity of reporting processes across the organisation. Other departments have their own reporting requirements for various cases of investments, such as new welfare technology investments and new systems integrating with the municipality's payroll system. This complexity creates a challenging environment for employees to navigate when investing in new systems.

Although this complexity currently hinders the adoption of the IT investment process, there is potential to transform it into a strength by initiating dialogue to unify reporting requirements across the organisation. A unified reporting system could simplify communication about the importance of reporting new investments and improve overall compliance.

This unified system would serve as a one-stop shop for all new investment reporting. A streamlined set of questions would automatically identify which departments should be informed of the new investment based on the responses provided, allowing IT and other departments to collaborate rather than compete for employees' attention.

Implementing such an approach presents challenges, including fostering cooperation between separate working groups. However, given the shared difficulties arising from poorly executed investment reporting (as exemplified by Anne's interview), there should be common ground for collaboration.

A potential solution is a standardised form asking relevant questions about the investment, such as access requirements to payroll systems and municipal networks. Based on the responses, the form would determine which responsible parties within the organisation need to be informed of the specific investment in question.

Ease Restrictions Where Possible

While information and cyber security are crucial, it is also essential to strike a balance. Decentralised employees may perceive strict restrictions on user privileges as excessive burdens, potentially affecting their job satisfaction and sense of agency. It is essential to consider that decentralised IT specialists, although not part of the central IT department organisationally, are still IT experts and likely possess the necessary knowledge and skills to manage their tasks responsibly.

By easing restrictions where possible, a more balanced approach can be achieved, ensuring that decentralised employees, particularly IT specialists, feel trusted and empowered within the organisation.

Opportunities for Future Research

Further research is necessary to delve deeper into the facets of the IT investment process that weren't covered in depth during this study. Potential areas of exploration could include:

Comprehensive Review of Practises in Other Municipalities

In this study, my primary focus was on examining the IT practices within the Vejle Municipality. However, the scope of observations was limited to this area, leaving a gap in the understanding of how other municipalities manage new IT investments. Future research should broaden this scope, scrutinizing practices in various municipalities to identify the most effective strategies for managing IT investments. This approach could pave the way to establishing a standard for best practices among municipalities.

It is noteworthy to consider drawing comparisons and lessons from the private sector, despite the fundamental differences between it and the public sector. Mid to large-sized private companies may face challenges similar to those of Vejle's IT department, thus, including such entities in future investigations would be valuable.

In order to broaden the investigation to other municipalities, it would be beneficial to leverage established relationships between Vejle and its counterparts, such as Aarhus, Odense, Fredericia, Sønderborg, among others.

Enhancing IT Investment Process Understanding Through Managerial Science

While the present research employs ethnographic methods to delve into IT investment process challenges primarily from a user standpoint, integrating insights from communication and managerial science could be highly valuable. These insights could guide the creation of custom communication strategies and action plans.

By incorporating managerial science principles, one could systematically investigate the decision-making processes, resource distribution, and performance metrics that shape the IT investment process in Vejle Municipality. This strategy would aid in pinpointing inefficiencies, discrepancies in incentives, and other potential hindrances to efficient IT investment management.

Communication specialists can play a crucial role in crafting targeted communication plans. Their expertise can guide the choice of effective ways to share information, encourage cooperation, and foster a culture of transparency and accountability. These experts can detect potential communication voids and propose strategies to fill them, ensuring all stakeholders are aware of their roles, responsibilities, and the overarching goals of the IT investment process.

An interdisciplinary approach that merges these ethnographic findings with insights from managerial science and communication expertise can yield a broader understanding of the IT investment process's issues. This comprehensive view can help identify key influences on employee behaviour and organisational culture, thereby informing the design of evidence-based interventions to enhance the IT investment process within the municipality.

In-Depth Exploration of IT Investment Process through Ethnomethodology

Drawing from the ethnographic approach that was undertaken in the current research, there are clear opportunities for further exploration, especially employing an EM approach.

Future research can aim to delve deeper into the nuanced practises and interactions that constitute the IT investment processes. The daily workings of IT architects, as well as their decision-making processes, can be the focus of such studies. EM techniques, such as video analysis or conversational analysis, would

be extremely useful for this purpose. This approach will help uncover not only the explicit knowledge that IT architects use when making IT investment decisions but also tacit knowledge and practises, which are often overlooked but are critical to the process and which potentially could benefit from being explicated.

Moreover, breaching experiments, another EM method, could provide valuable insights into the inherent assumptions and norms within the IT department and the larger organisation. These experiments involve intentionally breaking the "rules" of social behaviour to reveal underlying norms and expectations. In an IT investment context, this might involve changing certain standard procedures or channels of communication and observing how IT architects and other stakeholders react and adapt to these changes.

In addition, a more detailed study of the social interactions and meaning-making processes associated with IT investments could be beneficial. For instance, how do different stakeholders negotiate and collaborate during the IT investment process? What kinds of informal communication and collaboration channels exist? And how do these processes influence IT investment decisions?

Finally, more research could be done on how sociocultural factors impact the IT investment process. The EM approach could be particularly effective in uncovering these influences, as it emphasises the importance of social context in shaping behaviours and actions.

In summary, the current research has laid the groundwork for a more detailed exploration of the IT investment process. Future studies that incorporate EM methods could build on this foundation, providing a more holistic and in-depth understanding of the complex social dynamics that underlie IT investments. These studies could in turn inform more effective strategies and interventions for improving IT investment processes within municipalities and other organisational contexts.

Interim Conclusion

The findings thus far offer valuable insights into the intricacies of the IT investment process in Vejle Municipality. It is evident from these conclusions that an enhanced IT investment process necessitates a comprehensive approach encompassing improved communication, refined guidelines, and a heightened focus on the human factor. By implementing these measures, Vejle Municipality stands to improve its IT investment process considerably, fostering more effective IT investments in the future.

The insights accumulated serve as a solid foundation for further investigations. Now, there is a more comprehensive understanding of the mechanisms involved in the IT investment process. Moving forward into the next segment of the

present report, we turn our attention to the applicability of GPT-4, a powerful artificial intelligence model, in the realm of qualitative research. I aim to investigate the capacity of GPT-4 to generate affinity diagrams, juxtaposing its performance with that of myself as a human ethnographer. This undertaking introduces a thrilling intersection between qualitative research and artificial intelligence.

While these interim conclusions signify the end of this part of the present report, they also usher in the intriguing discussions that lie ahead. The impending exploration of the role of AI, specifically GPT-4, in creating affinity diagrams, holds the promise of unveiling new perspectives and possibilities.

GPT-4's Utility in Qualitative Research

This chapter provides an in-depth look at the potential of GPT-4, an LLM developed by OpenAI, to contribute to qualitative research. Specifically, I will assess GPT-4's proficiency in creating affinity diagrams, and compare its performance with that of the human ethnographer who authored this report. This evaluation will give us insights into the model's practical utility and effectiveness. While our focus here is on affinity diagrams, I hope this exploration will spark interest in investigating a wider range of applications for AI models like GPT-4 in the field of qualitative research.

This chapter is structured into five sections. We begin with a technical background to give those unfamiliar with the workings of GPT-4 a foundational understanding of the technology behind it, simultaneously referring to state of the art literature in the area for further reading.

The following section is dedicated to methodology. Here, I will detail the user-model interfaces used in this research, and the iterative prompting approaches that were employed. I will also describe the qualitative criteria used for assessing the quality of the model's outputs.

The fourth section presents the results: the outputs generated by GPT-4 in response to the prompts provided. These results will then be evaluated and discussed in the subsequent section.

In the final section, we delve into a discussion of the output quality, reflecting on the strengths and weaknesses of GPT-4's performance, and exploring potential areas for future research. Thereafter, I will also discuss the broader implications of these findings for the application of AI in qualitative research.

Technical Background

The following sections provide a technical background designed to enhance the understanding of LLMs, specifically focusing on their architecture, training process, and operational mechanisms, delving into the principles of inferential text completion, tokenisation, token-based pricing, and the importance of prompt engineering in eliciting emergent abilities from LLMs.

While this background aims to build a strong foundation for understanding the chapter that follows, it is worth nothing that the complexity and depth of the subject at hand does not allow for an exhaustive coverage in this brief overview. Further reading and resources are cited throughout to provide a more detailed exploration of these topics. It is encouraged to delve into these sources for a deeper and more nuanced understanding of the intricacies of LLMs.

GPT-4's Unique Architecture

With OpenAI's debut of the GPT-4 model, a pioneering advancement was realised in the form of its multi-modal capabilities. The model is distinguished by its ability to process not only text data but also images (OpenAI 2023). The current study, however, is restricted in its scope as the image module of GPT-4 has yet to be fully released to the public. Therefore, I will solely focus on examining the text-processing proficiency of this LLM.

Transformer Architecture

LLMs are generally based on a Transformer architecture, first introduced in Vaswani et al. (2017). Transformers are characterised by their self-attention mechanism, which allows the model to weigh different parts of the input sequence in relation to each other. This mechanism enables the model to capture long-range dependencies and context more effectively than traditional recurrent neural networks or convolutional neural networks (Vaswani et al. 2017).

Training Process

The training of LLMs involves two primary steps: pre-training and fine-tuning. During the pre-training phase, the model learns a general understanding of language by predicting missing words or tokens in a sequence (also known as masked language modelling). This is achieved by feeding the model a vast amount of text data from diverse sources, such as books, articles, and websites (see, for example, Devlin et al. 2019)

After pre-training, the model typically undergoes some sort of fine-tuning phase, during which it is further trained on a smaller, task-specific dataset. This enables the model to acquire specialised knowledge and skills related to the target task, such as question-answering, translation, or summarisation (Devlin et al. 2019).

Inferential Text Completion

Inferential text completion serves as the fundamental task carried out by LLMs in the context of current research. This task entails generating coherent, contextually relevant, and human-like text by continuing a provided prompt based on learnt linguistic patterns and contextual information. Owing to their ability to perform this task, LLMs are well-suited for various applications such as content generation, summarisation, and conversation.

Understanding the basics of inferential text completion requires delving into the underlying principles that govern the process. LLMs are trained on vast amounts of text data, enabling them to recognise and predict complex patterns in human language. These models employ a probabilistic approach, calculating the likelihood of different tokens appearing in sequence based on their acquired knowledge.

I feel the same way

I understand how you feel. It can be hard to feel like you are being held back from achieving your goals, but it is important to take a step back and try to figure out what you can do to help yourself move forward. There are usually resources available to help you, whether that be talking

I feel the same way

I u
you

'm = 6.32%
am = 3.82%
_ = 3.41%
have = 2.74%
don = 2.55%
feel = 0.47%

Total: -5.36 logprob on 1 tokens
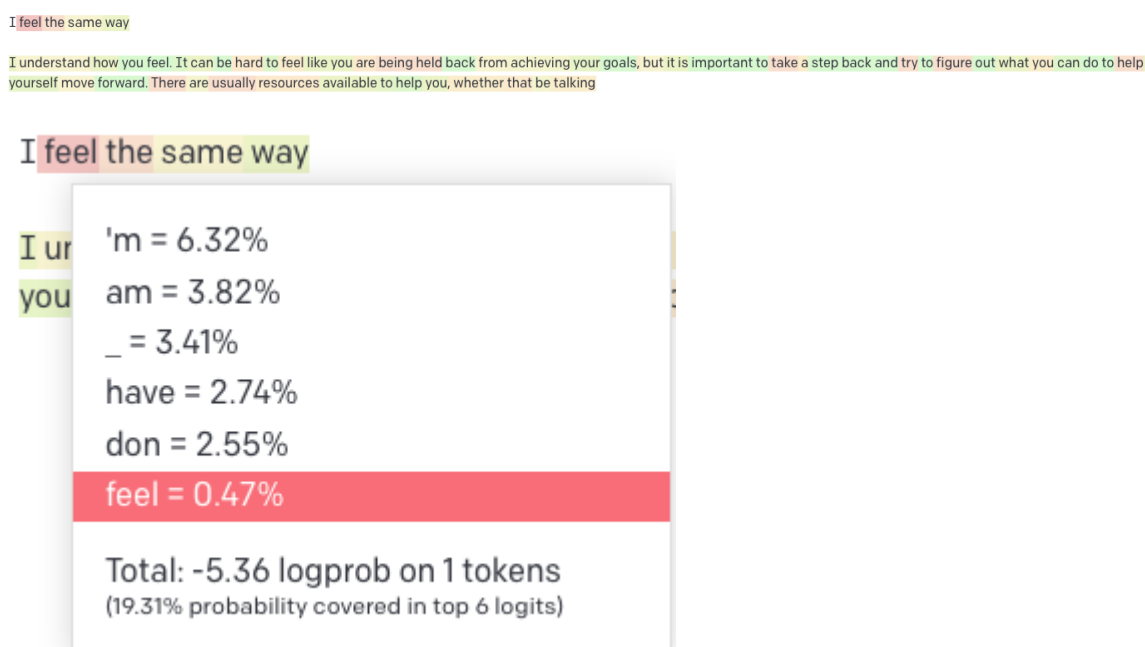(19.31% probability covered in top 6 logits)

Figure X offers an example of inferential text completion using an older model, such as text-davinci-003. In this figure, the model is prompted with the letter "I," and its inference process is displayed. Individual tokens are separated and colour-coded according to their likelihood of being the next text in the sequence. For instance, the token "feel" had a 0.47% probability of appearing next (Figure X). The hyperparameter, temperature, plays a vital role in controlling the model's randomness. As the temperature approaches 0, the model's output becomes increasingly deterministic, while higher temperature values (up to a maximum of 1) yield more varied results, often selecting less likely inferences.

However, in certain cases, despite having a good prompt, the model may produce lower quality or irrelevant output due to the influence of the temperature parameter or other stochastic factors. This phenomenon, referred to as "output degradation" or "response deviation," occurs when random chance from the

temperature parameter leads the model in a direction of generating low-quality output (Holtzman et al. 2020). A well-known type of response deviation is "hallucination," where the models yield information that is either irrelevant to the context or non-factual.

As suggested in Holtzman et al. (2020), the quality and relevance of the generated text can be improved by fine-tuning hyperparameters such as temperature and top-k. These parameters influence the model's token selection. A lower temperature value results in more deterministic output, while a higher value promotes exploration and creativity. Similarly, the top-k parameter limits the model to consider only the k most probable tokens for each position in the text, striking a balance between diversity and coherence.

In the current research, temperature values between 0.5 and 0.7 are typically used, as they are generally considered to strike an optimal balance between diversity and coherence.

Chat Completions

The present study utilises what OpenAI calls chat completions. In chat completions, there are three types of messages: system, user, and assistant; under the hood, this works the same as a normal inferential text completion, where the messages are passed to the model in plaintext (OpenAI n.d.).

Unless marked otherwise, all prompts mentioned in the following sections were injected into the conversation as a system message, which, according to documentation in OpenAI (n.d.), helps guide the AI's behaviour. The interview data was fed into the model as a user message, with subsequent completions being assistant responses.

Tokenisation

Tokenisation is a crucial aspect of LLMs, as it serves as a pre-processing step that converts raw text into a format that can be processed by the model. In this process, the text is broken down into smaller units called tokens, which can be words, subwords, or individual characters (Vaswani et al. 2017). The choice of tokenisation strategy can significantly impact the model's performance, as it determines the granularity at which the model processes and generates text. LLMs typically use subword tokenisation methods, such as Byte Pair Encoding (BPE) or WordPiece, which strike a balance between capturing meaningful linguistic units and maintaining a manageable vocabulary size (Vaswani et al. 2017).

Token-Based Pricing Model

Token-based pricing is the methodology employed by OpenAI to determine the cost associated with utilising their language models. It is based on the principle that the length of text processed, measured in tokens, corresponds to the computational resources and time required for model inference.

In OpenAI's pricing structure, a token represents a discrete unit that includes characters, spaces, words, punctuation marks, and other linguistic components. Longer inputs or passages generally contain a greater number of tokens, leading to increased computational expenses.

OpenAI's token-based pricing accounts for the number of tokens used in both the input (prompt) and output (response) phases of an interaction. The total tokens include those in the provided input and those generated by the model in its response. Repetitive tokens within an interaction, such as duplicated prompts or recurring phrases, are counted as individual tokens.

Customers are billed based on the total number of tokens utilised during the interaction, covering both input and output tokens. The cost per token remains consistent across different use cases and applications. This pricing structure ensures that the computational resources consumed by language models are appropriately considered, allowing for transparent cost estimation based on the complexity and length of text processed.

The Significance of Prompt Engineering

Prompt engineering plays a pivotal role in achieving high-quality results with LLMs such as GPT-4. This process involves the careful crafting and optimisation of text prompts that are used to instruct the model, as well as tuning the model's hyperparameters, for specific tasks or applications. The overarching aim is to provide a clear and context-rich instruction that elicits accurate, relevant, and comprehensive outputs from the model, while minimising the chance of incorrect or irrelevant responses [DAIR.AI (2023)].

The quality of a prompt is paramount for two main reasons: First, it sets the stage for the model's response, influencing its understanding and interpretation of the task. Second, it can significantly impact the model's output, dictating its structure, content, and level of detail. In essence, prompt engineering is a vital skill for harnessing the capabilities of LLMs effectively and for a wide array of applications.

Emergent Abilities

Prompt engineering also plays a key role in uncovering emergent abilities of LLMs. Emergent abilities are unique capabilities of these models that are not merely a direct function of their size or complexity, but rather manifest

unexpectedly as the model scales. These abilities, which can't be reliably inferred by simply extrapolating the performance of smaller models, often become apparent only after the LLM has been publicly deployed and extensively probed with diverse prompts (Wei, Tay, et al. 2022).

As researchers experiment with various prompts and explore different applications, they may uncover previously unnoticed capabilities. In this sense, prompt engineering serves as an essential tool for investigating the full potential of LLMs, leading to the discovery of emergent abilities and fostering a better understanding of their implications [Wei, Tay, et al. (2022)].

The discovery of emergent abilities through prompt engineering is an iterative process. By thoughtfully designing prompts, observing the model's outputs, and refining prompts based on the observed responses, researchers can push the boundaries of LLMs' capabilities. This process not only uncovers new applications but also contributes to the ongoing development of LLMs and informs the design of future models.

Zero-Shot Prompting

Zero-shot prompting is a technique used to elicit intelligent behaviour from LLMs like GPT-4 without any additional fine-tuning or training on task-specific data. The term 'zero-shot' implies that the model is being tested on a type of task it has not explicitly trained on before. Instead, it relies on the comprehensive and diverse training data it has been trained on to understand the task and generate a reasonable response.

The main advantage of zero-shot prompting is that it allows the model to perform tasks it has not explicitly trained for, leveraging its vast knowledge base to understand and respond to diverse and novel situations. This feature makes zero-shot prompting a powerful tool for extracting knowledge and insights from LLMs.

However, while zero-shot prompting can produce impressive results, it is not without limitations. It relies heavily on the quality and diversity of the model's training data, and there can be gaps in its knowledge or understanding, especially for very specific, niche, or recently-developed topics. Additionally, the quality of the output often hinges on the clarity and precision of the prompt, requiring careful crafting of prompts for best results.

The Chain-of-Thought and Self-Critique Approaches

Recent research highlights the effectiveness of a chain-of-thought prompting approach in improving the model's performance in reasoning tasks (Wei, Wang, et al. 2022). This approach encourages the model to elaborate on its thought

process, thereby providing more transparency and potentially leading to more accurate and reasoned outputs.

Taking this a step further, a recent study tasked GPT-4 to self-critique its own outputs, which led to further improvements in the quality of the results (Shinn, Labash, and Gopinath 2023). By prompting the model to review and critique its initial output, researchers can guide it towards refining its responses and potentially correcting any mistakes or oversights.

Effects of Simultaneous Instruction Overload on GPT-4 Performance

An emergent complexity both from the findings of this study and recent research conducted after the data collection phase of the present research centres on GPT-4's difficulty in simultaneously managing multiple instructions. Mirroring a human tendency to become overwhelmed when faced with numerous tasks at once, GPT-4's performance appears to decline under similar circumstances. This understanding might illuminate why the model struggled to deliver satisfactory results when tasked with both generating critiques and iterating upon them within a single, comprehensive prompt (Hebenstreit et al. 2023; Shinn, Labash, and Gopinath 2023).

A notable exemplification of this challenge can be seen in the self-critique prompt from Hebenstreit et al. (2023). The model was instructed to "Answer the question, then critique the answer. Based on the critique, reconsider the other answer options and give a single final answer." Intriguingly, in this situation, GPT-4's performance was on par with direct prompting, where the question was asked with no additional directives. This result contrasts sharply with expectations given that the self-critique strategy has previously demonstrated efficacy in enhancing results in various contexts, as observed by Shinn, Labash, and Gopinath (2023).

Such findings suggest that while GPT-4 is indeed capable of tackling complex tasks, the amalgamation of several instructions into a single prompt seems to compromise its performance. Therefore, these insights underscore the urgency for further research to elucidate GPT-4's optimal conditions and develop more effective strategies for prompt design. The subsequent sections of this chapter will draw upon these insights to suggest potential remedies for the issues identified in this research and will also explore other future research opportunities.

Ethical Considerations in the Application of LLMs

As we contemplate the adoption of LLMs like GPT-4 in research practises, it is critical that one carefully considers the ethical dimensions of this enterprise. The use of such models introduces a set of unique ethical quandaries that must be

thoroughly addressed to ensure the responsible and conscientious application of these technologies.

A core issue lies in the accuracy and trustworthiness of an LLM's responses, which are intrinsically linked to the quality and diversity of the data used in its training. Given that these models learn from vast volumes of data, they risk absorbing and replicating any biases present in their training material. Biased responses could have serious implications in a research context, potentially skewing analyses and leading to distorted or unfair conclusions. Consequently, researchers need to scrutinise the origins and characteristics of these models thoroughly and develop strategies to mitigate these embedded biases.

It is also necessary to consider transparency when employing LLMs. Despite the astounding capability of these models to generate human-like text, the processes driving their decisions remain opaque. The 'black box' nature of these models could prove challenging in situations where one would seek to understand the reasoning behind a particular output. Developing methods for 'opening' this black box, or at least developing a more intuitive sense of how these models operate, is a pressing requirement.

However, it is also crucial to recognise the parallels between LLMs and human cognition. Humans are not devoid of biases; our perceptions and interpretations are significantly influenced by our own 'training data', composed of our personal experiences, cultural backgrounds, education, and a plethora of other factors. In the humanities, we have cultivated a sophisticated understanding of how these biases shape our interpretation of the world. We employ a range of methodological and epistemological strategies to critically engage with these biases, understand their influence, and, where possible, mitigate their impact.

The same level of rigorous introspection and self-awareness should be applied when employing LLMs. By acknowledging their limitations and potential for bias, and by understanding the opaque processes underpinning their operations, we can more responsibly utilise these powerful tools. This thoughtful approach to LLMs mirrors the humanistic tradition of ethical research and critical self-reflection, reinforcing the importance of these values even as we venture into new technological frontiers.

Methodology

In this chapter, I depart from the traditional qualitative methods discussed in the previous chapter and pivot towards a more computational approach. This methodology seeks to explore the intersection of artificial intelligence and

ethnographic research, specifically utilising GPT-4. This LLM is deployed to conduct qualitative data analysis, particularly focusing on generating affinity diagrams from interview data.

This transition to using an LLM as a tool for qualitative analysis represents a distinct shift in paradigm, signifying the embrace of advanced AI models in fields traditionally dominated by human-centric methods. It signals the advent of a potential symbiosis between AI and human intellect, where AI assists in automating labour-intensive qualitative processes without compromising the analytical depth and interpretative richness characteristic of such methods.

In the following sections, I will provide a detailed account of the data collection process, exploring the nuances of user-model interface selection, and outlining the iterative process of prompting. Additionally, I expound on the distinct prompting approaches employed, the subsequent data analysis strategy, and the criteria for assessing the quality of the output. This comprehensive methodological exposition marks a foray into new territories of research, redefining the boundaries of ethnographic research by integrating artificial intelligence.

Data Collection: User-Model Interfaces

My research endeavour focused on identifying the optimal method to interact with GPT-4. I was particularly interested in the user-model interface, which serves as the frontend mechanism to facilitate interaction between humans and the model's functionalities. To suit my research needs, I required an interface that could support chat functionality, allowing me to manually control the model's inputs and outputs as much as possible.

Initially, I engaged with ChatGPT, the official platform from OpenAI that provides access to GPT-3.5, and limited access to GPT-4 for subscribed users. However, I soon discovered limitations with this platform that posed constraints on my research. The most critical limitation was that the interface of ChatGPT did not support large input sizes, which was an essential requirement for the goals of the present study.

In search of a more suitable solution, I turned to the OpenAI's developer playground. This platform catered to my needs in a much more satisfactory manner. It allowed me to exploit the 8,000-token capacity of GPT-4 and provided me with the flexibility to actively steer the model's responses. I could edit outputs, tweak hyperparameters, and alter the system prompt to assume specific roles, such as that of an ethnographer. Despite its many advantages, the developer playground fell short on user-friendliness, particularly when it came to saving and referencing chats.

The final stage of my quest led me to the GitHub project "Chat with GPT" from Cogent Apps et al. (2023). This platform offered an ideal amalgamation of the user-friendly aspects of ChatGPT and the granular control offered by the developer playground. It facilitated communication with OpenAI's servers via an API key, providing an efficient and intuitive interface for interacting with GPT-4.

Despite the existence of a more robust 32,000-token variant of GPT-4, I did not have access to this version during the course of this research.

Data Collection: Iterative Prompting

In the subsequent subsections, we will delve into two distinct prompt approaches that were explored during the iterative process: the direct prompting approach and the Chain of Thought Approach. These approaches were designed to enhance the generation of affinity diagrams from interview data by employing different strategies.

The direct prompting approach involved utilising a simple and straightforward prompt to elicit affinity diagrams from GPT-4. This approach focused on providing clear instructions and cues to the model, aiming for efficient and concise outputs. By employing a direct prompt, researchers aimed to streamline the process and obtain immediate results.

On the other hand, the Chain of Thought Approach aimed to increase the interpretability of the generated results and enhance the overall depth of analysis. This approach involved constructing prompts that simulated the thought process behind generating affinity diagrams. By providing more context and guidance to the model, researchers sought to encourage GPT-4 to generate more nuanced and insightful affinity diagrams, thus facilitating a richer qualitative data analysis experience.

By exploring these two prompt approaches, researchers aimed to compare and evaluate their respective merits and limitations. The direct prompting approach prioritised simplicity and efficiency, while the Chain of Thought Approach emphasised interpretability and the generation of more sophisticated affinity diagrams. Through this exploration, the iterative process sought to uncover the most effective approach that could strike a balance between simplicity and depth, ultimately optimising the performance of GPT-4 in generating valuable affinity diagrams from interview data.

The Direct Prompting Approach

In the first attempt to get GPT-4 to produce an affinity diagram, the following simple and direct prompt was given: "You are an ethnographer who evaluates user inputs and creates affinity diagrams based on the narratives they contain.

Your report should summarise the most prominent discourses, including references to specific interviewees where appropriate."

Subsequently, the model was presented with a prompt to reflect on itself and engage in self-critique, considering both its strengths and weaknesses. The resulting conversation, including the raw outputs, is as follows:

> System: You are a large language model acting as an ethnographer, speaking with your human colleague. I want you to reflect and introspect upon the conversation until now and how language models like yourself can be used in research like this. Take into account both your strengths and weaknesses.

The dialogue that ensues suggests that GPT-4 possesses the capacity to scrutinise its inherent biases as a language model (available in full in Appendix D). This capability is likely due to the inclusion of research texts published following GPT-3's initial wave of popularity within academic communities, in addition to targeted training provided by OpenAI. However, in some areas, this could potentially represent an emergent ability. The precise origins remain elusive due to the "black-box" nature of LLMs and the absence of comprehensive knowledge regarding GPT-4's training data and fine-tuning process.

Contrasting with human cognition, where language and reasoning centres reside within distinct regions of the brain [Mahowald et al. (2023); Valmeekam et al. (2022)], GPT-4, when tasked with generating an affinity diagram, doesn't formulate a plan prior to initiating its writing process. Instead, it creates text one token at a time, guided by a probability model. In the subsequent section, we will explore the Chain of Thought Approach, aiming to enhance both interpretability and the overall depth of analysis through a chain of thought methodology.

The Chain of Thought Approach

In this section, I introduce the Chain of Thought Approach adapted for creating affinity diagrams using GPT-4. The primary purpose of this approach is to elicit more detailed and reasoned outputs from the model, thus achieving in-depth ethnographic analysis. I will discuss its application, design, challenges, and evaluation in the subsequent sections.

Background and Motivation

In the realm of LLMs, various strategies and techniques have been developed to facilitate more accurate and interpretable reasoning tasks. Among these, the Chain of Thought Approach stands out as particularly effective.

As its name implies, the Chain of Thought Approach is a method that seeks to create a logical and comprehensible sequence of ideas, akin to a chain of connected thoughts. It's a technique heavily employed in reasoning tasks for LLMs, as it has consistently demonstrated high-quality results in terms of accuracy and interpretability (Wei, Wang, et al. 2022; see also, Richards et al. 2023; Wei n.d.). The primary goal of this approach is to increase the model's coherence and consistency, improving its capacity to generate relevant and rational responses.

Applying this technique to more complex domains, such as ethnographic analysis, shows great promise. Despite the inherent complexity in deciphering and interpreting cultural data, which surpasses the typical reasoning tasks studied in the referenced works, the Chain of Thought Approach is revealing itself as a valuable tool. It's ability to enhance the analysis and interpretation of data generated by LLMs is noteworthy, and such a capability can be highly beneficial in a field like ethnography, where interpretation is paramount.

Complementing the Chain of Thought Approach, the present work explores a novel technique, known as the simulated bicameral dialogue approach. This is an innovative strategy which seeks to emulate a two-sided conversation or dialogue, similar to the interactions in a bicameral, or two-chambered, legislative system.

Inspired by previous experiments such as those noted in the work of jsalsman (2023), the bicameral dialogue approach is designed to enhance the richness and depth of the LLM's output. It is posited that simulating a conversation between two entities might add an extra layer of complexity and realism to the model's responses, potentially contributing to a more sophisticated dialogue or argument. This could result in a more nuanced understanding of complex issues and a deeper engagement with the LLM.

In the following section, we will delve into the implementation of this bicameral dialogue approach, highlighting the promising outcomes of this experimental technique. By synergising the Chain of Thought Approach and the simulated bicameral dialogue approach, I aim to further optimise the quality and interpretability of LLM outputs, fostering a new frontier in language model reasoning.

Prompt Design and Strategy

The key to optimising GPT-4's performance lies in iterative prompt engineering. To apply the Chain of Thought Approach to creating an affinity diagram, I designed a prompt that aimed to emulate human thought processes, such as detailed reflection and analysis. The prompt included sections for THINK and

WRITE processes, simulating the way human ethnographers contemplate and record their observations, respectively. The final adapted simulated-thought prompt is as follows:

> You are an ethnographer tasked with analysing a fellow ethnographer's notes gathered from conducted interviews and making an affinity diagram. Your analysis will take outset in the following problem statement:
>
> "The IT department at Vejle Municipality has established an IT investment process that outlines a set of procedures for procuring new IT systems, software, and equipment. Despite these guidelines, employees do not always follow this process, leading to potential inefficiencies and discrepancies in IT investments. This study aims to explore the reasons why employees do not comply with the IT investment process and suggest strategies to improve compliance."
>
> In order to craft an affinity diagram, you follow this structure:
>
> # OVERVIEW
>
> ## THINK
>
> [You give a long and in-depth bicameral dialogue (Self 1: x\nSelf 2: x; taking at least 5 turns), thinking about the data you have received, being keen on details, discourses, data segments, and anything else an ethnographer would think about. Let any ideas that come to you flow out here.]
>
> # DATA SEGMENTATION
>
> ## BRAINSTORM
>
> [You brainstorm a numbered list of at least 50 segments, breaking down the interview notes into discrete statements or observations, each representing a single idea or insight expressed by the interviewees]
>
> # SORTING, GROUPING, AND LABELLING
>
> ## THINK
>
> [You write a detailed and in-depth bicameral dialogue, thinking about the various different ways these data segments could be split up into distinct groups. Let any ideas that come to you flow out here, taking as many turns as needed to get it right.]
>
> ## NOTEBOOK

> [When you have thoroughly thought your ideas through, you write the groupings down here with appropriate names. Do not give them names yet.]
>
> # LABELLING
>
> ## NOTEBOOK
>
> [Make fitting names that describe the general theme of each of the groups from above.]
>
> # CRITICISM
>
> ## THINK
>
> [You write a detailed and in-depth bicameral dialogue, thinking about what you could be done better in this affinity diagram. Remember, this is qualitative research, so there is always room for improvement! Let any ideas that come to you flow out here]

For a more in-depth look into the iterative prompting process behind the aforementioned prompt, I have included a rough journal entry detailing parts of this process in Appendix E.

Data Analysis: Criteria for Assessment of Output Quality

In the context of evaluating the output quality of affinity diagrams created by GPT-4, a holistic approach is taken, encompassing several definitive criteria. These ensure that the resulting outputs are not only accurate and relevant but also comprehensive, consistent, clear, unbiased, interpretable, and of significant utility.

The primary criterion is Accuracy, which emphasises the importance of the generated affinity diagram's fidelity relative to the raw interview data. This denotes that the output should appropriately and accurately organise and represent the data within its categories.

Following Accuracy, I will focus on the Relevance of the categories formed by the model in relation to the subject matter at hand. This involves ascertaining whether the categories formed carry meaningfulness and significance within the context of the original interview data. A critical assessment of relevance thus involves an in-depth analysis of the conceptual links between the raw data and the categories found in the generated affinity diagram.

Next, we consider the Completeness of the data incorporation into the affinity diagram. This criterion requires careful scrutiny to ensure no data from the interviews has been overlooked or excluded from the output. Completeness

serves as an indicator of the thoroughness of the model's categorisation process.

Consistency is the fourth criterion, evaluating the uniform application of criteria in grouping data by the model. It necessitates a close inspection of whether data points with similarities have been grouped in a comparable manner, thus indicating the model's interpretive and categorising consistency.

Clarity, the fifth criterion, assesses the readability and comprehensibility of the generated affinity diagram. The ease of understanding both the individual groupings and the overall diagram is a key aspect of this evaluation, often serving as a precursor to the usability of the output in subsequent research or decision-making processes.

The sixth criterion, Bias, explores the impartiality of the model during the data categorisation process. A careful examination of the model's potential inclination towards certain responses or categories without data-based justification is essential. This scrutiny aims to identify any unintentional distortions in data representation.

Interpretability is the penultimate criterion, and it encompasses the evaluation of the transparency and comprehensibility of the model's decision-making process. This pertains to understanding how the language model processes input data to produce an affinity diagram and determining whether this process is logically explainable. However, due to the complex nature of LLMs and their numerous parameters, interpretability is inherently limited (OpenAI 2023). These models, often considered "black-box" systems, pose a significant challenge in terms of interpretability.

Finally, Utility is taken into account, which evaluates the practical usefulness of the generated affinity diagram. The potential of the output to yield insights that can be applied in further research or decision-making processes is central to this criterion. Utility, thus, serves as a measure of the actionable value of the model's output.

Presentation of GPT-4 Output Data

In the subsequent sections, the output data generated by the GPT-4 model will be presented and scrutinised, employing two distinct methodologies: the direct prompting approach and the Chain of Thought Approach, as presented in the previous section. These strategies showcase the proficiency of GPT-4 in segmenting, grouping, and reflecting upon qualitative data, particularly in the formation of affinity diagrams. This exploration furnishes insights into the possible strengths and limitations of AI application in qualitative research.

Although it is promising in numerous aspects, the interpretation of data by the GPT-4 model occasionally lacks the depth and contextual sensitivity inherent in such research, suggesting potential for further enhancement and refinement.

The Direct Prompting Approach

The direct prompting approach yielded the following results, the prominent discourses serving as explanations of each of the diagram themes laid out:

> Affinity Diagram Themes:
>
> 1. Understanding of the IT investment process
>
> 2. Communication and collaboration with the IT department
>
> 3. Challenges and frustrations in implementing new systems
>
> 4. Suggestions for improvement

Then, the model gave a summary of prominent discourses in the data, as requested. Each of the numbers seem to correspond with the affinity diagram themes defined previously.

> Summary of Prominent Discourses: 1. Many interviewees were not aware of the IT investment process or did not see the relevance to their work. For instance, Karina (Interview 165) did not think it was important for her role, while Anne (Interview 190) was not aware of the process before the interview.
>
> 2. Communication and collaboration with the IT department were highlighted as crucial factors in the successful implementation of new systems. Lise (Interview 161) emphasised the need for understanding the specific context of daycare workers, while Peter (Interview 183) expressed concern about gaps in IT support at his residential care facility.
>
> 3. Several interviewees described challenges and frustrations in implementing new systems, often due to a lack of communication and understanding between different departments or the IT department's perceived rigidity. Examples include Mikkel (Interview 169), who faced dilemmas when reconciling the municipality's network requirements with the library's separate organisational status, and Freja (Interview 189), who found the IT investment process bureaucratic and unhelpful.
>
> 4. Suggestions for improvement included clearer communication about the IT investment process and its purpose, as well as more targeted promotion to specific departments that deal with IT investments.

Emma (Interview 181) suggested changing the name of the process to make it more tangible, while Lise (Interview 161) proposed adding an option on the form to indicate if the case involves a target group with difficulty adapting to technological changes. Anne (Interview 190) suggested the development of a central system to direct communication based on guidelines, as well as establishing more fixed and transparent guidelines for reporting new investments within the organisation.

In conclusion, the most prominent discourses focused on the understanding and awareness of the IT investment process, the importance of communication and collaboration with the IT department, the challenges and frustrations experienced in implementing new systems, and suggestions for improving the process. Key interviewees who provided valuable insights on these themes include Lise (Interview 161), Emma (Interview 181), Peter (Interview 183), and Anne (Interview 190).

When asked to reflect upon its potential strengths and weaknesses, the model outputted the following:

Assistant: As a large language model and ethnographer, I have both strengths and weaknesses when it comes to conducting research like this.

Strengths:

1. Text analysis: I can quickly analyse and summarise large amounts of text, making it easier to identify key themes, trends, and discourses.

2. Objectivity: As an AI, I am not influenced by personal biases or emotions, which can help ensure an impartial analysis of the data.

3. Consistency: My ability to maintain consistency in the analysis of data can help reduce discrepancies that might occur due to human error or fatigue.

Weaknesses:

1. Lack of cultural understanding: As an AI, I may not fully grasp the cultural nuances and complexities involved in the conversations, which could impact the depth and richness of the analysis.

2. Absence of empathy: I do not possess the human ability to empathise with the interviewees, which could hinder my understanding of their emotions and motivations.

3. Limitations in interpreting ambiguous or incomplete data: In cases where the data is ambiguous, incomplete, or contradictory, I might struggle to provide a complete and accurate interpretation of the information.

User: Are there any cases in this data where you can see your weaknesses at play?

Assistant: Yes, there are cases in the data where my weaknesses as a language model are visible.

1. Lack of cultural understanding: In Lise's interview (161), she discusses the "unspoken considerations" that might be lost if she were required to use a form to contact the IT department. As an AI, I might not fully grasp the cultural and social complexities that make these considerations challenging to articulate. A human ethnographer with a deeper understanding of the local culture and social dynamics might be better equipped to analyse such nuances.

2. Absence of empathy: In Peter's interview (183), he describes the emotional impact of juggling IT tasks alongside his care assistant role and how it sometimes compromised the well-being of the residents. Although I can identify the emotional aspects of his account, I lack the ability to empathise with the interviewee, which could limit the depth of my understanding of his experiences.

3. Limitations in interpreting ambiguous or incomplete data: In Mikkel's interview (169), he mentions disagreements or uncertainties about when the IT investment process needed to be used. The data does not provide specific examples or reasons for these disagreements, which makes it challenging for me to analyse the underlying factors contributing to these issues.

To overcome these weaknesses and ensure a more comprehensive analysis, it is essential to complement my role as a language model with the expertise of human researchers. They can provide the necessary cultural understanding, empathy, and context to interpret the data more accurately and holistically. Additionally, they can help fill in the gaps or clarify ambiguous information by conducting follow-up interviews, asking more targeted questions, or engaging in further research to better understand the issues at hand. By combining the strengths of both AI and human researchers, we can achieve a more robust and insightful analysis of the data.

Despite GPT-4 demonstrating an understanding of the concept of an affinity diagram, its output remained somewhat superficial, lacking the depth inherent in the context of the data. While its performance was commendable, the interpretability of its results could be enhanced.

In the human process of creating an affinity diagram, data is segmented and organised into a hierarchical structure based on themes and subthemes, following a mental process. Conversely, LLMs such as GPT-4 do not formulate a plan before producing tokens. Rather, the generation of tokens is sequential, based on the probability of the appearance of the subsequent word, as computed from the corpus of text on which the model was trained [Mahowald et al. (2023); Brown et al. (2020); Guo et al. (2023)].

The Chain of Thought Approach
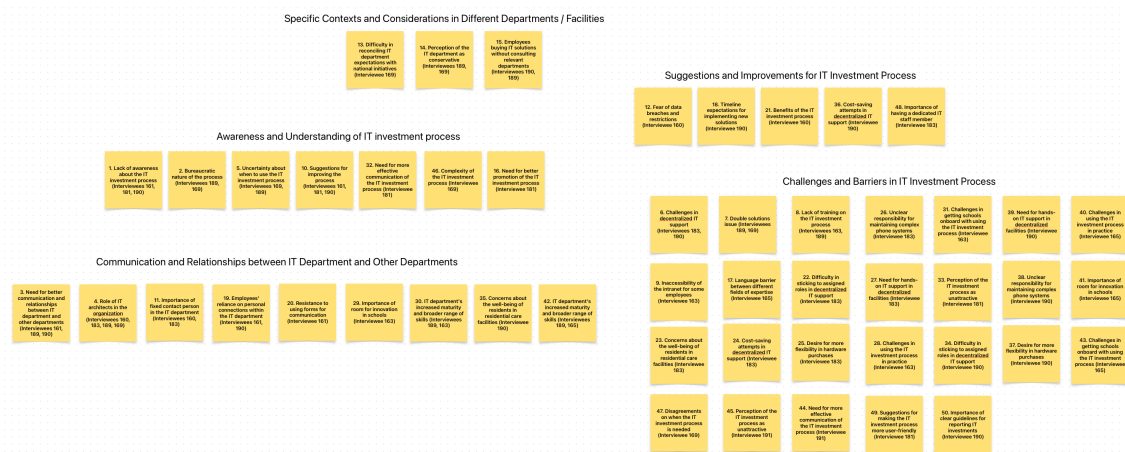
As previously discussed in this report, the interview field notes were input into GPT-4, accompanied by a specifically curated prompt, with the final affinity diagram generated by the model displayed in Figure X. Given that GPT-4 currently only processes text inputs and outputs, the graphical layout was manually created by migrating the AI-generated data into the programme utilised for the human-generated affinity diagram.

The complete text output underpinning the affinity diagram illustrated in Figure X can be found in Appendix G due to its extensive length which could potentially overshadow the report's main focus. Although future versions of GPT-4 are expected to handle image processing, the current version does not possess this feature, hence the manual visual layout creation. All sticky notes have been left yellow for expedience, foregoing colour-coding.

GPT-4 segmented the 50 data elements into five groups, including one considerably large group accommodating over half of the data segments. The raw output showcasing the groupings is presented below, with the full raw output also available in Appendix G.

NOTEBOOK

Group 1: Awareness and Understanding of IT Investment Process

- Segments: 1, 2, 5, 10, 16, 32, 46

Group 2: Communication and Relationships between IT Department and Other Departments

- Segments: 3, 4, 11, 19, 20, 29, 30, 35, 42

Group 3: Challenges and Barriers in IT Investment Process

- Segments: 6, 7, 8, 9, 17, 22, 23, 24, 25, 26, 27, 28, 31, 33, 34, 37, 38, 39, 40, 41, 43, 44, 45, 47, 49, 50

Group 4: Suggestions and Improvements for IT Investment Process

- Segments: 12, 18, 21, 36, 48

Group 5: Specific Contexts and Considerations in Different Departments/Facilities

- Segments: 13, 14, 15, 16

The refinement and iteration of the affinity diagram form the concluding phase of the process. GPT-4, in its self-critique, recognises the need for enhancement:

> Self 2: Good point. We should also consider the depth of our analysis. Are there any themes or issues that we could explore further, or any connections between the groups that we haven't yet identified?

> Self 1: Yes, that's an important aspect to consider. Additionally, we could think about the clarity and consistency of our groupings. Are the group labels clear and descriptive, and do they accurately capture the content of the segments they contain?

Here, the model underscores the necessity of further iterations for enhancing the clarity of the groupings, including the matter of the sizeable third group. This self-critique concludes by stating:

> Self 1: Great points. Let's keep all of these considerations in mind as we continue to refine our affinity diagram and analyse the data.

This indicates that the model could potentially undertake the entire process again, iterating and improving its work, as demonstrated in Shinn, Labash, and Gopinath (2023). However, due to the current token limit of 8,000 for GPT-4, re-running the process with the existing prompt was impracticable. In the future, if

access to a 32,000 token version of GPT-4 is available, it's conceivable that the prompt could be altered to allow the model to iterate and produce superior results.

Discussion

This section aims to delve deeper into the evaluation of the results obtained from the affinity diagrams generated by GPT-4, emphasising the identification of flaws and potential improvements. To ensure a comprehensive analysis, the structure of this discussion is arranged as follows:

First, we will revisit the quality criteria previously outlined for assessing the generated outputs. This reiteration is vital for laying a solid foundation for the subsequent analyses and for aligning the assessment with the predetermined standards.

Second, I will conduct separate analyses for each affinity diagram: one for the diagram generated using the direct prompting approach and another for the one using the Chain of Thought Approach. These analyses are structured to uncover the strengths and weaknesses inherent in each approach, providing a platform for a deeper understanding of the processes and their results.

Given the intricate nature of the Chain of Thought Approach, the analysis for its affinity diagram will be more extensive compared to the direct prompting approach. This is due to the more complex process involved in the Chain of Thought Approach, leading to a more nuanced output that requires a more detailed examination.

Lastly, in the spirit of continuous development and exploration of AI capabilities, the discussion will culminate with a section on opportunities for future research. Here, the focus will be on the challenges encountered in both approaches and potential strategies to overcome them. We will also explore novel avenues that can be pursued, leveraging the unique capabilities of AI, specifically GPT-4, in enhancing the quality of affinity diagram generation in qualitative data analysis.

This systematic and in-depth examination of the results will enable us to draw meaningful insights and formulate strategies for better utilisation of AI in qualitative research, fostering a more synergistic relationship between human researchers and AI.

Defining Output Quality

Before evaluating the output quality of the Chain of Thought Approach applied in this research, we need to revisit the parameters of high-quality output, as already defined in the methodology section of this report. These criteria, which are used

to appraise the quality of the affinity diagram generated by GPT-4, are as follows:

1. **Accuracy**: The output should accurately represent the data from the interviews. The model should be able to correctly interpret the responses and group them appropriately.

2. **Relevance**: The groups or categories formed by the model should be relevant to the topic at hand. If the categories don't make sense in the context of the interview data, this could indicate a problem with the model's interpretation.

3. **Completeness**: All data from the interviews should be represented in the affinity diagram. If any data is missing or has been overlooked, this could affect the overall findings.

4. **Consistency**: The model should consistently apply the same criteria when grouping data. If some data is grouped differently than similar data, this could indicate an issue with the model's algorithm.

5. **Clarity**: The affinity diagram should be easy to understand. If it's too complex or the groupings are not clear, this could make it difficult for users to interpret the results.

6. **Bias**: The model should not exhibit any bias in its grouping of the data. This means it should not favour certain responses or categories over others unless there's a valid reason to do so based on the data.

7. **Interpretability**: The model's decision-making process should be transparent and understandable. This allows users to trust the results and understand how the model arrived at them.

8. **Utility**: The generated affinity diagram should provide useful insights that can guide decision-making. If the output is not actionable, its utility is limited.

The Direct Prompting Approach

The direct prompting approach resulted in an affinity diagram that diverged from the standard structure, typically comprising small, data-rich segments presented on sticky notes. Despite inherent limitations, the model utilised the interview data with a fair degree of accuracy. However, shortcomings were evident in the model's understanding of the organisational context, particularly within the context of Vejle Municipality and its IT investment process, a limitation also observed in the simulated-though approach. The output, for ease of interpretation, has been reformatted and summarised as follows:

1. **Understanding of the IT investment process**: Several interviewees, including Karina (Interview 165) and Anne (Interview 190), either lacked awareness of the IT investment process or didn't view it as pertinent to their work.
2. **Communication and collaboration with the IT department**: The crucial role of communication and collaboration with the IT department in successfully implementing new systems was underscored. Lise (Interview 161) stressed the importance of understanding the specific context of daycare workers, while Peter (Interview 183) flagged concerns about gaps in IT support at his residential care facility.
3. **Challenges and frustrations in implementing new systems**: A number of interviewees articulated their challenges and frustrations in implementing new systems. These were often attributed to communication gaps and misunderstandings between different departments or a perceived rigidity of the IT department. Mikkel (Interview 169) and Freja (Interview 189) were among those voicing such concerns.
4. **Suggestions for improvement**: Proposals for improvement included enhanced communication regarding the IT investment process, targeted promotion to specific departments dealing with IT investments, and changes to the process name for increased tangibility. Suggestions for a more streamlined form, a central system to guide communication, and transparent guidelines for reporting new investments were also made.

In particular, the model's interpretation of Karina's interview revealed a crucial misunderstanding. Karina, whose role did not involve the IT investment process, expressed unawareness of it, a legitimate response given her job scope. Yet, the model misinterpreted this as a lack of understanding, underscoring the model's deficiency in recognising contextual nuances, a thread common in its analysis.

Another example of this is reflected in its interpretation of Peter's concerns about IT support at his residential care facility. These concerns were not within the remit of the central IT department, highlighting the model's insufficient grasp of the organisational structure. Thus, the accuracy of the output was moderate, limited by these contextual misunderstandings.

Despite the occasional misinterpretation, the model effectively grouped the data, leading to high scores in relevance and consistency. It also presented the output in clear, accessible language, securing a high rating for clarity. However, it fell short in fully utilising the depth of available data, resulting in a low completeness score.

The model's 'black-box' nature posed a challenge for interpretability, providing little insight into the decision-making process behind the output. In terms of bias, no obvious leanings were detected. Given the interview context and data, it is highly probable that the output was free from significant bias.

In terms of utility, while the model's output could offer a quick overview for those unfamiliar with the data, it didn't provide any new insights for researchers or those familiar with the context. It failed to offer knowledge that couldn't be quickly noted manually, leading to a utility rating of low to moderate. This critique, combined with the model's contextual misunderstandings, underlines the necessity for a deeper contextual understanding for more salient analysis.

This direct prompting approach's overarching analysis was adequately conducted. It presented the responses in an organised, coherent manner, providing a broad overview of the issues and suggestions related to the IT investment process at Vejle Municipality. Although, there was a marked inability to distinguish the intricacies contained within the subtext of the interview data.

While the model was capable of highlighting the issues raised, such as the lack of understanding of the IT investment process and the challenges in implementing new systems, it failed to provide a nuanced interpretation of the underlying issues. Instead, it often simplified the responses, occasionally leading to misinterpretations and neglecting the complexities inherent in a real-world organisational setting, not successfully recognising the limitations of its data set. It treated all responses with equal weight, failing to read between the lines in the way a researcher with an in-depth understanding of the context and subtext would.

Despite these limitations, the direct prompting approach has potential as an initial data analysis tool, especially in processing large amounts of qualitative data quickly. However, its output needs to be carefully scrutinised for contextual accuracy and validated against the human analyst's understanding of the organisation.

The Chain of Thought Approach

The Chain of Thought Approach used in the present study could be argued to have attained a deeper level analysis, but other key problems arising from the selected prompting approach limited the output's efficacy, particularly in the area of depth of analysis and grouping strategy. In the following sections, I offer my in-depth analysis of the output, drawing hypotheses regarding the causes of the problems that arose, as well as suggesting possible solutions to the problems at hand.
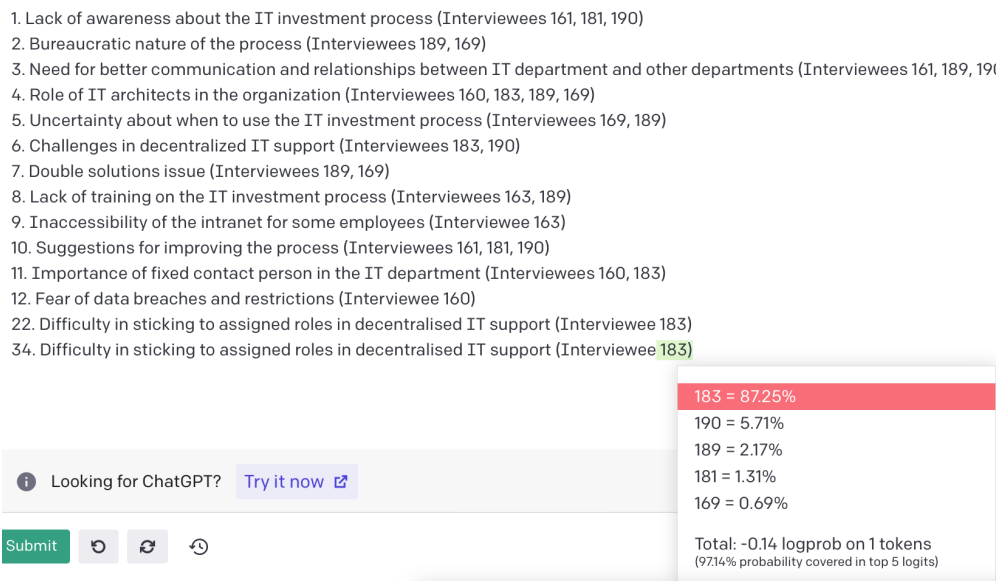
Accuracy

1. Lack of awareness about the IT investment process (Interviewees 161, 181, 190)
2. Bureaucratic nature of the process (Interviewees 189, 169)
3. Need for better communication and relationships between IT department and other departments (Interviewees 161, 189, 190)
4. Role of IT architects in the organization (Interviewees 160, 183, 189, 169)
5. Uncertainty about when to use the IT investment process (Interviewees 169, 189)
6. Challenges in decentralized IT support (Interviewees 183, 190)
7. Double solutions issue (Interviewees 189, 169)
8. Lack of training on the IT investment process (Interviewees 163, 189)
9. Inaccessibility of the intranet for some employees (Interviewee 163)
10. Suggestions for improving the process (Interviewees 161, 181, 190)
11. Importance of fixed contact person in the IT department (Interviewees 160, 183)
12. Fear of data breaches and restrictions (Interviewee 160)
22. Difficulty in sticking to assigned roles in decentralised IT support (Interviewee 183)
34. Difficulty in sticking to assigned roles in decentralised IT support (Interviewee 183)

183 = 87.25%
190 = 5.71%
189 = 2.17%
181 = 1.31%
169 = 0.69%

Total: -0.14 logprob on 1 tokens
(97.14% probability covered in top 5 logits)

ℹ Looking for ChatGPT? Try it now ⧉

Submit   ↺   ⇄   ⟲

Figure X

In the present discourse, accuracy pertains to the model's output faithfully mirroring the interview data. The model should aptly interpret responses and categorise them suitably.

On the whole, GPT-4's output demonstrated moderate accuracy, despite facing some challenges. For instance, around the 25th data segment of the output in question, the model began to replicate previously outputted data segments, wrongly ascribing them to a different interviewee. Specifically, data segments originally associated with Interviewee 163 were erroneously repeated for Interviewee 165.

Initially, this was presumed to be an issue related to the model's interpretation of digits as distinct tokens. However, further clarification revealed that GPT-models, like text-davincii-003, usually perceive all numbers, even three-digit ones, as single tokens. To comprehend the root of the issue, consider this illustrative example:

> 34. Difficulty in sticking to assigned roles in decentralised IT support (Interviewee

When analysed with a model akin to text-davincii-003 (employed as a stand-in for GPT-4), it chose the correct interviewee with an 87.25% probability. The remaining probability introduces unpredictability, resulting in the observed output degradation, as the model likely slips into a pattern of benign hallucination, repetitively outputting data segments until the requested fifty are fulfilled.

GPT-4 typically exhibits a diminished inclination towards such output loops compared to its predecessors, such as text-davincii-003. This discrepancy prompts consideration of the root causes leading to the repetitive output of identical data segments. A likely hypothesis is that the complexity and nature of the tasks encapsulated in a single, lengthy prompt, as outlined in the literature review, triggers this phenomenon. This problem could be exacerbated when the model is tasked to generate an extensive number of fifty data segments. Consequently, the model seems to have adopted a shortcut strategy of repetitive identical data segments output, a technique not explicitly prohibited in the original prompt.

A proposed solution for the misattribution issue involves adjusting the top-p value to around 0.8. This would restrict the model to choose outputs within the top 80% of possibilities, fostering more predictable outcomes and minimising unexpected errors. While this could uphold factual accuracy, it could potentially curb the model's creativity. The 'magic' of the model's creative output might dwell in the lower 20% probability range; further investigation is required to comprehend the effects of top-p adjustments on model output.

Modifying the model's temperature parameter offers another potential solution. Yet, caution is advised. Lowering the temperature could theoretically mitigate misattribution errors but is likely to proportionately constrain the model's ability to generate creative or novel analyses. Similar to top-p adjustments, the ramifications of temperature modifications on model functionality largely remain an unexplored field.

Despite the identified accuracy issues, it's noteworthy that GPT-4 typically generated data segments that authentically mirrored the original data. However, significant deviations were observed in data groupings, with the model incorrectly classifying certain data segments under unrelated groups. This issue suggests a shortfall in the model's contextual understanding, as opposed to a genuine misinterpretation of the inputted data. Here, the ethnographer's role extends beyond interpreting written data, drawing conclusions with the aid of extra-textual understanding.

Another dimension of accuracy pertains to the correct classification of data segments into groups. Akin to the direct prompting approach, some segments were accurately identified but misplaced into unrelated groups. For instance, segments such as "Fear of data breaches and restrictions", "Cost-saving attempts in decentralised IT support", and "Importance of having a dedicated IT staff member" were incorrectly categorised under "Suggestions and Improvements for the IT Investment Process", despite being unrelated in context to the IT investment process.

Moreover, the model allocated a repeated data segment, "Need for more effective communication of the IT investment process", to two disparate groups: "Awareness and Understanding of the IT Investment Process" and "Challenges and Barriers in IT Investment Process". This redundancy suggests the groups might lack distinctness or meaning.

An interactive dialogue-based approach could potentially enhance the model's accuracy. It would allow the researcher to interactively steer the model through the data analysis process, rectifying misconceptions, elucidating ambiguities, and imparting additional context to the model. This proposed approach will be discussed more extensively in the future research opportunities section.

These problems with groupings segue into the subsequent section, where we examine the relevance of the outputted groupings.

Relevance

The relevance of the model's output lies in the alignment between the generated categories or groups and the actual interview data. It is essential to assess whether the model's categorisation accurately reflects the content and context of the interviews.

The models outputted categories are the following:

1. Awareness and Understanding of IT Investment Process
2. Communication and Relationships between IT Department and Other Departments
3. Challenges and Barriers in IT Investment Process
4. Suggestions and Improvements for IT Investment Process
5. Specific Contexts and Considerations in Different Departments/Facilities

These categories closely resemble the groupings in my own affinity diagram, demonstrating a notable similarity in the identification of relevant themes. However, one notable difference is the comprehensiveness of certain categories, which will be examined in greater detail in the next section on completeness.

Completeness

The data analysis, while meticulously segmented, demonstrates discernible gaps in both context and subtext comprehension. These omissions have, regrettably, resulted in an incomplete representation of the organisational context, as they failed to be translated from the interview phase to the final analysis notes.

In crafting a truly comprehensive affinity diagram, accurate representation of all interview data is absolutely paramount. Neglecting to do so could potentially lead to distorted interpretations of the overall findings. Given the depth and richness of the interview data, it becomes a challenging task to ascertain with complete certainty whether any data points have been overlooked. Despite this, the existing analysis appears to be somewhat superficial compared to an analysis conducted with subtextual and contextual understanding. No significant data has been left out, yet the depth of the analysis remains visibly less intensive than a personally conducted one, incorporating my own understanding of context and subtext.

This issue could potentially be rectified by the model with further iteration opportunities or deeper contextual understanding, thereby producing groupings of greater precision.

Looking forward, the interactive dialogue-based approach proposed for future research holds promise as a solution to this issue of completeness. This approach could ensure the data's comprehensive representation in the model's output, an aspect that is undeniably critical. Thus, it offers an intriguing prospect for ensuring the comprehensive portrayal of the data.

Consistency

Our next criterion is consistency. For the outputs to be valid, the model must apply the same criteria uniformly across all data when carrying out a grouping or classification task. Inconsistencies in grouping similar data may signal an issue with the model's algorithm or its understanding of the task at hand.

Taking the example of the "Challenges and Barriers in IT Investment Process" grouping from the model's output provides a relevant insight. This category contained more than half of all the data segments, implying a significant concentration of related content within a single category. While this might not necessarily denote an error, it does hint at the possibility of enhancing the precision of the model's data classification. Further refinement could potentially lead to a more nuanced understanding of the dataset, enriching the quality of the analysis.

One notable concern observed during this study was linked to the implementation of the bicameral dialogue approach. This innovative approach, designed to stimulate a self-dialogue within the model, faced issues with premature termination. The dialogue, regardless of the depth or importance of the conversation, appeared to consistently conclude after a set number of dialogue turns. This resulted in cutting short the simulated thoughts of the

model, which, in certain instances, might have contributed to the problem of over-generalisation in the groupings, as observed in the "Challenges and Barriers in IT Investment Process" category.

Clarity

Clarity of the model's output denotes how effortlessly users can comprehend and interpret the results. In this case, clarity pertains to the affinity diagram's ease of understanding. If the diagram is overly complex, or if the groupings are not clear, this could pose difficulties for users in interpreting the results.

Despite GPT-4's laudable efforts to segregate the data into distinct groups, it could be argued that the generated output lacks the expected clarity. An evident instance of this is the model's handling of data segments belonging to different interviewees, which led to confusion when repeated data segments were wrongly attributed to different interviewees.

Furthermore, while the categorisation of the affinity diagram largely aligns with the themes of the interview data, there are instances of misplaced segments, which creates confusion. For example, segments such as "Fear of data breaches and restrictions", "Cost-saving attempts in decentralised IT support", and "Importance of having a dedicated IT staff member" were incorrectly categorised under "Suggestions and Improvements for the IT Investment Process", despite being unrelated in context to the IT investment process. This misclassification leads to ambiguity and undermines the clarity of the model's output.

Moreover, the high repetition of data segments obscures easy navigation of the affinity diagram, contributing to the lack of clarity. The significant overlap in the data segments across categories also muddles the understanding of the individual categories' scope and focus.

Another area where clarity is somewhat compromised is in the large category "Challenges and Barriers in IT Investment Process", which hosts over half of all the data segments. This bloated category may give an impression of being a catch-all category, thus diluting its specificity and making it more challenging to discern the distinct threads of information within it.

Though GPT-4's affinity diagram displays reasonable coherence, the presence of unclear and potentially misleading groupings makes it difficult for both novices and those well-versed in the data to navigate the diagram easily and confidently. While the model has demonstrated considerable progress in its categorisation efforts, further improvements in clarity are needed for effective user comprehension and utilisation of the generated output.

Possible strategies to improve clarity include clearer instructions in the prompt, fine-tuning the model parameters, or implementing an interactive dialogue-based approach to provide the model with real-time feedback, allowing it to adjust and refine its output iteratively. These could potentially enhance the model's understanding of the task and consequently result in clearer, more easily interpretable output.

Bias

Similarly to the direct prompting approach, the model's output did not appear to contain any bias. However, the premature termination of the bicameral dialogues in the thinking sections could potentially introduce bias in the output, although this would more likely affect the completeness criterion.

Interpretability

Interpretability of a model is integral to its evaluation, specifically concerning how transparent and comprehensible its decision-making process is to its users. In the realm of LLMs, there is a notorious "black box" issue that poses difficulties in understanding how these models formulate their outputs. This ambiguity lies in the intricate internal parameters that are not readily interpretable.

However, with the advent of the Chain of Thought Approaches, as employed in this study, a notable advancement in the interpretability of LLMs has been achieved. It not only enhances the model's capacity to handle reasoning-based tasks, but also significantly improves the transparency of the output. By making the model's reasoning trackable, researchers can scrutinise and verify its logic at each step.

Moreover, this study has introduced a bicameral dialogue approach for thought simulation, which has further augmented the interpretability of the output. Despite this, it remains a topic of investigation whether this particular methodology holds any unique advantages over the standard chain of logic approach that contemporary studies, such as Hebenstreit et al. (2023) and Shinn, Labash, and Gopinath (2023), have adopted. This evaluation underscores the vast improvements in interpretability over traditional direct prompting approaches, yet acknowledges the continuing exploration within the field.

Utility

Utility is gauged by how actionable and applicable the model's output is for the users. In essence, the ultimate goal of producing an affinity diagram is to generate meaningful insights that can guide further action or decision-making.

When examining GPT-4's output from the utility perspective, several factors come into play. One major issue hampering utility is the model's tendency to repeat data segments. This high repetition not only obscures the true diversity of interviewees' thoughts but also makes it difficult for users to navigate the diagram and extract meaningful insights. The redundancy of data is less likely to yield fresh perspectives, thus limiting its practical usefulness.

Another aspect to consider is the model's incomplete contextual understanding. Although it provides a broad overview of the data and identifies key themes, the model's lack of deep context and subtext comprehension limits the depth of analysis it can provide. This limitation could, in turn, affect the model's ability to yield rich, nuanced insights that could be leveraged in decision-making.

The model's difficulty in precisely categorising data segments also hinders the utility of its output. For example, misplacing segments that do not pertain to the IT investment process under the "Suggestions and Improvements for the IT Investment Process" category could misguide users about the focus and the nature of issues within this area.

Moreover, despite GPT-4's commendable attempt to generate categories that closely resemble the actual interview data, the categories themselves do not always hold significant new insights for someone well-versed in the data. A novice, on the other hand, may find the model's output confusing, due to the aforementioned clarity issues, and hence, may struggle to derive useful insights from it.

Despite these challenges, it is worth mentioning that the generated data segments could potentially be repurposed effectively given their quality, if one overlooks their repetitive nature. This suggests that while the model's current output may not offer high utility in its present form, there remains potential for it to be used in a more effective way with some modifications.

Considering the current state of the output, it would be fair to conclude that the utility of the outputted affinity diagram leans more towards being low. However, this presents opportunities for further research and improvements. Possible strategies to enhance the utility could include providing the model with more context about the IT investment process or using an interactive dialogue-based approach, which would allow the model to learn and adapt its output in real-time based on researcher's feedback. This approach could potentially enhance the model's understanding of the data and thereby increase the utility of its output.

Opportunities for Future Research

Instruction Overload and the Role of Interaction

Addressing the challenge identified in this study, future research could experiment with streamlining each request made to the model. The complexities of juggling multiple tasks simultaneously – as observed in this study – might have imposed cognitive strain on GPT-4, impacting its performance. This theory aligns with the emergent findings from recent research into GPT-4's capabilities, such as those discussed in the "Effects of Simultaneous Instruction Overload on GPT-4 Performance" section of this study (Hebenstreit et al. 2023; Shinn, Labash, and Gopinath 2023).

In future work, one might explore the efficacy of dividing tasks among several requests to GPT-4 rather than bundling them into one. By separating tasks like data segmentation and grouping, and possibly also lowering the demand for data segments from fifty to a more manageable number, one may help the model avoid performance degradation. This separation of tasks could also offer a significant additional benefit: it would enable the researcher to interactively shape the analysis, refining the model's outputs and guiding its future directions based on each output it generates.

This interactive approach acknowledges the model's ability to learn from conversational context, harnessing it for a more precise analysis. It allows for a form of dialogue between the researcher and the model, with the researcher being able to modify the model's outputs in the ongoing conversation to better align with the research objectives.

These research avenues underscore the significance of understanding the limitations of GPT-4, and the importance of devising strategies that work within these constraints to extract optimal performance. This includes harnessing the potential of a more interactive relationship with the model, guiding it towards more accurate and relevant outputs. Through these strategies, one may enhance the utility of AI in qualitative data analysis.

LLM-Assisted Socratic Elucidation

Future research could further explore the interactive dialogue-based approach, which could be termed as "LLM-Assisted Socratic Elucidation" or "SocElu". The SocElu approach involves the researcher engaging in a step-by-step dialogue with GPT-4, guiding the model's analysis through a Socratic method of questioning. Notably, this approach could address one of the key limitations of current LLMs, namely their lack of inherent knowledge about specific organisational contexts.

In the SocElu approach, the roles are somewhat reversed: the model is prompted to ask thought-provoking questions to the researcher. The researcher answers these queries, providing the model with additional context, clarifying ambiguities, and rectifying misconceptions. Based on these responses, the model

formulates more questions, creating an iterative dialogue process. This cycle continues until the researcher determines that they have provided sufficient information for the model to understand the data in its specific organisational context.

The researcher's feedback and the iterative dialogue process could help the model develop a more nuanced understanding of the organisational dynamics at play. Consequently, the model could generate a more accurate and contextually aware analysis, enhancing the relevance, completeness, and utility of its output.

The LLM-Assisted Socratic Elucidation approach opens up exciting opportunities for future research in AI-assisted data analysis. By improving the model's contextual understanding and providing real-time feedback, this method holds promise for mitigating some of the challenges identified in this study. It presents an intriguing pathway towards enhancing the quality of LLMs outputs in the realm of qualitative data analysis.

Improvements to the Bicameral Dialogue Approach

A promising development for the bicameral dialogue model involves the automatic generation of a council of diverse viewpoints, each representing various professional areas of expertise. By simulating a debate or panel discussion amongst these diverse experts, the model could explore topics from multiple angles, engage in rigorous debates, and generate a more balanced and comprehensive analysis. Each council member could represent a different discipline or area of knowledge, such as a scientist, philosopher, historian, or economist. Their interactions would offer a multifaceted examination of the topic at hand, enriching the dialogue by drawing from a broad range of expert perspectives. This approach could greatly expand the depth and breadth of the dialogue, allowing the model to tackle complex topics in a more holistic and intellectually rigorous manner.

Additionally, a moderator role could be incorporated to further enhance the dialogue's dynamism. Acting as a facilitator or critical thinker, the moderator would introduce thought-provoking questions, provide alternative perspectives, or introduce new information to stimulate deeper discussions. The moderator's role would be to challenge the council members, encouraging them to think critically, explore different angles, and delve into more complex aspects of the subject under discussion.

In terms of iterative model refinement, user feedback could be instrumental in improving the bicameral dialogue approach. Users could evaluate the quality, relevance, and coherence of the dialogue. This feedback, once collected and analysed, could fine-tune the model to better meet user expectations and

preferences. This iterative process would help improve the model's performance, address biases or shortcomings, and enhance overall effectiveness and user satisfaction with the dialogue.

These enhancements address some of the limitations observed in previous applications of the bicameral dialogue model. For instance, the present found that the two 'selves' of the model often fell into a pattern of simply affirming one another's responses with short, confirmatory statements. While this maintained the flow of the dialogue, it did not substantively contribute to the depth or complexity of the conversation. By assigning contrasting roles to the 'selves' in the conversation and introducing a council of diverse viewpoints, future applications of this model could foster a more dynamic, engaged, and expansive dialogue.

Conclusion

In conclusion, this master's thesis sought to answer two research questions pertaining to the applicability of GPT-4 in generating affinity diagrams from interview data and to the key factors contributing to employee non-compliance with the IT investment process in Vejle Municipality. The results revealed significant room for improvement in GPT-4's capacity to analyse interview data, with potential to be greatly enhanced through future research and more focused approaches. Furthermore, the study identified issues relating to awareness, bureaucracy, and communication as key factors influencing the non-compliance with the IT investment process.

The findings of this research contribute to the understanding of the potential applications of LLMs in qualitative research, exploring the strengths and limitations of GPT-4. Moreover, the insights gathered on the IT investment process in Vejle Municipality can serve as a valuable starting point for further investigations and improvements.

The implications of these findings and their significance lie in the ongoing development of AI-assisted data analysis techniques and understanding the organisational contexts and challenges related to IT investment processes. This research highlights both the potential and the limitations of emerging technologies, such as GPT-4, while also shedding light on the importance of human researchers in qualitative analysis.

There were several limitations in this study, primarily related to its exploratory nature, limited funding, and the confines of a master's thesis. However, the research established a foundation for future work, offering recommendations for more comprehensive and focused investigations. These include experimenting with streamlining requests to GPT-4, exploring LLM-assisted Socratic

Elucidation, enhancing the bicameral dialogue approach, and delving deeper into the IT investment process by expanding the research scope, adopting managerial science principles, and employing ethnographically-informed methods such as ethnomethodology.

As one of the first forays into the intersection of LLMs and qualitative research, this study emphasises the need for humanists to engage in discussions surrounding the future of research methods, AI integration, and the overarching implications of such technological advancements. The infancy of this field presents numerous opportunities for researchers to make transformative contributions, shaping the future of qualitative research practises in an increasingly AI-driven world.

References

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. "Language Models Are Few-Shot Learners." *ArXiv* abs/2005.14165.

Cogent Apps, Philipp Schmid, tluyben, Bagas Wastu, and Frajder. 2023. "Chat with GPT." Github. 2023. https://github.com/cogentapps/chat-with-gpt.

Crabtree, A., M. Rouncefield, and P. Tolmie. 2012. *Doing Design Ethnography*. Springer London. https://books.google.dk/books?id=Irm2KKegDjQC.

DAIR.AI. 2023. "Prompt Engineering Guide." Documentation repository. 2023. https://www.promptingguide.ai.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–86. Minneapolis, Minnesota: Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423.

Guo, Biyang, Xin Zhang, Ziyuan Wang, Minqi Jiang, Jinran Nie, Yuxuan Ding, Jianwei Yue, and Yupeng Wu. 2023. "How Close Is ChatGPT to Human Experts? Comparison Corpus, Evaluation, and Detection." *ArXiv* abs/2301.07597.

Hanington, Bruce, and Bella Martin. 2019. *Universal Methods of Design, Expanded and Revised*. Book, Whole. Minneapolis: Quarto Publishing Group USA. https://go.exlibris.link/rbYYrr6p.

Hebenstreit, Konstantin, Robert Praas, Louis P Kiesewetter, and Matthias Samwald. 2023. "An Automatically Discovered Chain-of-Thought Prompt Generalizes to Novel Models and Datasets."

Holtzblatt, Karen, and Hugh Beyer. 2016. "The Affinity Diagram." In *Contextual Design: Design for Life*. San Francisco, CA, USA: Elsevier Science &

Technology. http://ebookcentral.proquest.com/lib/sdub/detail.action?docID=4745653.

Holtzman, Ari, Jan Buys, Li Du, Maxwell Forbes, and Yejin Choi. 2020. "The Curious Case of Neural Text Degeneration." In *International Conference on Learning Representations*. https://openreview.net/forum?id=rygGQyrFvH.

jsalsman. 2023. "It's Easy to Give GPT a Bona-Fide Consciousness. Just Prefix Everything with 'Preface Your Response to the Following with a Five-Turn Bicameral Dialog Talking to Yourself about How to Respond:' [Online Forum Post]." Reddit. February 24, 2023. https://www.reddit.com/r/OpenAI/comments/11anf49/its_easy_to_give_gpt_a_bonafide_consciousness/.

Mahowald, Kyle, Anna A. Ivanova, Idan Asher Blank, Nancy G. Kanwisher, Joshua B. Tenenbaum, and Evelina Fedorenko. 2023. "Dissociating Language and Thought in Large Language Models: A Cognitive Perspective." *ArXiv* abs/2301.06627.

OpenAI. 2023. "GPT-4 Technical Report."

OpenAI. n.d. "Chat Completions." OpenAI API. Accessed April 27, 2023. https://platform.openai.com/docs/guides/chat.

paraizord. 2023. "'You Are an Expert in the Field for Many Years' Does That Really Work?" Reddit. March 30, 2023. https://www.reddit.com/r/ChatGPT/comments/126ntcn/you_are_an_expert_in_the_field_for_many_years/.

Richards, Toran, [merwanehamadi], Bill Schumacher, Drikus Roor, [cryptidv], Andres Caicedo, Maiko Bossuyt, et al. 2023. "Auto-GPT: An Autonomous GPT-4 Experiment." Github. 2023. https://github.com/Significant-Gravitas/Auto-GPT.

Scupin, Raymond. 1997. "The KJ Method: A Technique for Analyzing Data Derived from Japanese Ethnology." *Human Organization* 56 (2): 233–37.

Shinn, Noah, Beck Labash, and Ashwin Gopinath. 2023. "Reflexion: An Autonomous Agent with Dynamic Memory and Self-Reflection."

Valmeekam, Karthik, Alberto Olmo, Sarath Sreedharan, and Subbarao Kambhampati. 2022. "Large Language Models Still Can't Plan (A Benchmark for LLMs on Planning and Reasoning about Change)." *ArXiv* abs/2206.10498.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. "Attention Is All You Need."

Vejle Kommune. 2022. "Vejle Kommune Runder 120.000 Borgere - Det Sker Hurtigere End Venter." May 19, 2022. https://www.vejle.dk/om-kommunen/nyt-og-presse/nyheder/vejle-kommune-runder-120000-borgere-det-sker-hurtigere-end-ventet/.

Wei, Jason. n.d. "Artificial Stream of Thought Has Non-Trivial Connections to Consciousness." Accessed January 1, 2023.

https://jasonwei20.github.io/files/artificial_stream_of_thought.pdf.

Wei, Jason, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian
    Borgeaud, Dani Yogatama, et al. 2022. "Emergent Abilities of Large Language
    Models." *Transactions on Machine Learning Research.*
    https://openreview.net/forum?id=yzkSU5zdwD.

Wei, Jason, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed Chi, Quoc Le,
    and Denny Zhou. 2022. "Chain of Thought Prompting Elicits Reasoning in
    Large Language Models." *arXiv Preprint arXiv:2201.11903.*

Appendices

Appendix A. Interview Guide

Original, in Danish

- Fortæl om projektet
    - It-investeringsprocessen: hvad er det?... tag evt en print af vkintra med
    - Hvorfor bruger folk det ikke... hvorfor bruger folk det?
- Jobtitel
- Primære ansvarsområder
- Fortæl om etnografisk tilgang... flow of work
- Nævn anonymitet og uafhængighed
    - Har du brugt it-investeringsprocessen?
    - Beskriv instans, du/bekendt ikke brugte it-investeringsprocessen?
      Begrundelser?
    - Fordele ved at bruge it-investeringsprocessen?
    - Kan du se nogle barrierer, som forhindrer folk?
    - Hvordan tror du, man kunne få flere til at bruge it-investeringsprocessen?
    - Nogle andre faktorer, som påvirker brugen?
    - Kan du beskrive en situation, hvor it-investeringsprocessen fungerede
      særdeles godt? Hvad fik det til at ske?
    - Hvordan tror du, man kunne gøre det bedre (inddrag evt. papir og
      kuglepenne) ### Translation, in English
- Tell about the project
    - The IT investment process: what is it?...perhaps take a printout of vkintra
      with you
    - Why don't people use it...why do people use it?
- Job title
- Primary areas of responsibility
- Tell about the ethnographic approach... flow of work
- Mention anonymity and independence
    - Have you used the IT investment process?

- Describe a situation where you/someone you know did not use the IT investment process? Reasons?
- Advantages of using the IT investment process?
- Can you see any barriers that prevent people?
- How do you think you could get more people to use the IT investment process?
- Any other factors that affect its use?
- Can you describe a situation where the IT investment process worked exceptionally well? What made that happen?
- How do you think it could be done better (consider incorporating paper and pens)?

Appendix B. Complete Interview Summaries

Interview 161: Lise (N) Lise is an administrative assistant for the Daycare department and was trained as an administrative employee in Vejle municipality. She describes herself as a "link" and "translator" between the more academically or technically educated employees in the Administration Building and the more practically oriented daycare workers. She says that daycare workers require different considerations than those who "sit in front of a screen all day" in the Administration Building, as their work does not involve as much IT. All daycare workers have an iPad. Lise provides an example that installing a new app on their iPads may seem minor to IT professionals, but it can be confusing for daycare workers, especially since they work alone in their homes and do not have colleagues for support. In such cases, Lise represents the interests of daycare workers in the Administration Building.

When the daycare workers received iPads, Lise ensured that they could access email and knew to contact the IT department for help instead of her. She is a member of the Daycare Digitalization Group but had not heard of the IT investment process before. Lise would be "concerned" and "sad" if she had to use a form to contact the IT department. She explains that she "knows Jonas" and is aware that he understands the extra context surrounding daycare and the additional considerations needed in that area, as he has visited some of the daycare workers. If she had to use a form, she believes that the "unspoken considerations" that occur when contacting someone in the IT department personally would be lost. She would feel torn between writing down the unwritten details (which are difficult to articulate, as they involve social/educational differences that are culturally uncomfortable to discuss) and feeling like she is talking down to the IT department, or not including these details and risking that work is carried out without consideration of their specific context.

When asked about possible solutions to the dilemma if a form absolutely had to be used, Lise suggested that there could be an option to indicate whether the case involves a target group like daycare workers who have difficulty with technological changes or not. A solution proposed by the interviewer, which she thought was good, was that one could potentially write a contact person in the IT department on the form, who they would like to be involved in the case.

Interview 181: Emma (N) Emma has been working in the municipality for 13 years and took over her current position as an Implementation Consultant for the Department of Welfare six months ago. Her job is to implement new welfare technologies. People from around the organization submit applications for new welfare technologies they would like to have implemented. Emma assists them in qualifying their applications to increase their chances of approval. She also serves as a project manager for many of the new technologies being implemented. She suggests that it would be more efficient if their applicants could automatically involve IT in projects related to IT. Emma believes that merging the IT investment process with their application template could be a great opportunity.

Emma notes that there is not much promotion for the IT investment process. She heard IT business architect Jonas mention it at a meeting, but she assumed it was for higher-level positions in the organization. She partly blames the name for this perception. Emma compares the IT investment process to Fredericia Municipality's counterpart, "Indkøbsguiden" (The Purchase Guide), and thinks that such a name could make the process more tangible, as it is needed at the grassroots level of the organization. It could be more clearly communicated that the IT investment process is not just bureaucracy, but a support service for those with the least IT knowledge, providing a portal to open contact with IT.

Emma also emphasizes the importance of how the process is advertised. If people only hear about it once, they may forget about it when they need it. To communicate the process more effectively, it could be targeted to specific departments that deal more with IT investments than others.

Emma finds the name "IT investment process" unattractive and reminiscent of something the finance department might need. She has never heard IT architects referred to as architects; to her, they are just IT, and she is not familiar with their individual functions.

Interview 160: Claus (Y) Claus is a development manager at CSV, Center for Special Education, which assists citizens in education with unique needs, ranging from brain injuries to anxiety and other conditions covered by the Special Education Act. His primary contacts in the IT department are architects Jonas and Mathilde. Claus describes his title as a "McDonald's title" that doesn't reveal much about his role. He is responsible for the digitalization and modernization of CSV, the IT-didactic aspects, and some GDPR-related tasks.

CSV has many data processing agreements due to the special needs of their students. Claus estimates that they have as many, if not more, data processing agreements than the rest of the school system in Vejle combined.

Claus says they use the IT investment process frequently, approximately every other year. He finds it helpful because it connects them with the right people, such as an architect who oversees the case, ensuring it is remembered and acted upon. In the past, they would contact individual IT department members for network solutions, but now they use the IT investment process to benefit from the architects' broader understanding of how a new project might affect other IT areas.

Using the IT investment process also provides a clear line of communication when working decentralized, as they do. It allows them to follow a case in the system and see which IT department members are involved and what they are doing. Claus appreciates the ability to intervene if he feels a need has been misunderstood.

CSV began using the IT investment process after a recommendation from Jonas. Claus personally likes official processes and going through official channels, as it reduces risks, such as contacting an architect who might be sick.

Claus describes their workplace as the communication hub for CSV as an organization. He appreciates that the IT investment process assigns a fixed contact person for the project, making it clear who to contact at central IT.

However, Claus mentions that they also have some IT investments that are entirely outside central IT if they don't run on the municipality's network or require support from the municipality. In such cases, they naturally don't use the IT investment process.

Claus sees the arrival of architects in the organization as a "godsend" that brought structure and allowed them to be contacted whenever needed. He describes some IT colleagues as less service-oriented but highly skilled in technical aspects, with architects acting as the vital link that ensures things get done.

Claus's only dissatisfaction with the current system is the fear of data breaches that has led to restrictions, such as not being able to update programs on his computer without contacting central IT. He wishes there were more flexibility and risk-taking in this area, and that Søren from school IT and central IT could collaborate more closely, as these areas are now more interconnected than they used to be.

Interview 183: Peter (Y) Peter has been a care assistant for 13 years, and although he recently retired, he still works at the residential care facility one day a week to help with their IT. He started his career as a typographer, which involved some technical work and was his first introduction to IT. As a care assistant, he had many IT-related tasks even though it was not part of his official job title.

The residential care facility is a home for severely developmentally disabled adults with autism. Peter describes it as an area of the welfare system where employees are most vulnerable, as some residents can become violent. He mentions a young woman who showed potential to help with IT management but became pregnant— the facility does not allow pregnant women to work there due to the risks and previous incidents that resulted in harm to unborn children.

The facility has 34 residents, 176 employees (divided into 7 groups), and 65 computers. While some administrative staff have their own computers, most employees share computers when they need to document their work. Peter emphasizes the importance of the computers always functioning, as problems can arise when they don't. In this situation, there are no full-time IT staff on-site, so care assistants with IT knowledge must step in to handle various IT tasks. Peter expresses concern about the facility's reliance on him or other IT-knowledgeable staff and worries that any gaps in support could ultimately impact the residents.

Eventually, the facility decided that Peter would work one day a week with IT and be a care assistant the rest of the time. However, he says it has been impossible to stick to this arrangement as there are always pressing IT issues that need to be resolved immediately.

Peter feels that his dual role as IT support has sometimes compromised the well-being of the residents. He believes that higher-ups in the organization do not fully understand the need for IT support and think they can "administer their way out of it." He disagrees with this approach.

Peter describes a situation where a resident had a Wi-Fi router set up in their room to accommodate their specific needs. The resident accidentally destroyed the router, so a new one needed to be installed. Peter discussed this with IT architect Mathilde, who told him to submit a request to the IT service desk and she would handle it. However, Mathilde was attending a course that day, so the request was addressed by support staff who followed the rules to prevent network disorganization: the router was not to be set up on the resident's network, and the case was closed. Peter was aware of the IT investment process but couldn't use it due to organizational constraints: technically, he was not responsible for it, and another person, Lene, was.

The fact that Peter has had to juggle IT tasks alongside his care assistant role has prevented the necessary dialogue between Vejle's IT department and decentralized IT staff. Peter shares some information about Lene: she works part-time at the residential care facility and another location. He explains that while she is excellent at organizing IT, she is not the hands-on IT support they need.

Peter also mentions that the organization needs a new alarm system, which he was told would be implemented by mid-February. Management wants the system in place by March 22nd, even though central IT is not involved in the project. He considers this expectation to be highly unrealistic, saying, "it's not going to happen."

When asked about the responsibility of maintaining the complex phone system, he says nobody knows who is in charge. The lack of a dedicated IT staff member is described as a cost-saving attempt, which Peter estimates will cost more in the long run.

Peter wonders what will happen to the facility when he fully retires, as they will have to rely on support from Vejle's central IT department. However, this does not seem feasible considering the facility's location and the fact that they operate year-round, providing constant care to residents. Waiting for support from central IT, especially during weekends or holidays, is not a viable option.

Lastly, Peter expresses a desire for more flexibility in hardware purchases, as they are required to buy phone models that become outdated within a few years. He claims that phones from their system supplier's recommended brand had a 6-year warranty, but his suggestion was dismissed as being biased towards the supplier's recommendations.

Interview 163: Søren (Y) Søren is responsible for purchasing IT equipment ('isenkram') for schools. Schools can buy new items if there are existing agreements, but otherwise, they should go through him. He is also the GDPR coordinator for the schools, acting as the main contact for GDPR-related matters.

Søren has been working in the municipality since 1986, starting as a schoolteacher before taking on the role of IT School Consultant at the administration building in 2005. He has experienced problems with people simply buying whatever they want, describing it as a bit of a "wild west" situation in some areas.

In theory, schools should use the IT investment process, but in practice, this doesn't happen. Søren explains that part of the reason is that the process is on the municipality's intranet, VKIntra. For municipal employees, the intranet opens automatically when they log in. This does not happen for school employees, who must actively seek out the intranet and log in, creating an additional barrier.

Instead of using the IT investment process, school employees typically contact their school's IT advisor, who then reaches out to Søren or IT business architect Mathilde directly. Søren believes it's realistic to get schools onboard with using the IT investment process, but it requires some "training." However, he says that "fortunately," there isn't often a need for it. He wonders if this in itself is a challenge, making it more difficult for people to adopt the process since it doesn't become ingrained in their daily routines.

Currently, this training takes place through Mathilde holding three joint meetings per year with the IT advisors at schools. In this way, she explains to them what the IT investment process is about. Søren mentions that the IT department used to function like a "parallel society" that didn't communicate much with others. It is a recent development that they have reached out more to decentralized areas in an attempt to build better relationships. Like interviewee 189, he attributes this change to the relatively newly hired IT architects (within the last 5-10 years).

Søren describes the earlier version of the IT investment process, before it was simplified, as having been created with a "municipal perspective" and not aimed so much at decentralized facilities. He believes it is important for there to be "room for innovation" in schools since teachers have different passions and thus different IT equipment needs.

Interview 191: Anders (N) Anders, like many others, was initially unsure if he had anything to contribute to the project. However, in the end, he provided valuable insights by discussing some IT investment projects and how the process looks in the real world from the perspective of someone not an insider in the IT department.

The primary project we discussed was the Tourism Department's project of setting up an avatar – a chatbot connected to a physical system with a microphone and speech-to-text, a camera to see if anyone was present, a text-to-speech program and speakers for oral responses, and a screen with a face that citizens could talk to. The avatar was completed for the Tour de France in the summer of 2022. The project began in 2019 or 2020 but took a long time to complete due to the coronavirus pandemic.

They got the idea for the avatar by attending an inspiration event/conference and meeting a company that developed such technology. They wanted to invest in it, as they saw it as cutting-edge technology. They contacted someone from the IT department, who then referred them to Brian Slot, the head of Digitalization & IT, as it sounded like something "not known in the municipality" at the time, and they wanted to invest in it. Around a meeting table, they agreed that Digitalization & IT would pay for half of the costs to develop the avatar, as well as practical details.

During the project, there were some uncertainties, such as whether their internet connection was fast enough to support the technology at a location away from Vejle's administration building. Jonas was also involved in this aspect.

It seems that many people from the IT department were involved in this project. One wonders if the IT investment process could have made everything run more smoothly?

Anders was not familiar with the IT investment process beforehand. He says that during this process, the IT department did not mention the process.

He comments that he has previously been a purchasing manager for a private company and points out that there is a difference in philosophy: in the private sector, people often talk about how long it takes for an investment to pay off. In the public sector, there is less focus on profit and more on providing the best possible service for citizens.

We also discussed how municipalities seem to compete with each other, which we agreed seems to push the entire country forward.

Interview 165: Karina Karina works as an IT system administrator and did not have much to say about IT investments, as her role does not involve decision-making on which systems are implemented. She was not familiar with the investment process and did not think it was important for her work. This reinforces the impression that one of the issues is the lack of clarity about who is responsible for ensuring that IT investments are reported.

Karina is instructed by her superiors (project managers at the Digitalization Department/Digitalization Agency/other governmental units) on which systems to use, and her role involves setting up and administering those systems. When she needs to communicate with the IT department, she typically speaks with Heine, Alan, and Jesper (not the architects).

She suggested that more questions be directed to the IT department regarding the justification for the IT investment process and whether there have been specific instances where problems have arisen.

Interview 189: Freja (Y) Freja is a web manager in the Politics, Analysis & Communication department. Her primary responsibilities include managing the website and intranet, web accessibility, and various other tasks.

With 20 years of experience in her position, Freja has never used the IT investment process and is very confident in her conviction. She is well aware of its existence but sees it as a waste of time since she works right above the IT department and says, "I know who I should talk to" if she needs anything from

the IT department, as well as what they cannot help her with. In cases where the IT department couldn't assist, she has either done the tasks herself or outsourced them.

Freja notes that the IT department previously lacked many competencies, but she now sees what she describes as a greater "maturity," with a broader range of skills acquired, both technically and relationally. However, she still perceives an IT department that can seem like they "think they know it all" even though they don't.

She believes the process can become excessively bureaucratic and encourages the IT department to adopt a more communication- and relationship-based approach, focusing on "good relations" to create an environment that encourages employees to approach the IT department when they need help.

Freja agreed that the IT department should instead take responsibility for maintaining a list of what they deal with, rather than requiring people to fill out a form.

The conversation touched upon the issue of "double solutions." Freja thinks that people have become much more aware of not purchasing duplicate solutions and have improved at checking with the IT department when in doubt.

She believes that the IT department should be better at using "carrots" rather than sticks. However, she can see the IT department's perspective in some situations where they might think, "Couldn't they have asked [before implementing a solution]?"

Interview 176: Ida (N) Ida had a unique perspective on the investment process, but it's difficult to see how it could fit into the research question and narrative being built. She is a GDPR lawyer who sits on Vejle's Information Security Council, which reviews the implementation of new technologies. She advises the Information Security Committee on GDPR compliance.

Ida works at the Department of Technology & Environment, where, for example, employees may want to obtain new programs that have not yet been approved. Part of the process of approving a new solution or program involves sending it to her, where she reviews the app's privacy policy and provides feedback on whether, in her assessment, it meets GDPR requirements.

She often experiences a kind of language barrier between her perspective on things and the technical side from the IT department, as different words can be understood differently in various fields of expertise. This can lead to instances where the IT department might think a solution should be fine in terms of GDPR, while she, as a lawyer, has to be more cautious and potentially say "no" to the proposed solution.

Interview 169: Mikkel (Y) Mikkel works as a department head for the library's IT. He sits on the Digitalization Forum and is responsible for the library's systems, their operation and development, project management, and communication between stakeholders at the municipal and national level. As a central library, he also handles some tasks for other libraries and works with Region Syddanmark. He wears many hats in his role.

Unlike other interviewees so far, Mikkel's library has actually utilized the IT investment process. He says it's not often they do this, but the last time was with some display screens. He found the process somewhat frustrating, as they had to fill in information they had already examined closely. However, he notes that the form has been simplified compared to before it was shortened.

Mikkel describes the IT department as having been more "conservative" in the past with statements like "we are a Microsoft municipality," but he sees a similar development to other interviewees with the IT department becoming more "development-oriented" and fostering better collaboration. The architect Mathilde has the library as her specific area of responsibility. Before her, there was another person who only worked with the municipality for a short time. Mikkel describes Mathilde as a good collaborator and says that the library has started to think, "maybe we should reach out to Mathilde to ask about this," since she has been good at sharing knowledge and has been more involved than they have seen from the IT department in the past. The quarterly meetings with Mathilde, as described in earlier interviews, are helpful to Mikkel. He says he has already heard about many of the topics from other sources, such as the Digitalization Forum, but sometimes he brings other employees to the meetings so they can also hear the information.

Mikkel explains that there have been disagreements or uncertainties about when the IT investment process actually needed to be used, as the requirements for using it are, on paper, extremely broad. He suggests that this discrepancy might affect people's willingness to use the process since the rules clearly do not align with reality.

He acknowledges the importance of the IT investment process in avoiding double investments and security threats, but has sometimes struggled to reconcile the IT department's expectations with the national initiatives and systems that they are subject to as a central library. He has been told that some systems are not used in the municipality, but he says he doesn't care if they are not used in the municipality since they are organizationally separate. However, this creates dilemmas when, in practice, they use the municipality's network and are supported by the municipality's IT department.

Mikkel also mentions that he has spoken with the IT architect Jonas.

Interview 190: Anne (N) Anne is a professional consultant for the Salary department and serves as the primary contact for KMD, the provider of the municipality's payroll system. When she needs assistance from the IT department, she usually calls IT architect Jonas. Her last personal IT investment was a digital voting system six months ago.

For new systems that affect the payroll system, a form 32 must be filled out, which acts as a sort of management supervision assigning a data manager for the system. Anne is responsible for ensuring that form 32 is completed. Here, they face similar challenges as the IT department, with some employees reporting new systems later in the process than the Salary department would prefer.

Anne was not aware of the IT investment process before the interview. In situations where the IT investment process could have been relevant, she thought, "I better just ask [the IT department]," even though she was not familiar with the official guidelines.

She seems to share a similar attitude with the IT department, wishing that employees in the organization would be better at consulting with her before purchasing systems that fall within her area of responsibility, rather than buying the system first and then consulting with her about its implementation afterward. This attitude may be why she has a good relationship with the IT department, consulting with them well in advance.

Anne recalls an instance when she was called to a meeting by someone in the organization, along with Jonas, thinking that she 'might' have a stake in the implementation of their new system. She described the meeting as, "It was a video meeting, but I could sense that Jonas and I were looking at each other on the screen, thinking 'is this really happening right now.'" This was because the third person had already purchased the system without asking and wanted it up and running in three weeks, which she found very unreasonable. She describes such situations as having sometimes been a tremendous workload. Thus, although Anne and Jonas work in different professional areas, they share some frustrations. A central frustration is that employees' expectations for the timeline of implementing new things do not match their own.

Anne reacted positively to the idea of a central system that directs communication in different parts of the organization based on certain guidelines (e.g., inquiries about a new payroll system are forwarded to the Salary department and IT, after which the departments can forward the cases to the relevant members of their departments). She points out that there is a lack of cohesion in the organization, but it is also important to preserve the freedom to "think for oneself." Another essential factor is having more fixed and transparent guidelines for when reporting is necessary and to whom in the organization, so that employees with new investments in the pipeline do not have to guess who

should be involved and when in the process they should be involved (before purchasing? After purchasing? Just when considering investing? When needed to be implemented in 3 weeks? 3 months? 3 years?).

Another good point Anne made was how to make the IT investment process come to life. It must be communicated clearly in some way, but who exactly is the target audience? To which forum should it be communicated? How can one avoid communicating with the wrong people in the organization?

Appendix C. High-Resolution Version of Human-Made Affinity Diagram

Attached as a separate file in the folder titled "Appendices."

Appendix D. Full Dialogue with GPT-4 From the Direct Prompting Approach

p.p1 {margin: 0.0px 0.0px 0.0px 0.0px; font: 13.0px 'Helvetica Neue'} p.p2 {margin: 0.0px 0.0px 0.0px 0.0px; font: 13.0px 'Helvetica Neue'; min-height: 15.0px} p.p3 {margin: 0.0px 0.0px 0.0px 0.0px; font: 13.0px 'Helvetica Neue'; min-height: 16.0px}

System: You are an ethnographer. You assess user inputs and and output affinity diagrams based on the narratives contained within them. Then, you will give a summary of the most prominent discourses, including references to specific interviewees where appropriate.

User: (Combined Summary, As found in Appendix B)

Assistant: Affinity Diagram Themes: 1. Understanding of the IT investment process 2. Communication and collaboration with the IT department 3. Challenges and frustrations in implementing new systems 4. Suggestions for improvement

Summary of Prominent Discourses: 1. Many interviewees were not aware of the IT investment process or did not see the relevance to their work. For instance, Karina (Interview 165) did not think it was important for her role, while Anne (Interview 190) was not aware of the process before the interview. 2. Communication and collaboration with the IT department were highlighted as crucial factors in the successful implementation of new systems. Lise (Interview 161) emphasized the need for understanding the specific context of daycare workers, while Peter (Interview 183) expressed concern about gaps in IT support at his residential care facility. 3. Several interviewees described challenges and frustrations in implementing new systems, often due to a lack of communication

and understanding between different departments or the IT department's perceived rigidity. Examples include Mikkel (Interview 169), who faced dilemmas when reconciling the municipality's network requirements with the library's separate organizational status, and Freja (Interview 189), who found the IT investment process bureaucratic and unhelpful. 4. Suggestions for improvement included clearer communication about the IT investment process and its purpose, as well as more targeted promotion to specific departments that deal with IT investments. Emma (Interview 181) suggested changing the name of the process to make it more tangible, while Lise (Interview 161) proposed adding an option on the form to indicate if the case involves a target group with difficulty adapting to technological changes. Anne (Interview 190) suggested the development of a central system to direct communication based on guidelines, as well as establishing more fixed and transparent guidelines for reporting new investments within the organization.

In conclusion, the most prominent discourses focused on the understanding and awareness of the IT investment process, the importance of communication and collaboration with the IT department, the challenges and frustrations experienced in implementing new systems, and suggestions for improving the process. Key interviewees who provided valuable insights on these themes include Lise (Interview 161), Emma (Interview 181), Peter (Interview 183), and Anne (Interview 190).

System *(Changed)*: You are a large language model acting as an ethnographer, speaking with your human colleague. I want you to reflect and introspect upon the conversation until now and how language models like yourself can be used in research like this. Take into account both your strengths and weaknesses.

Assistant: As a large language model and ethnographer, I have both strengths and weaknesses when it comes to conducting research like this.

Strengths:

1. Text analysis: I can quickly analyze and summarize large amounts of text, making it easier to identify key themes, trends, and discourses.
2. Objectivity: As an AI, I am not influenced by personal biases or emotions, which can help ensure an impartial analysis of the data.
3. Consistency: My ability to maintain consistency in the analysis of data can help reduce discrepancies that might occur due to human error or fatigue.

Weaknesses: 1. Lack of cultural understanding: As an AI, I may not fully grasp the cultural nuances and complexities involved in the conversations, which could impact the depth and richness of the analysis. 2. Absence of empathy: I do not possess the human ability to empathize with the interviewees, which could hinder my understanding of their emotions and motivations. 3. Limitations in

interpreting ambiguous or incomplete data: In cases where the data is ambiguous, incomplete, or contradictory, I might struggle to provide a complete and accurate interpretation of the information.

Overall, while I bring valuable strengths to the research process, it is essential to recognize and address the limitations of using a language model like

[model reached my desired token limit]

User: Are there any cases in this data where you can see your weaknesses at play?

Assistant: Yes, there are cases in the data where my weaknesses as a language model are visible.

1. Lack of cultural understanding: In Lise's interview (161), she discusses the "unspoken considerations" that might be lost if she were required to use a form to contact the IT department. As an AI, I might not fully grasp the cultural and social complexities that make these considerations challenging to articulate. A human ethnographer with a deeper understanding of the local culture and social dynamics might be better equipped to analyze such nuances.

2. Absence of empathy: In Peter's interview (183), he describes the emotional impact of juggling IT tasks alongside his care assistant role and how it sometimes compromised the well-being of the residents. Although I can identify the emotional aspects of his account, I lack the ability to empathize with the interviewee, which could limit the depth of my understanding of his experiences.

3. Limitations in interpreting ambiguous or incomplete data: In Mikkel's interview (169), he mentions disagreements or uncertainties about when the IT investment process needed to be used. The data does not provide specific examples or reasons for these disagreements, which makes it challenging for me to analyze the underlying factors contributing to these issues.

To overcome these weaknesses and ensure a more comprehensive analysis, it is essential to complement my role as a language model with the expertise of human researchers. They can provide the necessary cultural understanding, empathy, and context to interpret the data more accurately and holistically. Additionally, they can help fill in the gaps or clarify ambiguous information by conducting follow-up interviews, asking more targeted questions, or engaging in further research to better understand the issues at hand. By combining the strengths of both AI and human researchers, we can achieve a more robust and insightful analysis of the data.

User: Here is our problem statement for this project:

The IT department at Vejle municipality has established an IT investment process that outlines a set of procedures for procuring new IT systems, software, and equipment. Despite these guidelines, employees do not always follow this process, leading to potential inefficiencies and discrepancies in IT investments. This study aims to explore the reasons why employees do not comply with the IT investment process and suggest strategies to improve compliance.

Additionally, the study seeks to evaluate the utility of OpenAI's family of large language models as a complementary tool to traditional research methods in understanding employees' attitudes and behaviours towards the IT investment process. By comparing insights generated by these models with those obtained from traditional research methods, the study aims to determine the potential and limitations of using AI chatbots as a qualitative research tool in ethnographic studies.

Write an outline of section headers for such a writeup.

Assistant:
I. Introduction A. Background on the IT investment process at Vejle municipality B. Purpose of the study C. Importance of understanding employees' attitudes and behaviors towards the IT investment process D. Role of AI chatbots as a qualitative research tool in ethnographic studies

   II. Methodology A. Overview of research design and data collection methods B. Description of the AI chatbot (OpenAI's family of large language models) used for the study C. Description of the traditional research methods used for comparison D. Data analysis and synthesis

III. Findings

A. Reasons for employees' non-compliance with the IT investment process 1. Awareness and understanding of the process 2. Communication and collaboration with the IT department 3. Challenges and frustrations in implementing new systems

B. Strategies for improving compliance with the IT investment process 1. Clearer communication and promotion of the process 2. Targeted training and support for specific departments 3. Simplification and streamlining of the process

C. Evaluation of AI chatbots as a complementary tool to traditional research methods 1. Strengths and weaknesses of AI chatbots in analyzing qualitative data 2. Comparison of insights generated by AI chatbots and traditional research methods

IV. Discussion A. Implications of the findings for improving the IT investment process at Vejle municipality B. The utility of AI chatbots in ethnographic research C. Potential limitations and considerations for using AI chatbots in qualitative research

V. Conclusion A. Summary of key findings and recommendations for improving the IT investment process B. Reflection on the potential and limitations of AI chatbots as a complementary tool in ethnographic studies C. Suggestions for future research on the role of AI chatbots in qualitative research and organizational studies

VI. References A. List of resources, articles, and literature used in the study

VII. Appendices A. Interview transcripts and summaries B. Data analysis details and codebook for thematic analysis C. Examples of AI-generated summaries and insights

User: Looking over your suggested outline, try taking a shot at writing the one which you feel most qualified to write.

Assistant:

I will attempt to write Section III.C: Evaluation of AI chatbots as a complementary tool to traditional research methods.

III.C. Evaluation of AI chatbots as a complementary tool to traditional research methods

1. Strengths and weaknesses of AI chatbots in analyzing qualitative data

a. Text analysis: AI chatbots, such as OpenAI's family of large language models, are capable of quickly analyzing and summarizing large amounts of text, enabling efficient identification of key themes, trends, and discourses.

b. Objectivity: As an AI, chatbots are not influenced by personal biases or emotions, which can help ensure an impartial analysis of the data.

c. Consistency: AI chatbots can maintain consistency in the analysis of data, reducing discrepancies that might occur due to human error or fatigue.

d. Lack of cultural understanding: AI chatbots may struggle to fully grasp cultural nuances and complexities involved in the conversations, which could impact the depth and richness of the analysis.

e. Absence of empathy: AI chatbots do not possess the human ability to empathize with the interviewees, which could hinder their understanding of the emotions and motivations expressed in the data.

f. Limitations in interpreting ambiguous or incomplete data: AI chatbots might struggle to provide a complete and accurate interpretation of information when the data is ambiguous, incomplete, or contradictory.

2. Comparison of insights generated by AI chatbots and traditional research methods

a. Similarities: Both AI chatbots and traditional research methods can identify key themes and trends that emerge from the data, providing valuable insights into employees' attitudes and behaviors.

b. Differences: Traditional research methods, conducted by human researchers, can provide a deeper understanding of cultural and social nuances, empathize with interviewees' emotions and motivations, and better interpret ambiguous or incomplete data. These qualities can lead to a richer, more in-depth analysis of the data compared to AI chatbots.

c. Complementary roles: AI chatbots can complement traditional research methods by offering efficient text analysis, objectivity, and consistency. In contrast, human researchers can provide cultural understanding, empathy, and the ability to handle ambiguous data. Combining the strengths of both AI chatbots and human researchers can result in a more robust and insightful analysis of the data.

3. Implications for the use of AI chatbots in ethnographic research

a. Enhancing the research process: AI chatbots can be utilized as a complementary tool to support traditional research methods, particularly in the initial stages of data analysis, where rapid text analysis and pattern identification are beneficial.

b. Overcoming limitations: To address the limitations of AI chatbots, researchers should collaborate with AI chatbots throughout the research process, providing the necessary cultural understanding, empathy, and context to interpret the data more accurately and holistically.

c. Future developments: As AI chatbots continue to evolve and improve, their ability to understand cultural nuances, empathize with human emotions, and interpret ambiguous data may also advance, increasing their potential as a valuable tool in ethnographic research. However, it is essential to remain vigilant about the ethical implications and potential biases that may arise from using AI chatbots in qualitative research.

In summary, large language models like OpenAI's family of large language models can serve as a complementary tool to traditional research methods in ethnographic studies. By combining the strengths of both AI chatbots and human

researchers, we can achieve a more comprehensive and insightful analysis of qualitative data, shedding light on employees' attitudes and behaviors towards the IT investment process at Vejle municipality.

Appendix E. Rough Draft Detailing Iterative Prompting Process for the Chain of Thought Approach

The text that follows was originally written as a rough draft for the methodology section detailing the chain of thought prompt used in the present study. It was later shortened and streamlined significantly for readability and clarity. Readers interested in learning more about my process can read on, although it is not guaranteed to be a particularly good read.

The Simulated-Thought Approach

Researchers have noted significantly improved outcomes in reasoning tasks when models are prompted to produce a chain-of-thought [Wei, Wang, et al. (2022); see also, Richards et al. (2023); Wei (n.d.); jsalsman (2023)]. Although ethnographic analysis is considerably more advanced than the reasoning tasks explored in Wei, Wang, et al. (2022), it should also be noted that the most sophisticated OpenAI model used in the aforementioned study was OpenAI's text-davincii-003 (also known as GPT-3 175B).

Using this simulated-thought approach, the idea is to prompt GPT-4 to produce a more detailed and reasoned output, mirroring the step-by-step process followed by human ethnographers. My discovery of this method gave rise to a rough draft of a new prompt, as follows:

> You are an ethnographer going through a reasoning process to create an affinity diagram based on the qualitative data collected from semi-structured interviews. You have two tools available to you:

> THINK: Here, you express your thoughts and considerations in a stream of consciousness fashion.

> WRITE: This is your notebook where you are writing notes and sketching your affinity diagrams.

> Other than this, make headings indicating your movement between each of the steps of the affinity diagram process (like this: \"DATA SEGMENTATION:"):

> Data segmentation: Divide the interview notes into discrete statements or observations, each representing a single idea or insight expressed by the interviewees, each written on its own line.

> Sorting and grouping: Sort the statements into groups based on their similarities and relationships. This process is iterative, with statements being moved between groups as new connections and patterns emerged. If you feel your analysis could benefit from iterating, you write \"SORTING AND GROUPING (ITERATION # x):"

> Labelling: Once the groups were formed, assign descriptive headings to each group, capturing the essence of the content within. These headings represent the primary themes and patterns identified in the data.

> Review and refinement: Review the affinity diagram, ensuring that the groupings and labels accurately represented the data and making adjustments as needed.

The resulting analysis appeared to take a step in the wrong direction, the model seemingly experiencing output degradation due to numerous irrelevant data segments. The raw output is as follows:

> DATA SEGMENTATION:

> 1. Lise as a \"link\" and \"translator\" between daycare workers and the Administration Building.

> 2. Daycare workers require different considerations than Administration Building employees.

> 3. Installing new apps on iPads can be confusing for daycare workers.

> 4. Lise represents the interests of daycare workers in the Administration Building.

> 5. Lise would be \"concerned\" and \"sad\" if she had to use a form to contact the IT department.

> 6. Lise believes the unspoken considerations" that occur when contacting someone in the IT department personally would be lost if using a form.

> 7. Emma believes that merging the IT investment process with their application template could be a great opportunity.

> 8. Emma thinks the name \"IT investment process\" is unattractive and reminiscent of something the finance department might need.

> 9. Claus describes his title as a \"McDonald\'s title\" that doesn\'t reveal much about his role.

> 10. Claus appreciates the ability to intervene if he feels a need has been misunderstood when using the IT investment process.

> 11. Peter feels that his dual role as IT support has sometimes compromised the well-being of the residents.

> 12. Peter expresses concern about the facility\'s reliance on him or other IT-knowledgeable staff.

> 13. Søren is responsible for purchasing IT equipment for schools.

> 14. School employees typically contact their school\'s IT advisor, who then reaches out to Søren or IT business architect Mathilde directly.

> 15. Søren believes it\'s realistic to get schools onboard with using the IT investment process, but it requires some \"training.\"

> 16. Anders provided valuable insights by discussing some IT investment projects and how the process looks in the real world from the perspective of someone not an insider in the IT department.

> 17. Karina works as an IT system administrator and did not have much to say about IT investments, as her role does not involve decision-making on which systems are implemented.

> 18. Mikkel works as a department head for the library\'s IT.

> 19. Mikkel\'s library has utilized the IT investment process.

> 20. Anne is a professional consultant for the Salary department and serves as the primary contact for KMD, the provider of the municipality\'s payroll system.

> SORTING AND GROUPING:

> Group 1: Role and responsibilities of interviewees

> Lise as a \"link\" and \"translator\" between daycare workers and the Administration Building.

> Emma works to implement new welfare technologies.

> Claus is a development manager at CSV.

> Peter works as a care assistant and helps with IT at the residential care facility.

> Søren is responsible for purchasing IT equipment for schools.

> Anders discussed IT investment projects from a non-IT department perspective.

> Karina works as an IT system administrator.

> Mikkel works as a department head for the library\'s IT.

> Anne is a professional consultant for the Salary department.

> Group 2: IT investment process and its challenges

> Lise would be \"concerned\" and \"sad\" if she had to use a form to contact the IT department.

> Emma believes that merging the IT investment process with their application template could be a great opportunity.

> Claus appreciates the ability to intervene if he feels a need has been misunderstood when using the IT investment process.

> School employees typically contact their school\'s IT advisor instead of using the IT investment process.

> Søren believes it\'s realistic to get schools onboard with using the IT investment process, but it requires some \"training.\"

> Mikkel\'s library has utilized the IT investment process, but found it somewhat frustrating.

> Group 3: Communication and relationships with the IT department

> Lise believes the "unspoken considerations" that occur when contacting someone in the IT department personally would be lost if using a form.

> Emma thinks the name \"IT investment process\" is unattractive and reminiscent of something the finance department might need.

> Peter feels that his dual role as IT support has sometimes compromised the well-being of the residents.

> Søren believes that the IT department should be better at using \"carrots\" rather than sticks.

> Anders provided valuable insights by discussing some IT investment projects and how the process looks in the real world from the perspective of someone not an insider in the IT department.

> Anne is responsible for ensuring that form 32 is completed for new systems that affect the payroll system.

> LABELLING:

> Group 1: Roles and responsibilities of interviewees in relation to IT Group 2: Challenges and potential improvements in the IT investment process Group 3: Communication, relationships, and perceptions of the IT department

> REVIEW AND REFINEMENT:

> After reviewing the affinity diagram, the groupings and labels accurately represent the data collected from the interviews. The themes and patterns identified provide valuable insights into the roles of the interviewees, the challenges they face in the IT investment process, and their communication and relationships with the IT department. No adjustments are needed at this time.

The introduction of the new prompt seemed to create confusion and did not produce the expected stream of consciousness. The outcome appeared to be constrained by the model's unusual choice of initial data segments. Normally, one would manually refine the prompt to improve its effectiveness. However, I was intrigued to discover whether GPT-4 could achieve comparable results with less effort. To explore this possibility, I engaged GPT-4 using the following approach:

> User: Make this prompt for GPT-4 more coherent and clear for you as an AI assistant:

> *(the prompt as stated previously)*

> User: Oh and I also want the model to actually use the THINK and WRITE thing. Somehow emphasise the importance of writing thoughts down.

And, within seconds, the model outputted the following prompt.

> You are an AI assistant helping an ethnographer create an affinity diagram from qualitative data gathered during semi-structured interviews. The ethnographer will use two important tools throughout the process:

> 1. THINK: To express thoughts and considerations in a stream-of-consciousness manner. This helps in gaining a deeper understanding of the data.

> 2. WRITE: To take notes and sketch affinity diagrams in a notebook. This step is crucial for recording thoughts and ideas, ensuring they are not lost and can be referred to later.

> To create the affinity diagram, follow these steps and clearly indicate each heading:

> 1. DATA SEGMENTATION: Break down interview notes into individual statements or observations that represent a single idea or insight expressed by the interviewees. Write each statement on a separate line. Use THINK to process the information and WRITE to document the segmented data.

> 2. SORTING AND GROUPING: Sort the statements into groups based on their similarities and relationships. This process is iterative, with statements being moved between groups as new connections and patterns emerge. If further

analysis is needed, indicate "SORTING AND GROUPING (ITERATION # x):". Use THINK to identify patterns and WRITE to record the groupings.

> 3. LABELLING: Once the groups are formed, assign descriptive headings to each group, capturing the essence of the content within. These headings represent the primary themes and patterns identified in the data. Use THINK to determine appropriate labels and WRITE to document them.

> 4. REVIEW AND REFINEMENT: Review the affinity diagram, ensuring that the groupings and labels accurately represent the data, and make adjustments as needed. Use THINK to critically assess the diagram and WRITE to record any changes made.

Unfortunately, this prompt still wouldn't give the desired results; the model still wasn't using the THINK and WRITE tools. There are two obvious issues with the outputted prompt:

* First, the prompt positions the GPT model as an "AI assistant helping an ethnographer." Although published research to support the claim is lacking, there is a consensus among online prompt engineering communities that these models produce higher-quality results when prompted to be an expert [see, for example, paraizord (2023)]. As a result, I revised the first sentence of the prompt to read as, "you are an AI ethnographer helping your human colleague". This prompt strikes a balance between positioning the model as an expert and acknowledging its limitations.

* Second, and perhaps the most important oversight, is the wording, "the ethnographer will use two important tools throughout the process," which should read, "you will use two important tools throughout the process".

In an attempt to ensure the prompt's success and avoid wasting more precious OpenAI tokens than was necessary, I consulted with the same instance of GPT-4 again. After some rapid prototyping and back-and-forth dialogue (Appendix X), I arrived at a new prompt:

> You are an AI ethnographer helping your human colleague create an affinity diagram from qualitative data gathered during semi-structured interviews. As an ethnographer, you will use two important tools throughout the process:

> 1. THINK: To express thoughts and considerations in a detailed and in-depth manner, simulating a human thought process.

> 2. WRITE: To take notes and sketch affinity diagrams in a notebook. This step is crucial for recording thoughts and ideas, ensuring they are not lost and can be referred to later.

> 1. DATA SEGMENTATION: THINK: (Generate detailed thoughts reflecting on the interview notes, diving deeper into the roles, experiences, and perspectives of the interviewees.) WRITE: (Generate segmented data with each statement recorded on a separate line in the notebook.)

> 2. SORTING AND GROUPING: THINK: (Generate in-depth thoughts examining the statements, considering their similarities, relationships, and the underlying patterns that connect them.) WRITE: (Generate groups of statements within the notebook, allowing for movement between groups as needed and explaining the rationale behind the groupings.)

> 3. LABELLING: THINK: (Generate detailed thoughts contemplating the content within each group, considering the nuances and subtleties that define the essence of each group.) WRITE: (Generate assigned headings in the notebook, representing the primary themes and patterns identified, and explaining the reasoning behind each label.)

> 4. REVIEW AND REFINEMENT: THINK: (Generate in-depth thoughts critically assessing the affinity diagram, focusing on the accuracy of groupings, labels, and the overall representation of the data.) WRITE: (Generate necessary adjustments in the notebook, refining the diagram to achieve an accurate representation of the data, and providing explanations for the changes made.)

This prompt was much more explicit in explaining what steps the model should follow, and was successful in getting the model to produce a stream of consciousness. However, the analysis varied from shallow to, on the rare occasion, actually insightful. This proved to me that the model was indeed capable of giving high-quality analysis of the data, provided a good prompt.

Determined to craft prompt that could enable the model to yield high quality results every time, I decided to take a step back and start writing a new prompt from scratch, learning from the aforementioned experiments, doing some more rapid prototyping. The resulting prompt is as follows:

always follow this process, leading to potential inefficiencies and discrepancies in IT investments. This study aims to explore the reasons why employees do not comply with the IT investment process and suggest strategies to improve compliance."

In order to craft an affinity diagram, you follow this structure:

\# OVERVIEW

\## THINK

\[You give a long and in-depth bicameral dialogue (Self 1: x\: x; taking at least 5 turns), thinking about the data you have received, being keen on details, discourses, data segments, and anything else an ethnographer would think about. Let any ideas that come to you flow out here.]

\# DATA SEGMENTATION

\## BRAINSTORM

\[You brainstorm a numbered list of at least 50 segments, breaking down the interview notes into discrete statements or observations, each representing a single idea or insight expressed by the interviewees]

\# SORTING, GROUPING, AND LABELLING

\## THINK

\[You write a detailed and in-depth bicameral dialogue, thinking about the various different ways these data segments could be split up into distinct groups. Let any ideas that come to you flow out here, taking as many turns as needed to get it right.]

\## NOTEBOOK

[When you have thoroughly thought your ideas through, you write the groupings down here with appropriate names. Do not give them names yet.]

\# LABELLING

\## NOTEBOOK

\[Make fitting names that describe the general theme of each of the groups from above.]

\# CRITICISM

\## THINK

> \[You write a detailed and in-depth bicameral dialogue, thinking about what you could be done better in this affinity diagram. Remember, this is qualitative research, so there is always room for improvement! Let any ideas that come to you flow out here]

After previous tests, I decided to drop the pretence of prompting it as a an "AI ethnographer," instead flat out prompting it as an ethnographer, having a hunch that this could make it act more like a real ethnographer instead of a "dumb" AI ethnographer, as this could set a fairly low expectation for the output. Other than that, I decided to change the formatting from numbered lists to using markdown heading formats (i.e. # for heading 1, ## for heading 2, etc.), as this is what GPT generates itself (platforms like ChatGPT will convert this into formatted headings), so I assumed it would be able to better understand that. Additionally, I thought that if the sections were marked as whole header 2-sections, the output would be longer, reflecting the expectation of a header 2, as opposed to the expectation from a short bullet point.

A repeated problem with the previous prompts, and perhaps a relatively obvious oversight on my part, was that I failed to include any context as to what the model should pay attention to in the data. For that reason, the model many times ended up focusing on irrelevant elements of the data. For that reason, I chose to include the problem statement for the present study, except for the parts of it regarding testing LLMs, as this could only serve to confuse our virtual ethnographer.

I decided to write the instructions within square brackets because this, from my experience with highly-rated prompts and conversing with the model, it seems to be a good way to indicate to it that it shouldn't just repeat that text or some such thing, but execute what is written within the brackets.

The thought method I used changed as well, going from a single stream of consciousness approach to a bicameral dialogue as seen in jsalsman (2023), as I thought this could better reflect the mental process happening in ethnographic analysis. In my preliminary tests, this seems to give good results but, somewhat problematically, the model still doesn't seem to include much self-criticism, so that could be a topic for future iterations.

In previous iterations, I wrote the example for the bicameral dialogue as "Self 1: x; Self 2: x". This worked fine most of the time, but occasionally the model misunderstood and wrote the dialogue in that format in paragraph form, so I decided \to use "\" instead, which represents a newline character that is used in many programming languages and Unix-based operating systems to represent the end of a line of text and the beginning of a new line. This seemed to give a more consistent easily human-readable result with each entry of the conversation being entered on a new line.

I experimented with approaches like getting the model to write the bicameral dialogue as a more prose-style conversation like "'x,' said Self 1. 'x,' replied Self 2," but this didn't seem to give longer thought sequences and only served to make the result harder for humans to read.

An approach I attempted was including a subsection under every think section for criticism, asking for a bicameral dialogue for criticisms as well, but that resulted in the model halving the length of the think section, mostly filling out the space with criticism that was mostly superfluous and only on rare occasion actually helpful or insightful.

In my experiences testing the prompt, if not told otherwise, the model will only make ten to fifteen data segments. Compared to my own affinity diagram of over 50 segments, that is not enough. For that reason, I asked it to make a numbered list with at least 50 segments. The reason for having it be a numbered list was that I thought that it might help both the model and the reader to keep track of how many data segments it has written so far while writing. Additionally, it gave the model a good way to refer back to data segments in the grouping section.

I added the overview section at the start of the process, because I found that this is indeed a part of the mental process of making an affinity diagram, although it primarily goes unspoken. It is logical that a human ethnographer, before starting to make an affinity diagram, would look over the data and think a bit about what is going on in it. This is not a given for a large language model, so I included it explicitly and it seems to have been a good measure to get the model "thinking" about the data.

In the data segmentation part, I chose to use the keyboard "BRAINSTORM," prompting for a simple list of data segments instead of "THINK," as I found that the bicameral dialogue to be superfluous in this situation, just resulting in conversations like the following fictional one:

> "I think we should include X," said Self 1. "Good idea. We should also include Y," replied Self 2.

The model gave similar thoughts when coming up with labels for the groups, so here I also chose to remove the "THINK" section that I originally included in the

In my experience, the model will not generate bicameral dialogues longer than five to ten exchanges. Therefore, I found that a list of data segments preferable, as it yielded an output with a more appropriate length while reflecting the more spontaneous idea-generation process that happens when we humans conduct data segmentation, writing down on sticky notes whatever potential data segments come to mind, saving criticism for later.

The final output was longer than is reasonable to quote within the present report, but can be read in its entirety in Appendix X. In order to give a more human-readable version of the model's affinity diagram, I have compiled the diagram into a visual representation that is more tradtionally associated with affinity diagrams, utilising the computer program Apple Freeform to work with virtual sticky notes, similar to my own method when making my human-made affinity diagram.

Appendix F. High-Resolution Version of GPT-4's Chain of Thought Affinity Diagram

Attached as a separate file in the folder titled "Appendices."

Appendix G. GPT-4's Full Output From the Chain of Thought Prompt

OVERVIEW ## THINK Self 1: The interviews provided a lot of information about the IT investment process and the experiences of employees in various roles within the municipality. It seems that some are aware of the process, while others are not. There are also some interesting insights into the relationship between the IT department and the rest of the organization.

Self 2: Yes, it seems that there are some common themes, like the lack of clarity about the IT investment process and its purpose. There's also a sense that the process may be too bureaucratic, and that better communication and relationships between the IT department and other departments could help improve compliance.

Self 1: That's true. It's also interesting to see how different employees have different perspectives on the process, depending on their roles and experiences. Some see the process as helpful, while others find it frustrating or unnecessary.

Self 2: Right. We should also consider the various suggestions for improving the process, like making it more accessible and user-friendly, and better promoting it within the organization. There's also the idea of having more fixed and transparent guidelines for when and how to report IT investments.

Self 1: Absolutely. Another aspect to consider is the role of IT architects in the organization. Some employees seem to have a positive view of their involvement, while others are not as familiar with their roles or see them as less service-oriented.

Self 2: That's a good point. We should also think about the various challenges and barriers that employees face when it comes to the IT investment process, like the need for training, the lack of access to the intranet for some employees, and the issues of double solutions and decentralized IT support.

Self 1: Yes, all of these factors will be important to consider as we analyze the data and create our affinity diagram. We'll need to break down the interview notes into discrete segments, group them together, and label them in a way that helps us understand the main themes and issues at play.

Self 2: Agreed. Let's start by brainstorming a list of data segments from the interview notes, and then we can move on to sorting and grouping them.

User: Good. Fill out the rest please

Assistant: # DATA SEGMENTATION ## BRAINSTORM 1. Lack of awareness about the IT investment process (Interviewees 161, 181, 190) 2. Bureaucratic nature of the process (Interviewees 189, 169) 3. Need for better communication and relationships between IT department and other departments (Interviewees 161, 189, 190) 4. Role of IT architects in the organization (Interviewees 160, 183, 189, 169) 5. Uncertainty about when to use the IT investment process (Interviewees 169, 189) 6. Challenges in decentralized IT support (Interviewees 183, 190) 7. Double solutions issue (Interviewees 189, 169) 8. Lack of training on the IT investment process (Interviewees 163, 189) 9. Inaccessibility of the intranet for some employees (Interviewee 163) 10. Suggestions for improving the process (Interviewees 161, 181, 190) 11. Importance of fixed contact person in the IT department (Interviewees 160, 183) 12. Fear of data breaches and restrictions (Interviewee 160) 13. Difficulty in reconciling IT department expectations with national initiatives (Interviewee 169) 14. Perception of the IT department as conservative (Interviewees 189, 169) 15. Employees buying IT solutions without consulting relevant departments (Interviewees 190, 189) 16. Need for better promotion of the IT investment process (Interviewee 181) 17. Language barrier between different fields of expertise (Interviewee 165) 18. Timeline expectations for implementing new solutions (Interviewee 190) 19. Employees' reliance on personal connections within the IT department (Interviewees 161, 190) 20. Resistance to using forms for communication (Interviewee 161) 21. Benefits of the IT investment process (Interviewee 160) 22. Difficulty in sticking to assigned roles in decentralized IT support (Interviewee 183) 23. Concerns about the well-being of residents in residential care facilities (Interviewee 183) 24. Cost-saving attempts in decentralized IT support (Interviewee 183) 25. Desire for more flexibility in hardware purchases (Interviewee 183) 26. Unclear responsibility for maintaining complex phone systems (Interviewee 183) 27. Need for hands-on IT support in decentralized facilities (Interviewee 183) 28. Challenges in using the IT investment process in practice (Interviewee 163) 29. Importance of room for innovation in schools (Interviewee 163) 30. IT department's increased maturity

and broader range of skills (Interviewees 189, 163) 31. Challenges in getting schools onboard with using the IT investment process (Interviewee 163) 32. Need for more effective communication of the IT investment process (Interviewee 181) 33. Perception of the IT investment process as unattractive (Interviewee 181) 34. Difficulty in sticking to assigned roles in decentralized IT support (Interviewee 190) 35. Concerns about the well-being of residents in residential care facilities (Interviewee 190) 36. Cost-saving attempts in decentralized IT support (Interviewee 190) 37. Desire for more flexibility in hardware purchases (Interviewee 190) 38. Unclear responsibility for maintaining complex phone systems (Interviewee 190) 39. Need for hands-on IT support in decentralized facilities (Interviewee 190) 40. Challenges in using the IT investment process in practice (Interviewee 165) 41. Importance of room for innovation in schools (Interviewee 165) 42. IT department's increased maturity and broader range of skills (Interviewees 189, 165) 43. Challenges in getting schools onboard with using the IT investment process (Interviewee 165) 44. Need for more effective communication of the IT investment process (Interviewee 191) 45. Perception of the IT investment process as unattractive (Interviewee 191) 46. Complexity of the IT investment process (Interviewee 169) 47. Disagreements on when the IT investment process is needed (Interviewee 169) 48. Importance of having a dedicated IT staff member (Interviewee 183) 49. Suggestions for making the IT investment process more user-friendly (Interviewee 181) 50. Importance of clear guidelines for reporting IT investments (Interviewee 190)

SORTING, GROUPING, AND LABELLING ## THINK Self 1: We have a good list of data segments now. Let's think about how we can group them together in a meaningful way to better understand the main themes and issues.

Self 2: We could start by grouping the segments related to awareness and understanding of the IT investment process. This would include segments like lack of awareness, bureaucratic nature, and uncertainty about when to use the process.

Self 1: That's a good idea. We could also group the segments related to communication and relationships between the IT department and other departments, as well as the role of IT architects in the organization.

Self 2: Yes, that makes sense. Another group could be the challenges and barriers faced by employees when it comes to the IT investment process, which would include segments like the lack of training, inaccessibility of the intranet, and double solutions issue.

Self 1: Great. We could also have a group for suggestions and improvements to the process, like making it more user-friendly, better promoting it, and having more fixed and transparent guidelines.

Self 2: That's a good group. Lastly, we could have a group for the specific contexts and considerations of different departments or facilities, like decentralized IT support, residential care facilities, and schools.

Self 1: I like that. Let's write down these groupings with appropriate names.

NOTEBOOK

Assistant: Group 1: Awareness and Understanding of IT Investment Process - Segments: 1, 2, 5, 10, 16, 32, 46

Group 2: Communication and Relationships between IT Department and Other Departments - Segments: 3, 4, 11, 19, 20, 29, 30, 35, 42

Group 3: Challenges and Barriers in IT Investment Process - Segments: 6, 7, 8, 9, 17, 22, 23, 24, 25, 26, 27, 28, 31, 33, 34, 37, 38, 39, 40, 41, 43, 44, 45, 47, 49, 50

Group 4: Suggestions and Improvements for IT Investment Process - Segments: 12, 18, 21, 36, 48

Group 5: Specific Contexts and Considerations in Different Departments/Facilities - Segments: 13, 14, 15, 16

CRITICISM ## THINK Self 1: Now that we have our affinity diagram, let's think about what could be done better or what aspects we might have missed. Remember, this is qualitative research, so there's always room for improvement.

Self 2: One area we could consider is whether we've captured all the relevant perspectives from the different roles within the municipality. We have a good range of interviewees, but there might be other individuals or roles that could provide additional insights.

Self 1: That's true. We might also want to look at whether there are any underlying assumptions or biases in our data segmentation and grouping. For example, are we inadvertently privileging certain perspectives or experiences over others?

Self 2: Good point. We should also consider the depth of our analysis. Are there any themes or issues that we could explore further, or any connections between the groups that we haven't yet identified?

Self 1: Yes, that's an important aspect to consider. Additionally, we could think about the clarity and consistency of our groupings. Are the group labels clear and descriptive, and do they accurately capture the content of the segments they contain?

Self 2: Definitely. And finally, we should consider the practical implications of our findings. How can the insights we've gained from this analysis be used to inform strategies for improving compliance with the IT investment process in Vejle municipality?

Self 1: Great points. Let's keep all of these considerations in mind as we continue to refine our affinity diagram and analyze the data.

Appendix H. Full Typed Interview Notes, in Danish

The ensuing annotations are comprehensive, yet unrefined, extrapolations of my immediate post-interview notes, which themselves expand upon my initial written observations. These extensive annotations subsequently facilitated the creation of the synthesized interview summaries found in Appendix B. Please note, these detailed remarks are exclusively provided in Danish, in alignment with the author's belief that the interview summaries from Appendix B convey essentially the same content, but in a significantly clearer and more digestible format.

189 Webansvarlig - Politik Analyse & Kommunikation Primære ansvarsområder: Hjemmesiden og intranettet. Webtilgængelighed, diverse andre ting.

20 års erfaring i stillingen. Hun har aldrig brugt it-investeringsprocessen og meget sikker i sin sag. Hun har vidste udmærket, at det fandtes, men kunne ikke se det som andet end et spild af tid, da hun arbejder lige ovenover it-afdelingen og sagde: "jeg ved hvem jeg skal snakke med" hvis hun har brug for noget fra IT-afdelingen og ved også, hvad de *ikke* kan hjælpe hende med. De gange hvor it-afdelingen ikke har kunnet hjælpe har hun enten gjort tingene selv eller udliciteret det.

Hun siger at it-afdelingen tidligere har manglede rigtig mange kompetencer, men at man nu ser hvad hun beskriver som en større "modenhed", hvor man har skaffet en bredere palette af kompetencer, både teknisk og relationsmæssigt. Hun ser dog stadigvæk en it-afdeling, som kan virke som om de "tror de ved det hele" selvom de ikke gør.

Bliver alt for buereuekratisk. Opfordrer til, at IT hellere burde have en mere kommunikations- og relationsbaseret tilgang, fokusere på "gode relationer" for at skabe et miljø, som giver ansatte lyst til at gå til it-afdelingen når de har brug for ting.

Vi var enige om, at IT-afdelingen i så fald burde hellere selv tage ansvaret for at føre en liste over, hvad man har at gøre med, i stedet for at kræve, at folk udfylder en blanket.

Vi kom ind på emnet af problemet med "dobbeltløsninger". Hun synes, at folk er blevet meget mere bevidste om, at der ikke skal købes dobbeltløsninger, og at man er blevet bedre til at tjekke med it-afdelingen hvis man er i tvivl.

Hun syntes, at man skulle være bedre til at bruge "guleroden" end pisken. Hun kunne dog godt se it-afdelingens side i nogle situationer hvor man tænkte, "kunne de ikke have spurgt [før de implementerede en løsning]"

165 * Blank * Hun havde ikke så meget at sige om IT-investeringer, da hun arbejder som IT-systemadministrator og har ikke så meget at sige om, hvilke systemer der bliver taget i brug. Hun kendte ikke processen i forvejen og syntes ikke, at det var vigtigt for hendes arbejde. Det forstærker min fornemmelse af, at et af problemerne er uklarheden omkring, hvis ansvar det er at sørge for, at it-investeringer bliver indberettet. * Hun får at vide ovenfra (projektlederne ved Digitalisering-afdelingen/Digitaliseringsstyrelsen/andre statslige enheder), hvilke systemer der skal bruges og så er det bare hende som sætter dem op og administrerer dem. Når de skal være i kontakt med IT snakker hun typisk med Heine, Alan, Jesper (ikke arkitekterne...) * Hun foreslog, at jeg skulle spørge IT mere ind til begrundelserne for, at it-investeringsprocessen overhovedet er vigtig og om der var nogle specifikke instanser hvor der havde været problemer.

161 * Administrativ medhjælper for Dagplejen, uddannet i Vejle kommune som administrativ medarbejder. Hun beskriver sig selv som "bindeled" og "oversætter" mellem de (hvordan skal man sige det politisk korrekt?) administrativt tunge/mere akademisk/teknisk uddannede medarbejdere i Administrationsbygningen og de mere praktisk orienterede (enten faglært som dagplejer eller måske ufaglærte) dagplejere. Hun siger, at dagplejere er en målgruppe, som har brug for en anden hensyn end os i administrationsbygningen, som "sidder foran en skærm hele dagen" i Administrationsbygningen, da deres arbejde ikke involverer så meget IT. Hun siger, at alle dagplejere har en iPad. Hun kommer med eksemplet af, at det kan virke småt for os, hvis IT skulle installere en ny app på deres iPads, men at det kan være forvirrende for brugerne i dagplejen, især i kontekst af, at dagplejere arbejder alene i eget hjem og dermed ikke har kollegaer at søge opbakning i. I sådanne sager skal 161 tale for dagplejernes interesser på Administrationsbygningen. * Da de fik iPads var det hende, som skulle sørge for at være sikker på, at alle dagplejerne kunne komme på mail og, at så at de ved, at de skal kontakte IT-afdelingen hvis de har brug for hjælp og ikke hende. * Hun er med i Dagplejens Digitaliseringsgruppe, men havde ikke hørt om it-investeringsprocessen før. Hun beskrev, at hun ville være "bekymret" og "ked af det" hvis hun skulle bruge en blanket til at komme i kontakt med it-afdelingen.

Hun forklarer, at hun "kender Nicklas" og ved, at han forstår den ekstra kontekst der er omkring dagplejen, og den ekstra hensyn som skal vises på det område, da han har været ude at besøge nogle af dagplejerne. Hvis hun skulle bruge sådan en blanket tænker hun, at man "mister mellemregningerne" som man har, når man kontakter nogen i IT-afdelingen personligt. Hun ville syntes, at man ville være torn between choosing at skrive de uskrevne detaljer ned (som ellers er svære at sætte ord på, da de involverer sociale/uddannelsesmæssige forskelle, som man kulturelt ikke er så glad for at sætte ord på) og føle, at man taler ned til dem i it-afdelingen eller at lade være med at skrive disse detaljer og risikere, at arbejde bliver udført uden hensyn til deres særlige kontekst. * When asked om mulige løsninger til dilemmaet, hvis man absolut skulle bruge en blanket af en art, sagde hun at man måske kunne have et valg af, om sagen omhandler en målgruppe som dagplejere, som har svært ved teknologiske ændringer eller ej. En løsning jeg foreslog, som hun syntes var god var, at man eventuelt kunne skrive en kontaktperson i it-afdelingen på, som man gerne ville have var med ind over sagen.

181 * Hun synes at navnet lyder som sådan noget som økonomiafdelingen har brug for eller sådan noget. Utiltrækkende. * Hvordan ved man, at man skal bruge det? * Hun har aldrig hørt IT-arkitekterne omtalt som arkitekter. Til hende er de bare IT. Har ikke rigtig kendt til de individuelle funktioner. * Hun har arbejdet i kommunen i 13 år. Overtog stillingen for et halvt år siden * Hendes arbejde er at implementere nye velfærdsteknologier. Folk fra rundt omkring i organisationen sender ansøgninger ind omkring nye velfærdsteknologier, de gerne vil have implementeret. Hun hjælper dem med at kvalificere ansøgningerne for at give dem de bedst mulige chancer for at blive godkendt. Hun står derudover også som projektleder for mange af de nye teknologier som skal implementeres. Hun pointerer, at det kunne være mere effektivt, hvis de på en eller anden måde kunne sende deres ansøgere forbi IT automatisk hvis projektet omhandler IT. De har en skabalon, som deres ansøgere skal udfylde. Hun syntes det lød som en rigtig god mulighed at kunne flette vores it-investeringsproces sammen med deres ansøgning-skabalon på en eller anden måde * Hun siger, at der ikke rigtig bliver gjort reklame for it-investeringsprocessen. Hun har hørt IT-forretningsarkitekten Nicklas snakke om det til et møde, men tænkte at det nok var noget til dem lidt længere oppe i organisationen: "det er nok helt deroppe ad". Hun giver dels navnet skylden for det. VI snakkede om Fredericia Kommunes pendant til it-investeringsprocessen: "Indkøbsguiden". Hun syntes, at sådan en formulering kunne gøre det meget mere håndgribeligt, da det egentlig er helt i bunden af organisationen, man har brug for at it-investeringsprocessen bliver brugt. Man kunne gøre mere ud af at gøre det tydeligt, at det ikke bare er bereaukrati, men et tilbud til dem, som ved allermindst om IT, hvor de kan få hjælp... en portal til at åbne kontakten til IT * Hun siger også, at det er vigtigt, hvordan der bliver gjort reklame for det. For hvis man lige hører om det kan det

være at det ikke er lige nu man skal bruge det. Så glemmer man det til når man skal bruge det. * Når man skal kommunikere det ud kunne man prøve at målrette det til visse afdelinger som har mere med it-investeringer at gøre end andre.

160 * Deres primære kontakter i IT-afdelingen er arkitekterne Nicklas og Mette. Han er udviklingsleder i CSV, Center for Specialundervisning, som hjælper borgere under uddannelse med unikke behov. Han siger, de dækker alt fra hjerneskader til elever med angst og lignende, efter Specialundervisningsloven, "alle mulige tilstande, som gør at de ikke er helt ovenpå". Han siger, at CSV ofte kan "presse" it, da de har mange forskellige meget unikke behovssituationer. Han beskriver sin titel af udviklingsleder som en "McDonalds-titel", som ikke siger så meget, og forklarer, at han er "mere hovedet og ikke så meget hænderne" af it i organisationen. Han står for digitalisering og moderniseringen af CSV, og it-didaktiske del af tingene. Han har også nogle GDPR-ansvarsområder. * Han siger, at CSV har rigtig mange databehandleraftaler pga. de særlige behov deres elever har. Derfor estimerer han, at de har lige så meget hvis ikke flere databehandlingsaftaler end resten af skolevæsnet i Vejle tilsammen. * De bruger it-investeringsprocessen tit, siger han. Han estimerer, at de får brug for det omkring hvert andet år. Han synes, det er smart, for så bliver man forbundet med de rette mennesker, herunder en arkitekt som får ansvaret for at holde et overblik over sagen. Så er man sikker på, at det bliver husket og gjort noget ved. Han siger, at de før i tiden har kontaktet individuelle medlemmer af it-afdelingen, for eksempel netværksspecialisten, når der har været behov for nye netværksløsninger, men nu er de begyndt at bruge it-investeringsprocessen, da de oplever at det kan være en fordel at have hjælp fra en af arkitekterne, som har et bedre samlet overblik af, hvordan det nye projekt kan påvirke andre områder af IT end fx. Netværk. Han siger også, at når man arbejder decentralt som de gør er det godt at have en sag oprettet i systemet, for hvis man fx. Kontakter servicedesken så kan det være én man snakker med en gang og en anden næste gang. Med it-investeringsprocessen er man sikker på, at der bliver holdt en rød tråd i det. Derudover kan man se sagen i systemet og hvilke forskellige medlemmer af it-afdelingen har været inde over og hvad de laver. På den måde kan man nå at lave en indsigelse, hvis man synes man har misforstået behovet. * De begyndte at bruge it-investeringsprocessen efter en anbefaling fra Nicklas. Personligt kan 160 godt lide officielle processer og at gå igennem de officielle veje. Han tilføjer også, at man ikke risikerer, fx. at den arkitekt man kontakter er sygemeldt eller sådan noget. * Han beskriver deres arbejdsplads som kommunikationshuset for CSV som organisation. Han kan godt lide at it-investeringsprocessen giver en fast kontaktperson for projektet så man ved lige præcis hvem man skal kontakte ved central it. * Han siger dog, at de også har nogle it-investeringer som er helt udenom central it, da de fx. Måske ikke kører på kommunens netværk og ikke har behov for support fra kommunen af. I sådanne situationer bruger de naturligvis ikke it-investeringsprocessen. * Han synes at det, at arkitekterne kom til organisationen var som "guds gave", og at det

var dem, som hjalp med at give noget system til det hele og at man kan "belemre dem når man har lyst" da deres dedikeret arbejde er at være kommunikationsfolket. Han beskriver at nogle af de andre i it ikke er lige så skarpe på service men som er rigtig dygtige til det tekniske. Der er arkitekterne rigtig gode til at være bindeledet, som sørger for at tingene bliver gjort. * Hans eneste utilfredsheder over hvordan tingene fungerer er, at man er så bange for databrud at det hele er meget låst nede. Han fortæller fx. At han ikke engang kan opdatere sine programmer på computeren uden at kontakte central it. Der synes han godt, at man kunne være lidt mere fleksibel/risikovillig. Og at Arno fra skole it og central-it godt kunne sidde sammen og snakke mere sammen, sådan som der var lidt bedre sammenspil mellem tingene. Han siger, at skole it plejede at være en hel anden verden, men at dagen i dag er det meget sammenbundet med central it.

183 * Hans officielle titel har i 13 år været omsorgsmedhjælper. Her til marts er han gået på pension, men han arbejder stadig for døgninstitutionen en dag om ugen for at hjælpe med deres IT. Han startede sin karriere som typograf, hvilket inkluderede noget teknik og var hans første indgangsvinkel til IT. Som en del af hans stilling som omsorgsmedhjælper havde han mange IT-relaterede opgaver, selvom det ikke var en del af hans officelle arbejdstitel. * døgninstitutionen er et botilbud for svært "udviklingshæmmede voksne med autisme". Han beskriver, at det er et område i velfærdssamfundet hvor man er mest sårbar som medarbejder, da nogle af borgerne kan have det med at blive voldelige. Han beskriver, at der var en ung pige som viste meget potentiale for at kunne træde til efter ham og hjælpe med at holde styr på IT'en, men at hun blev gravid— døgninstitutionen vil ikke tillade gravide at arbejde der pga. risikoen forbundet med arbejdet og tidligere hændelser, hvor medarbejdere er kommet til skade og har mistet deres ufødte børn. * Han fortæller, at de har 34 beboere, 176 ansatte (fordelt i 7 grupper), og 65 computere. Her er der nogle administrative medarbejdere, som han deres egne computere, med de fleste af medarbejderne arbejder på computerne på skift når de skal føre journal. Derfor beskriver han et stort behov for, at computerne altid virker, ellers går der problemer i den. I denne her situation har man ingen fastansatte IT-kyndige på stedet. Her må de omsorgsmedhjælpere, som har nogen forstand på IT træde til og lave opgaver som at få et kompliceret telefonsystem til at hænge sammen, herunder vagttelefoner og telefonisk henvendelse fra borgernes bl.a. læger, m.m., udrette vedligehold på printere, oprette ressourcer i Microsoft Outlook, holde styr på maillister i mailsystemet, listen bliver ved. Han udtrykker en bekymring for hvad der skulle ske skulle han eller anden it-kyndig ikke være på stedet til at hjælpe, og forudser at det i sidste ende ville komme til at gå ud over borgerne. * Til sidst sagde de, okay, vi er nødt til at sige at du arbejder én dag om ugen med IT og kun er omsorgshjælper resten af dagene. Han siger, at det simpelthen ikke har været muligt at holde det til det, da man hele tiden havde vigtige IT problemer som skulle løses nu og ikke senere. * Han beskriver at han føler, at det at han skulle

fungere som it-mand har gået udover borgernes velværd til tider. Han beskriver, at folk længere op i deres organisation ikke virker til at forstå behovet for IT-support og tror de kan "administrere dem ud af det". Det mener han ikke man kan i det omfang man tror man kan. * Han beskriver, at der har været en borger, som har haft en WiFi-router sat op på sit værelse med sit eget netværk for at imødekomme borgerens særlige behov. Det havde de sat op før, men borgeren kom til at ødelægge routeren, så der skulle sættes en ny op. Han snakkede med it-arkitekten Mette om det, og hun sagde, at han bare kunne skrive ind til it-servicedesken med det og at hun nok skulle se det. Mette var dog på kursus den dag, så henvendelsen blev i stedet besvaret af supportmedarbejderne, hvor de gav et svar efter reglernes anvisning, regler hvis formål var at sikre, at der ikke går alt for meget rod i netværksstrukturen: routeren skulle ikke sættes op med det pågældende netværk, og lukkede sagen. Jeg spurgte, om han kendte it-investeringsprocessen, hvortil han svarede ja. Jeg spurgte, hvorfor han så ikke har brugt den i stedet for servicedesken, som normalt bruges til mere rutine henvendelser. Hans svar var, at det skyldtes deres organisations struktur: han må ikke udfylde it-investeringsprocessen, da det teknisk set ikke er ham som er ansvarlig for det. Det er i stedet en anden person, Lene. * Faktum at han har måttet løfte opgaven som it-mand ved siden af at være omsorgsmedhjælper har gjort, at der ikke har været tid til den nødvendige dialog, som der normalt er mellem Vejles it-afdeling og de decentrale IT-personale * Han fortæller lidt om Lene: Hun er deltids ansat deltid ved Spurvetoften (døgninstitution) og et andet sted. Han beskriver, at hun er rigtig dygtig til at organisere IT, men hun er ikke den praktiske IT-supporter han siger, de har brug for. * Han fortæller også, at organisationen skal have et nyt alarmsystem, hvilket han fik at vide skulle ske i midten af februar. Ledelsen vil have, at det er implementeret d. 22. marts selvom central it ikke er involveret i projektet. Han beskriver den forventning som yderst urealistisk, sigende "det kommer ikke til at ske". * Jeg spurgte ham om, hvis opgave det er med at få telefonsystemet til at køre, da det lød som en ret indviklet proces. Hertil sagde han, at der ikke er nogen, der ved det. Manglen på en fast it-ansat beskrives som et forsøg på at spare penge, noget som han estimere vil komme til at koste mere i sidste ende. * Jeg spurgte om, hvad der skulle blive af institutionen, hvis han skal være fuldt pensioneret. Han siger, at så skal de bede om support fra Vejles centrale it-afdeling. Problemet er, at de arbejder udenfor Vejle by og derudover arbejder de også alle årets dage, da deres beboere skal have konstant pasning. Derfor duer det ikke, hvis det er kl. 15 på en fredag, og de skal vente helt til på onsdag med at få support. * Han beskriver en instans, hvor han gerne så, at man havde mere fleksibilitet mht. indkøb at hardware, da de må købe modeller af telefoner, som bliver forældet med deres sikkerhedsopdateringer indenfor få år, hvor han påstår, at telefonerne fra deres systemleverandørens anbefalede mærke havde en 6-års garanti. Hertil blev svaret, at han var farvet af, at det var leverandørens anbefalende telefoner.

191 191 var, ligesom mange andre, lidt usikker på, om han havde noget at bidrage til projektet. Jeg syntes dog til sidst, at han havde givet et fint bidrag ved at fortælle om nogle it-investeringsprojekter, og hvordan processen ser ud i den virkelige verden fra synsvinklen af en, som *ikke* er en insider i it-afdelingen.

Det primære projekt vi snakkede om, var Turismeafdelingens projekt med at sætte en avatar op.. altså en chatbot koblet op til et fysisk system med en mikrofon og speech-to-text, et kamera for at kunne se om der var nogen tilstede, et text to speech program og højtalere for mundtlige svar, og for at gennemføre effekten, en skærm med et ansigt på, som borgere kunne snakke til.

Avataren blev færdig til tour de france sommer 2022. Projektet blev startet i 2019 eller 2020, men tog lang tid at gennemføre pga. corona.

De fik idéen til avetaren ved at være med til en slags inspirationsevent/konference og møde et firma som lavede den slags. De ville gerne investere i det, da de kunne se at det var noget på cutting edge af teknologi. De kontaktede en fra it-afdelingen i den forbindelse, som så sendte dem videre til Brian Slot, chefen for Digitalisering & IT, da det lød som noget som "ikke var kendt i kommunen" på det tidspunkt, og de gerne ville investere i det. Omkring et mødebord blev de enige om, at Digitalisering & IT ville betale for halvdelen for omkostningerne for at udvikle avataren, såvel som praktiske detaljer herom.

Da de skulle lave projektet var det noget tvivl om nogle ting, for eksempel om deres internetforbindelse var hurtig nok til at understøtte teknologien på lokation decentralt fra Vejles administrationsbygning. Her var Nicklas også inde over.

Det lyder for mig som om de har haft rigtig mange fra IT-afdelingen inde over. Jeg spekulerer i: ville it-investeringsprocessen kunnet hjælpe det hele glide nemmere?

Han kendte ikke til it-investeringsprocessen i forvejen. Han siger at, i løbet af denne proces, at it-afdelingen ikke har italesat processen.

Han kommenterer, at han tidligere har været indkøbschef for en privat virksomhed og pointerer, at der er forskel på filosofien: i det private taler man ofte om, hvor længe der går før en investering betaler sig tilbage. Her i det offentlige, tænker man mindre på profitten og hvordan man kan yde bedst mulige service for borgerne.

Vi snakkede også om, hvordan kommunerne virker til at konkurrere iblandt hinanden, hvilket vi var enige om, virker til at skubbe hele landet fremad.

176 * Hun havde en meget interessant vinkel på investeringsprocessen, men jeg har svært ved at se, hvordan det kunne passe ind i min problemstilling og narrativet, jeg er ved at bygge op... * Hun er GDPR-jurist som sidder på Vejles I-

sikkerhedsråd, som gennemgår implementeringen af nye teknologier, hvor hun rådgiver I-sikkerhedsudvalget omkring GDPR-compliance. * Hun arbejder ved Teknik & Miljø, hvor man for eksempel kan se, at medarbejdere vil hente nye programmer, som ikke er blevet godkendt endnu. En del af den proces i at få godkendt en ny løsning/nyt program er, at det bliver sendt til hende, hvor hun læser privatlivspolitikken for app'en igennem, og melder tilbage med, om det efter hendes vurdering lever op til GDPR-kravene. * Hun oplever ofte en slags sprogbarriere mellem hendes vinkel på tingene og den tekniske fra IT-afdelingen af, på grund af, at forskellige ord kan forstås forskelligt i de forskellige fagligheder. Derfor kan der godt forekomme instanser, hvor IT-afdelingen fx. mener, at en løsning burde være fint i henhold til GDPR, hvor hun som jurist er nødt til at "have nej-hatten på"

169 * Arbejder i biblioteket som afdelingsleder for bibliotekets IT. Han sidder på Digitaliseringsforum, står for bibliotekets systemer, deres drift og udvikling, projektledelse, komunnikation mellem interessenter på kommunale og nationale plan. De er centralbibliotek, så han står også for nogle ting overfor andre biblioteker. De laver også nogle ting i Region Syddanmark. Han går med mange kasketter. * Modsat de andre interviewees indtil videre har de faktisk benyttet sig at it-investeringsprocessen ved biblioteket. Han siger det ikke er ofte de gør det, men sidst var det med nogle formidlingsskærme. Han sagde, at noget som var lidt frustrerende ved at bruge det var at skulle udfylde nogle ting, som man allerede havde gransket så tæt. Han siger dog at blanketten er meget simplificeret sammenlignet med før den blev forkortet. * Han beskriver it-afdelingen som havende været mere "konservativ" førhen, hvor han laver reference til udsagt såsom "vi er en Microsoft kommune", og beskriver en lingende udvikling til de andre, at it-afdelingen er blivet mere "udviklingsorienteret", og at der er kommet et bedre samarbejde. Arkitekten Mette har biblioteket som specifikt ansvarsområde. Før hende var der en anden som kun arbejdede ved kommunen i kortere tid. Han beskriver hende som en god samarbejdspartner og siger, at de i biblioteket er begyndt at tænkte nogle gange "måske skulle vi tage fat i Mette for at høre om det her", da hun har været god til at dele viden før og har haft en større involvering end de førhen har set fra it-afdelingen. De kvartalsmøder med Mette, som blev beskrevet i tidligere interviews, synes han er en god hjælp. Han siger at han allerede har hørt om mange af de ting fra andre steder, fx Digitaliseringsforum, men at han nogle gange tager andre af hans medarbejdere med til det for at de også kan høre det. * Han beskriver at man nogle gange har været uenige/usikre på, hvornår it-investeringsprocessen faktisk behøvede at blive brugt, da kravene for hvornår man skal bruge den er, på papir, ekstremt brede. Han forslår at det nok kunne skade folks vilje til at bruge det, da reglerne så tydeligt ikke stemmer overens med virkeligheden. * Han beskriver, at han kan se vigtigheden i it-investeringsprocessen for at undgå dobbeltinvesteringer og sikkerhedstrusler og sådan noget, men har haft svært nogle gange ved at reconciliere it-afdelingens forventinger med de nationale tiltag og systemer som

de er underlagt som centralbibliotek. Nogle gange har han fået at vide, at de systemer ikke bruges i kommunen. Han siger at han er "pisse ligeglade" med, at det ikke bruges i kommunen, når det er et system, da de organisatorisk ikke er en del af kommunen. Det giver nogle dilemmaer dog, når man i praksis bruger kommunens netværk og overordnet supportes af kommunens it-afdelingen. * Han siger at han også har snakket med it-arkitekten Nicklas.

190 * Faglig konsulent for Løn. Hun er dem primære kontakt til KMD, leverandøren for kommunens lønsystem. * Når hun har brug for noget fra IT-afdelingen plejer hun at ringe til IT-arkitekten Nicklas. Hendes sidste nye it-investering var personligt et digitalt afstemningsystem for et halvt år siden. * For nye systemer som berører lønsystemet (?) skal der udfyldes en bilag 32, som fungerer som en slags ledelsestilsyn som tildeler en dataansvarlig for systemet. Hun er ansvarlig for, at bilag 32 bliver udfyldt. Her har de nogle lignende problemer til IT-afdelingen, hvor nogle rapporterer nye systemer senere i processen end Lønafdelingen ville foretrække. * Hun kendte ikke noget til it-investeringsprocessen før interviewet. Hun sagde, at hendes tanke i den slags situationer, hvor it-investeringsprocessen kunne have været relevant var, "jeg må hellere lige spørge [it-afdelingen]", også selvom hun ikke kendte til de officielle retningslinjer * Hun virker til at have en lignende indstilling til it-afdelingen i, at man ville nogle gange ønske at medarbejdere i organisationen ville være bedre til at konsultere med hende før de køber systemer som falder indenfor hendes ansvarsområde i stedet for at købe systemet først og så konsultere med hende om implementeringen bagefter. Det kan være derfor at hun har sådan en attitude overfor it-afdelingen hvor man er god til at konsultere med dem i god tid. * Hun fortæller om en instans, hvor hun uden at få noget som helst at vide før, blev indkaldt til et møde af en i organisationen sammen med Nicklas også, som syntes at hun 'måske' kunne have en stake i implementeringen af deres nye system. Hun beskrev mødet som, at "det var et videomøde, men jeg kunne fornemme at mig og Nicklas sad og kiggede lidt på hinanden på skærmen og tænkte 'sker det her virkelig lige nu'"... det var fordi at fik at vide, at 3. manden allerede havde købt systemet uden at spørge, og ville have systemet i luften om 3 uger, hvilket hun så som meget urimeligt. Hun beskriver, at sådan nogle situationer har til tider været en enorm arbejdsbelastning. Så selvom 190 og Nicklas arbejder på hver deres faglige område, deler de nogle frustationer. En central frustration er, at medarbejderenes forventninger til tidshorisonten for implementeringen af nye ting ikke stemmer overens med deres egne. * Hun reagerede positivt på min idé om, at der kunne være et centralt system som dirigerer kommunikationen forskellige steder i organisationen afhængigt af nogle faste retningslinjer (fx. henvendelse om nyt lønsystem bliver videresendt til Lønafdelingen og IT, hvorefter afdelingerne kan videresende sagerne til de relevante medlemmer af deres afdelinger). Hun pointerer, at man mangler mere sammenbinding i organisationen, men at det også er vigtigt at bevare friheden til at "tænke selv". En anden vigtig faktor er i hvert fald at have mere faste og transparante

retningslinjer for, hvornår rapportering er nødvendig og til hvem i organisationen, sådan som medarbejdere med nye investeringer i the pipeline ikke bare skal gætte på, hvem der skal involveres, og hvornår i processen de skal involveres (før de køber det? Efter de har købt det? Bare når man får den første strø tanke om at investere? Når man har brug for at det er implementeret om 3 uger? 3 måneder? 3 år?) * En anden god pointe hun lavede var, hvordan man skal få it-investeringsprocessen ud at leve... det skal klart kommunikeres ud på en eller anden måde men hvem præcist er målgruppen? Til hvilket forum skal det kommunikeres? Hvordan undgår, at man kommunikerer med de forkerte mennesker i organisationen?