

Designing LLM driven tools
**to support collaboration in
complex multi-stakeholder
systems**

MASTER THESIS

Karel te Marvelde

March 28, 2025

MASTER THESIS

Designing LLM driven tools to support collaboration in complex multi-stakeholder systems

AUTHOR

Karel te Marvelde
Msc Design for Interaction
Faculty Industrial Design Engineering
Delft University of Technology

CHAIR

Dr. I.J. (Ingrid) Mulder
Department of Human-Centered Design

MENTOR

Ir. S.L. (Sterre) de Jager
Department of Design, Organization and Strategy

IN COLLABORATION WITH

Marguerite Evenaar
SE.Lab



Participatory City Making
Delft Design Labs

ACKNOWLEDGEMENTS



With this project, I set out to contribute new insights to the field of design. My research presented me with the opportunity to thoroughly explore complex collaborative systems. I now say that I am part of the growing group of individuals that are trying to overcome the unique challenges presented by complex problems. Furthermore, it allowed me to develop new skills and become proficient in Large Language Models and LLM powered tools, a topic that I will continue working on even after this research has been completed. However, I could not have written this thesis without the support of the following people:

To Ingrid and Sterre, thank you for your patient, insightful, and kind supervision throughout this project. Your guidance was invaluable during the many moments I felt stuck, and I'm especially grateful for the trust you showed me, even when I didn't always successfully articulate my process.

To Marguerite, thank you for involving me in your work with SE.Lab and for sharing your knowledge, exchanging ideas with me, and brainstorming new opportunities together.

Gina, Hans, Irene, Leonoor, Marina, and Sabine, thank you for sharing your experiences and allowing me to pick your brains. Your insights were truly appreciated.

Finally, I want to thank my family, friends and my partner Vita for supporting me throughout this project.

I hope that you will find this research insightful and hopefully it might even inspire some of you - *Karel te Marvelde*

ABSTRACT

The world is grappling with increasingly complex challenges, from global issues like climate change to local concerns such as rising asthma rates in communities near industrial zones. Tackling these problems requires collaboration between stakeholders across different domains, resulting in highly intricate, multi-stakeholder systems. These complex collaborations introduce unique challenges that make finding effective solutions difficult. While current design methodologies offer frameworks to support collaboration, their effectiveness diminishes as complexity grows, highlighting the urgent need for new, adaptive tools.

This research explores the potential of Large Language Model driven tools to enhance collaboration within these complex systems. It identifies key barriers to effective cooperation, including misaligned stakeholder values, communication breakdowns, and entrenched power dynamics. By addressing these challenges, LLM-powered tools offer a promising new approach to facilitate more inclusive, efficient, and adaptive collaborative processes. Through a combination of literature review and expert interviews, three critical themes that impact the success of collaboration in complex systems were identified: Value Alignment, Communication & Certainty, and Power Structures. Traditional design methods are evaluated against these themes, highlighting their limitations in adequately addressing the complexities inherent in multi-stakeholder collaborations. To overcome these limitations, the study investigates the potential of LLM-based tools, notably leveraging OpenAI's ChatGPT-4o model, to facilitate improved stakeholder interactions, streamline communication, and balance power dynamics.

These tools scale traditional design methods by automating the identification and analysis of stakeholder values, making it feasible to work with large, diverse groups. They improve communication by tailoring information to different audiences, translating jargon, and reducing misunderstandings, ultimately lowering uncertainty and increasing stakeholder engagement. LLMs also support value alignment by extracting and comparing stakeholders' goals and perspectives from textual data, helping to identify shared priorities and potential conflicts early in the process. Furthermore, they contribute to flattening power hierarchies by democratizing access to information, enabling all participants, regardless of expertise or status, to contribute meaningfully. By generating meeting summaries, co-creation materials, and scenario analyses, LLMs support more inclusive, agile decision-making and help maintain momentum in collaborative projects.

LLM-powered tools act as cognitive and communicative amplifiers in collaborative systems. They don't replace human interaction but enhance it—by improving understanding, managing complexity, and reducing the friction that typically arises in stakeholder collaboration.

TABLE OF CONTENTS

1. INTRODUCTION	P. 06
1.1 SE.Lab and Meerwijk	
1.2 Report structure	
2. COLLABORATION	P. 10
2.1 Exploring the literature	
2.2 Expert interviews	
2.3 Identifying the common themes	
3. VALUE, COMMUNICATION AND POWER	P. 18
3.1 Value alignment	
3.2 Communication & certainty	
3.3 Power structures	
4. COMPLEX SYSTEMS AND PROBLEMS	P. 33
4.1 Increasing complexity	
4.2 The challenges of complexity	
4.3 Using large languages models	
5. PROBLEM DEFINITION	P. 42
6. METHODOLOGY	P. 44
7. LARGE LANGUAGE MODELS	P. 45
7.1 Ethics and risks	
7.2 Prompt engineering and testing	
8. EXPLORING THE POSSIBILITIES OF LLMS	P. 53
8.1 Generating ideas	
8.2 Testing and developing LLM tools	
9. DESIGNING A LLM POWERED TOOLBOX	P. 69
9.1 Making a tool selection	
9.2 Combining all insights	
10. CONCLUSIONS	P. 95
11. DISCUSSION AND RECOMMENDATIONS	P. 96
12. PERSONAL REFLECTION	P. 98

1. INTRODUCTION

In 2018 the Dutch Ministry of Public Health put out a report on the aging Dutch population stating the following:

"The Dutch population is gradually aging. This means that the proportion of elderly people in the total population is increasing. Currently, there are 1.3 million people aged 75 and older. By 2030, this number will rise to 2.1 million, and by 2040, it will reach a staggering 2.5 million. Of the current 75+ population, 92% live independently, and even two-thirds of those aged 90+ still live independently.

Nearly a quarter of the 75+ individuals living at home use assistance and care services from multiple domains. The demand for care will increase: while most people over 75 are still vital, about 38% of this group is considered vulnerable, according to the SCP (Netherlands Institute for Social Research). As a result, by 2030, there will be 1 million vulnerable elderly people.

At the same time, the number of available informal caregivers is expected to decrease. Currently, an individual aged 85+ can rely on an average of 15 relatively "younger" elderly (aged 50–75). By 2040, this number will drop to just six. Moreover, due to the increasing aging population, the demand for suitable housing for the elderly will rise.

These developments present a significant challenge for all of us."

The problem put forth in this report demonstrates a complex or "wicked" problem, an issue with no clear single solution, which involves many interconnected factors and stakeholders with possibly conflicting interests. These wicked problems can't be solved with standard linear problem solving methods and require an alternative approach (Roberts, 2001). There are numerous papers published describing how wicked problems can be tackled, including notable ones written by Nancy C Roberts. In her papers, Roberts mentions three main types of strategies for solving wicked problems. The first of which are authoritative strategies, which focus on reducing complexity by reducing the amount of stakeholders involved. The second are competitive strategies, in which the opposing interests and views of stakeholders are used to pit the stakeholders against each other in the hopes of boosting innovation. The final type are collaborative strategies, where all stakeholders involved are incentivized to work together in order to find the best solution for everyone. It is this last group of strategies that has gained the most traction in recent years (Roberts, 2001; Khademian, 2008; Hagemann & Kluge, 2017; Strachwitz et al., 2021).

Although the collaborative approach appears to hold the most promise for addressing wicked problems, it still has its limitations. As a result, even in 2025, many complex issues, such as the rapidly aging Dutch population, remain without a definitive solution.

1.1 SE.LAB AND MEERWIJK

In neighborhood of Meerwijk, Social Enterprise Lab (SE.Lab) leads a model pilot to transform it into a living example of integrated solutions for pressing transition challenges. The pilot focuses on three key pillars: providing appropriate care, enabling comfortable living with innovative housing solutions, and fostering an engaged community through shared ownership and collaboration. Together, these pillars form the foundation of "Caring Neighborhoods," where localized support enables residents, particularly the elderly, to live meaningful, active, and connected lives.

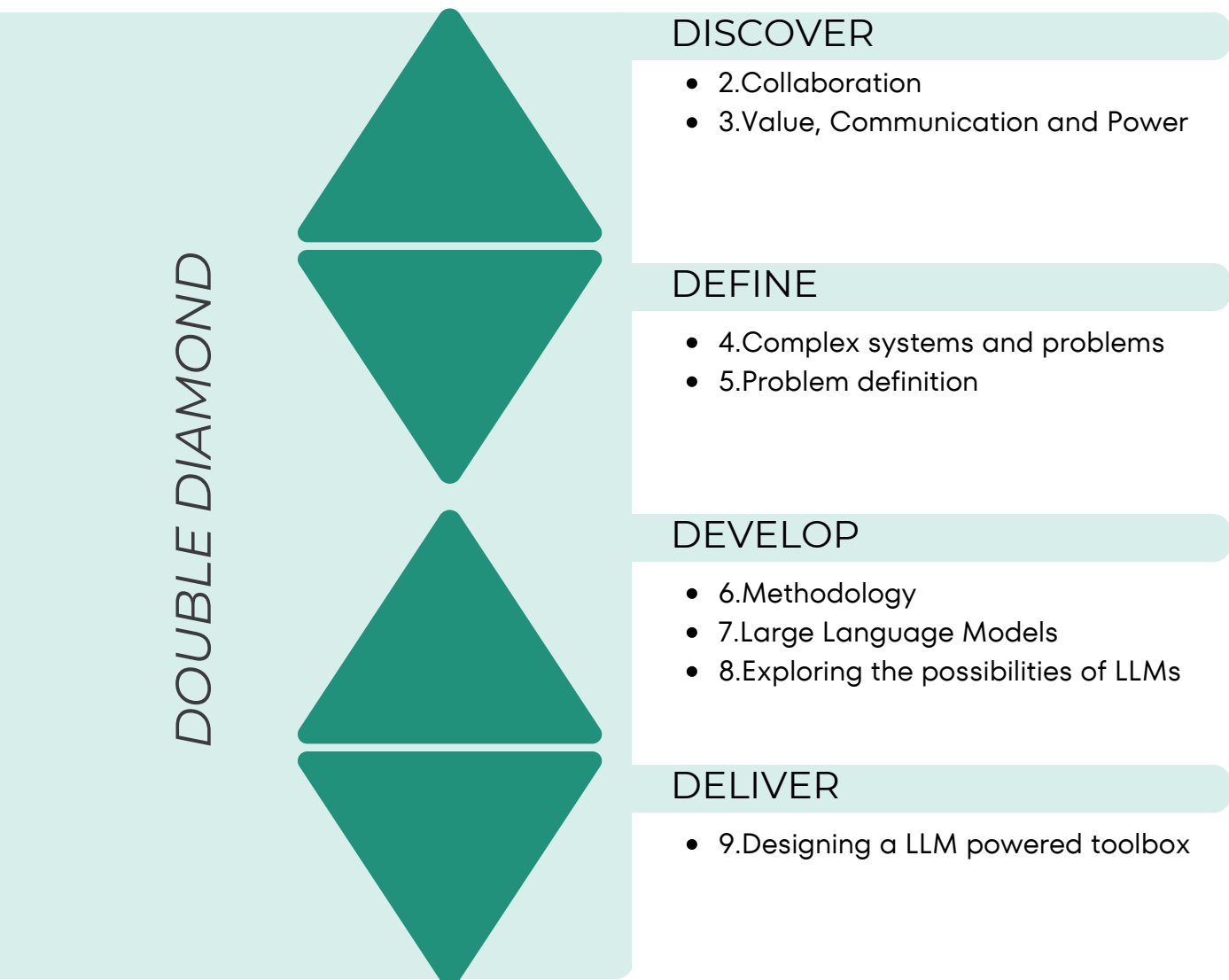
This collaborative project not only experiments with new organizational approaches but also establishes a growing network of changemakers, individuals who help enable transition in their direct environment. Over the coming year, SE.Lab and its partners will facilitate co-creative workshops, implement monitoring and evaluation strategies, and disseminate their findings through conferences and publications. Their aim is to embed sustainable practices, empower local actors, and scale these innovative models to broader contexts. SE.Lab leverages transitions as powerful solutions for today's complex societal challenges. Through a unique combination of strategic insight, practical execution, and systemic thinking they guide both public and private partners in designing and implementing sustainable (societal) change. With a pragmatic and research-driven methodology SE.Lab is dedicated to translating vision into reality. In doing so they contribute to building resilient structures that can anticipate and adapt to the evolving societal context for now and the future.

The Meerwijk pilot case is, however, not without its issues. Just like most actors that try to tackle complex problems, SE.Lab has encountered a number of obstacles that could hinder their endeavors. Aligning diverse actors in a complex system is a challenge. SE.Lab recognizes that bringing together actors from the "living world" (local community members) and the "system world" (institutional stakeholders) is difficult. These groups often operate with different values, languages, and priorities, making alignment challenging. SE.Lab has also observed a resistance to change. Established systems and traditional approaches often resist innovation, limiting the capacity for experimental methods to take hold. This resistance can come from individuals, organizations, or systemic rigidity in parts of our society. Innovation may be stifled if key stakeholders are unwilling to adapt or prioritize co-creative methods. Finally, demonstrating the value of co-creative processes is critical, but a lack of proven methodologies for measuring and validating these approaches complicates advocacy for their adoption. Without clear evidence of success, gaining support from stakeholders and scaling efforts becomes more difficult. SE.Lab has found that stakeholder alignment, unwillingness to deviate from existing structures, and a demand for proof of success are key factors that are preventing successful collaboration.

SE.Lab, together with the TU Delft Participatory City Making Design Lab, wants to explore how collaborative endeavors, such as in Meerwijk, can be structurally improved, not only to guarantee a more impactful outcome but also to stimulate a global transition in the field of social innovation for complex problems.

1.2 REPORT STRUCTURE

Collaboration is increasingly recognized as the key to addressing complex societal challenges and SE.Lab is at the forefront of the global shift towards a new, collaborative model. However, significant obstacles remain, largely due to the complexity of modern problems. To better understand and overcome these challenges, this research applies the Double Diamond design and innovation method. This framework consists of two key phases: divergent thinking, where possibilities are explored, and convergent thinking, where decisions are refined and focused. A major advantage of this approach is that it ensures researchers define their focus clearly before delving into solution development. For a topic as broad as collaboration, this structured approach is essential. Given its many possible interpretations and applications, collaboration presents an almost infinite solution space. The Double Diamond method helps navigate this complexity, ensuring that research remains both systematic and impactful.



The first half of the first diamond is the discovery phase, aimed at understanding all relevant problems, their causes, and their ripple effects. **Chapters 2 and 3** initiate this research by mapping key factors that significantly influence collaboration, drawing from literary research and expert interviews. From these factors, a set of unifying themes is established to guide the rest of the study. These themes are then explored further in a literature review on collaboration, examining their potential benefits when properly managed and the risks of neglecting them. Additionally, design theory schools that address these themes are identified, with specific tools and methods provided as examples.

The second half of the first diamond is the defining phase, aimed at synthesizing discoveries to identify key issues. **Chapters 4 and 5** focus on the aspects of complex systems and problems that may explain why existing design theory alone is insufficient to address these challenges. From these findings, the need for a new model of collaboration emerges, prompting an analysis of current limitations to lay the foundation for the next section of this report. Finally, Large Language Models (LLMs) are introduced, emphasizing their strengths and potential applications in enhancing collaboration on complex problems.

The first half of the second diamond is the development phase, where brainstorming, testing, and experimentation with potential solutions take place. **Chapters 6, 7, and 8** explore the possibilities of LLM-based tools. First, the ethical implications and risks of using LLMs are examined, leading to the creation of guidelines for safe and responsible use in design. Next, prompt engineering, the process of designing LLM interactions, is studied to understand how to develop the most effective models. Finally, several ideas are implemented and tested using OpenAI's ChatGPT-4o model, providing insights into the possibilities, limitations, strengths, and weaknesses of LLMs.

The second half of the second diamond is the delivery phase, where the total number of possible solutions is narrowed down into a final set of solutions. **Chapter 9** brings together the key insights gathered throughout this research. A final selection of tools is made and mapped onto a collaborative project structure to create a LLM powered toolbox. Additionally, a handbook is provided that instructs future researcher and developers on how to continue building LLM powered tools for specific collaborative contexts.

The text in this paper has been reviewed and refined with the assistance of a large language model to improve spelling, grammar, and clarity.

2. COLLABORATION

Humans rely on collaboration, it is how we have become the dominating species on this planet. No single person has to be able to do everything because we are able to make use of a vast network of other individuals that can help us (Boyd & Richerson, 2009; Tomasello et al., 2012). However, as our society becomes increasingly more complex, we encounter more complex challenges, challenges that our current collaborative capabilities are not quite capable of tackling. In order to uncover methods that improve our collaboration, we first need to identify which factors influence the success of collaboration.

Collaboration in multi-stakeholder systems involves many factors. To help streamline this research, the most important factors are identified and grouped together to form clusters with a common theme. These collaborative themes will then be used as the focus for further literature research in the next chapter.

To identify these important factors, both literary research as well as interviews with stakeholders from the context were performed. The literary research focuses on topics such as collaboration, collaboration theory, factors impacting collaboration and collaboration in multi-stakeholder systems. The experts come from a variety of backgrounds, each working on projects where they collaborate with multiple stakeholders.

KEY TERMINOLOGY

- **Factors & Themes:** Factors are specific elements that influence collaboration, such as trust or clarity. Themes are broader patterns or underlying causes that give rise to these factors. For example, both trust and clarity relate to how stakeholders communicate. Because of this shared connection, they are grouped under the theme of communication.
- **Value:** The term value can have many meanings, but in this research, it is used in two distinct ways. First, values refer to the principles or priorities that stakeholders consider important during a project, these shape their decisions and behavior. Examples include transparency, equality, or profit. Second, value also refers to the intended outcome of the collaboration. It serves as a measure of a project's success. For instance, if the goal of a project is to reduce homelessness, then value is achieved when fewer people are homeless at the end of the project
- **Communication & Certainty:** Communication refers to all exchanges of information between stakeholders. Certainty describes the level of confidence and trust stakeholders feel when engaging in a collaboration.
- **Power Structures:** Power structures refer to both the explicit and implicit relationships between stakeholders. These can be formal, such as hierarchical roles within government organizations, or informal, like the dynamic between a boss and an employee.

2.1 EXPLORING THE LITERATURE

The collaborative problems observed by SE.Lab in complex systems, such as Meerwijk, provide a starting point for this research. However, additional exploration of the factors that impact collaboration is needed. To do this, various sources were analyzed to find the most commonly mentioned factors which have an impact on collaboration (Martín-Rodríguez et al., 2005; Rybníček & Königsgruber, 2018; Jiang & Ritchie, 2016; Tang & Shen, 2012; Letaifa, 2014; Weber & Khademian, 2008; Cui et al., 2021; Valk & Kratovič, 2021; Mysore et al., 2019; Hara et al., 2003; Gajda, 2004; Vangen, 2016). These factors are:

Trust & Relationship Building: Trust and relationship building are highlighted as fundamental to successful collaboration. Trust enhances communication, reduces conflicts, and fosters long-term cooperation among stakeholders. It is built through consistent, transparent interactions and mutual respect. Weak trust can lead to misalignment, disengagement, and adversarial relationships.

- *"Trust amongst organizations was described as a basic requirement for the dynamic formation of collaborative networks. Previous positive partnerships can establish trust and rapport between stakeholders and easily create shared goals for collaborative management."* (Jiang & Ritchie, 2016)
- *"Many authors consider trust as another important relationship factor in fostering a collaboration between industry and universities. Mistrust, in turn, influences the information flow and can lead to a departure from the original focus of a collaboration project. Therefore, partners need to spend sufficient time on establishing mutual trust."* (Rybníček & Königsgruber, 2018)
- *"Intolerance is rooted in the trust and culture of the engaged globally distributed stakeholders, and when trust is weak, conflicts and blame games emerge."* (Mysore et al., 2019)

Power Dynamics & Competition: Power dynamics and competition significantly influence stakeholder collaboration, either fostering productive cooperation or causing conflicts and disengagement. When power is unevenly distributed, dominant stakeholders may exclude others from decision-making, leading to resentment, mistrust, and resistance.

- *"These demands for resources is influenced by power and politics that need to be interpreted and assessed against the adverse intolerance levels of stakeholders who assign scarce resources within the engagement"* (Mysore et al., 2019)
- *"Collaboration is negatively affected by competition, adherence to hierarchy, and the desire to dominate, all of which create grounds for conflict."* (Valk & Kratovič, 2021)
- *"According to some authors, successful collaboration between health care professionals requires a shift from traditional hierarchical structures toward more horizontal structures. In fact, traditional structures do not facilitate the emergence of key conditions for collaboration, such as shared decision-making or open and direct communication."* (Martín-Rodríguez et al., 2005)

Communication & Information Clarity: Clear, open communication facilitates knowledge sharing, decision-making, and stakeholder alignment, while poor communication leads to misunderstandings, conflicts, and inefficiencies.

- *"The impact of communication has been extensively addressed in the investigated literature. With respect to communication, the frequency of communication is vital to create a shared understanding."* (Rybníček & Königsgruber, 2018)
- *"Miscommunication between stakeholders can result in incorrect assumptions, delays, and conflicts, all of which hinder collaboration and project success."* (Mysore et al., 2019)
- *"When communication structures are decentralized and flexible, they support teamwork, shared decision-making, and ultimately foster collaboration."* (Martín-Rodríguez et al., 2005)

Stakeholder Alignment & Shared Goals: Stakeholder alignment and shared goals are crucial for successful collaboration, ensuring that all parties work toward mutually beneficial outcomes. When stakeholders have divergent interests or unclear objectives, conflicts, inefficiencies, and disengagement arise. A shared vision enhances cooperation, trust, and long-term success, making stakeholder alignment a requirement for effective multi-stakeholder engagement and collaborative problem-solving.

- *"Shared understandings of what stakeholders can collectively achieve together through collaboration are crucial for long-term success."* (Jiang & Ritchie, 2017)
- *"Divergent interests and unclear objectives among stakeholders often result in conflicts, inefficiencies, and disengagement, ultimately harming collaborative efforts."* (Letaifa, 2014)
- *"Collaboration is not always without problems, and one should be consciously prepared to deal with potential barriers, which may include differences in goals and priorities among stakeholders."* (Valk & Kratovič, 2021)

Governance & Leadership: Governance and leadership play a pivotal role in ensuring effective collaboration by providing structure, decision-making clarity, and accountability. Strong leadership fosters stakeholder trust, alignment, and problem-solving, while weak governance leads to miscommunication, conflicts, and inefficiencies.

- *"A commitment to governance within government means that interdependencies are inevitable and that leadership plays a prominent role in structuring collaboration."* (Weber & Khademian, 2008)
- *"The initiatives that have been analyzed show the importance of concrete aspects of organizations, such as the management of human resources and leadership"* (Martín-Rodríguez et al., 2005)
- *"Group members and even the group leaders themselves often did not know who the group leaders were.. This suggests the importance of structural mechanisms of organization as a stimulus for seeing connections that lead to both complementarity and integrative levels of collaboration"* (Hara et al., 2003)

2.2 EXPERT INTERVIEWS

The literature review highlighted five overarching trends that influence the success of collaboration. To gain an even deeper understanding, seven individuals with experience in multi-stakeholder projects, each addressing various challenges, were interviewed to share their insights on collaboration and the obstacles they encountered. The following individuals participated in this research, all of whom work with multi-stakeholder systems in some capacity and can be considered experts through their experience.

- **Gina Gommer** is part of the AMS Institute's Living Labs, with a focus on the Energy Labs. In this role, she facilitates projects that bring together a diverse range of stakeholders, including government, commercial enterprises, and academic institutions.
- **Hans Roeland Poolman** has played strategic roles across high-tech, telecom, and industrial sectors. Most recently he has worked with AMS institute to solve complex multi-stakeholder societal problems such as the energy transition.
- **Irene Teerink** was a trainer and advisor at the expertise center for Bureau Jeugdzorg. She developed training and coaching programs for both youth care employees and external parties. She collaborated with behavioral scientists, youth care trainers, community organizations, education professionals, and many other stakeholders to identify practical challenges and determine the best methods to use in youth care.
- **Leonoor van Dam van Isselt** has two roles: she is trained as a physician specializing in elderly rehabilitation and oversees the rehabilitation section in Delft. Additionally, she is a research professor in Leiden, conducting research on elderly rehabilitation and teaching various student groups.
- **Marguerite Evenaar** is Director of Social Enterprise Lab U.A., an organization that leverages transitions as powerful solutions for today's complex societal challenges. In her work she uses strategic insights and a research-driven and pragmatic approach in guiding cross-sector collaborations, translating complex analyses into actionable change that bridges traditional frameworks and builds resilient, adaptive ecosystems.
- **Marina Bos de Vos** is an assistant professor and researcher at the TU Delft currently working on value alignment theory and methodology for complex multi-stakeholder systems.
- **Sabine van Gastel** is a Parkinson's healthcare specialist in both hospitals and nursing homes. She has a management function and plays a central role facilitating communication between different stakeholders such as: hospitals, nursing homes, apothecary and general practitioners.

The interviews followed an open structure (Appendix A), focusing on two main topics and basing further questioning on the responses of the participants. The two overarching topics were: what collaboration looked like in their sector or work environment, and whether an established structure or methodology was already in place to support it. From their responses four main topic clusters were created (Appendix B).

Remarks about collaboration, mentioning the difficulty of collaboration and the impact of bad collaboration on the success of a project.

- Marina - *"At the start, everything remains quite broad and high-level, so everyone can generally agree. But as the project progresses, you begin to uncover misunderstandings and realize where assumptions have been made. Often, these assumptions are never explicitly addressed and linger in the system for a long time until a conflict arises that makes it clear some stakeholders are not aligned."*
- Sabine - *"The main obstacles that have prevented this (Reforms in the healthcare sector) from happening so far are because people generally find it difficult to work together with people outside of their domain."*
- Leonoor - *"Collaborating on a larger scale is challenging because larger systems often do not communicate with each other. For example, cooperation between different home care organizations can be difficult. Within hospitals, it is somewhat easier since the system is less fragmented."*

Remarks about power dynamics that exist between stakeholders in collaborative projects.

- Irene - *"When providing training to employees, we frequently encountered resistance. This resistance likely stems from the top-down nature of decision-making, where employees feel unheard in their day-to-day work. Additionally, they face conflicting demands—being pushed for both quality and quantity—forcing them to complete numerous tasks in a short time frame. In some organizations, it is also difficult to admit struggles or fears."*
- marina - *"In collaborations, there are often certain parties that hold all the power. In discussions, they are always very open, but when push comes to shove, they are the ones who make the final decision. Some choices simply aren't made because certain parties know that a dominant party won't participate if those options are chosen, which leads to other participants dropping out as well."*
- Sabine - *"I feel like there is a lot of knowledge among the caretakers but these people in the top don't use it. I feel like they use too much of a top down approach. They give commands instead of collaborating. Hierarchy still plays a large role. Another thing that prevents proper collaboration is that everyone has opinions about people and organizations in the healthcare sector and those opinions influence their decision making."*

Remarks about commitment and uncertainty and how stakeholders need proof of success before joining, especially if there is a monetary aspect involved in the project.

- Marguerite - *"These organisations don't fully involve themselves in the project because you can't immediately promise them an outcome. They send someone that participates but it's just them, not the whole organisation."*
- Gina - *"As long as all interests are safeguarded, stakeholders are much more willing to take risks. However, this needs to be clearly communicated in advance."*
- Hans - *"Uncertainty and risk play a significant role in these types of collaborations. There is often a large difference between stakeholders in terms of how much risk they are willing to take. The question is how to mitigate risk for those who find it difficult to handle."*

Remarks about shared values and goals and how, when these are not aligned, collaboration becomes difficult.

- Hans - *"Currently, you often end up with a sort of zero-sum game where, if I give something to someone, I don't immediately see how I get something in return. This creates a race to the bottom, which I don't believe in at all. In these kinds of complex situations, you really need to focus on: Where does our shared value lie? But also, what are our conflicting values? You need to map these out, and it has to happen as early as possible."*
- Gina - *"People often want to step into a meeting and just 'bam, bam, bam,' discuss things as they usually would, and then topics like values or joint value creation aren't addressed at all. If you don't clearly define from the start how each stakeholder derives value from the project, you end up with a kind of expectation management throughout the project, which eventually results in the project falling apart."*
- Irene - *"We use methodologies that already incorporate core values. For example, Signs of Safety and solution-focused work are built on foundational values, which also serve as the criteria for measuring success. However, the organizations we worked with often had their own values, which sometimes led to conflicts."*

2.3 IDENTIFYING THE COMMON THEMES

By combining insights from SE.Lab, literary analysis, and expert interviews, this research compiles a dataset encompassing many different factors that influence collaboration. Solving for each factor individually would be beyond the scope of this study and would only address surface-level issues. Instead, this research aims to tackle the underlying challenges by identifying trends among these factors. Based on the factors identified the following themes to categorize the factors in where chosen (Figure 1).



Figure 1: The three collaborative themes

The three identified themes can be described as follows:

Value alignment: Already identified by SE.Lab as the difference between the system world and living world and further mentioned by many of the other experts, the theme of value alignment or misalignment encompasses stakeholders ability to identify what they see as their core objectives, goals and vision. Being able to align and create shared value as a result is a core theme for collaboration. Value, in this research, refers to both the inherent drivers for stakeholders, as well as the end result that is worked towards collaboratively .

Communication and Certainty: Interlinked concepts, the theme of communication and certainty revolves around the exchange of information and its effects on participation. When uncertainty is high and trust is low, it affects the success rate of collaboration. Communication plays a critical role, as it is both affected by uncertainty and the tool for reducing uncertainty. Having good communication and creating certainty for stakeholders is thus crucial for good collaboration.

Power structures: When there is collaboration there are power structures, and how these structures are shaped has a significant impact on the success of collaboration. Mentioned by both SE.Lab and experts, the way people see themselves in relation to other stakeholders greatly influences how they work together. Identifying how these dynamics can be shaped into a productive structures is critical for good collaboration.

It is important to note that whilst the three themes are mentioned individually both in literature as well as by experts, there is also overlap between them (Figure 2). For example, value alignment requires proper communication and the way communication takes place between stakeholders is heavily influenced by the power structure present between those stakeholders. These three themes lay at the core of collaborative endeavors and will be further investigated in the next chapters to identify how companies and organizations can benefit from proactively addressing these themes and how ignoring them impacts successful collaboration.

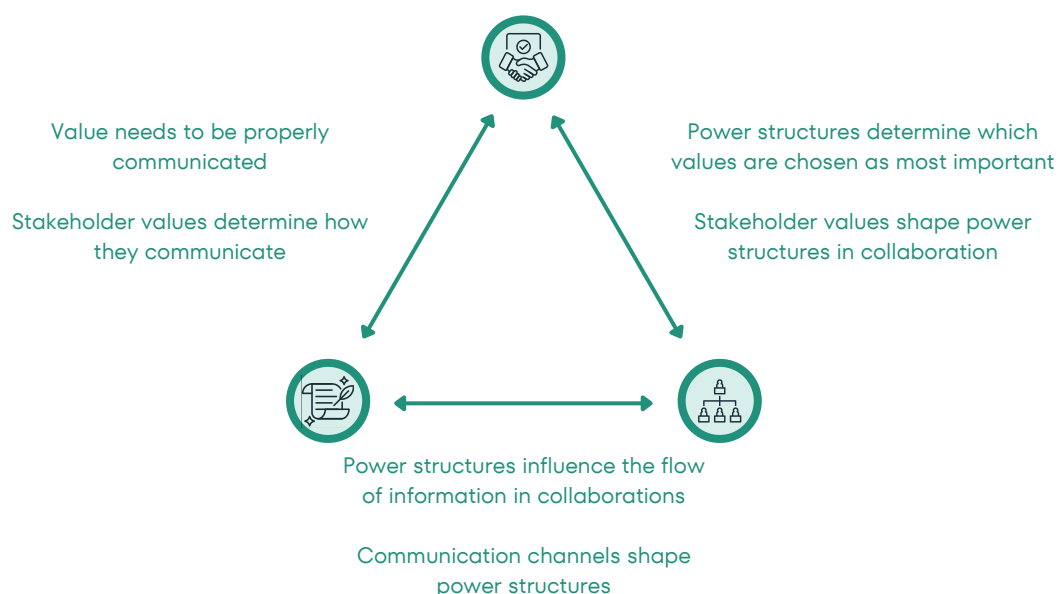


Figure 2: Examples of overlap between the three different themes

3. VALUE, COMMUNICATION AND POWER

As the three most important themes of collaboration have been identified in the previous chapter, this section of the report will delve deeper into the literature surrounding these themes. Each sub chapter in this sections examines the role of one of these themes within the context of collaboration. This involves outlining the positive impact of a theme on collaboration, and the consequences of ignoring it. Furthermore, this chapter explores how current design theory tries to deal with these themes in collaborative systems, listing specific design methodology and tools that are commonly used.

KEY TERMINOLOGY

Stakeholders: Stakeholders can be both individuals or groups that partake in a collaboration. Other terms often found in literature are Actors or Participants. Some examples of stakeholders are: Neighborhood residents, Hospital employees, Government representatives, etc.

Value alignment: Value alignment refers to the practice of stakeholders of identifying values that are important to them and properly communicating these to other stakeholders. When this happens, using various methods, the stakeholder can find common ground and align their values to improve collaboration.

Design methodology: Design methodology refers to the different schools inside the field of design. Example of design methodologies are Value centered design, communication design or participatory design.

Design tools: Design tools are specific tools and frameworks that have been developed by different schools of design. Examples of design tools are Service Blueprints, Mental models or Stakeholder maps.

3.1 VALUE ALIGNMENT

Value is both a core driver and an inherent aspect of collaboration. When collaborating, different stakeholders will focus on, prioritize and bring different values to the table. Companies and organizations from across all sectors are beginning to realize that identifying value can be extremely valuable for collaboration (Thomsen, 2004; Aschhoff & Vogel, 2018).

Value in Organisations and Companies

Thomsen (2004) has identified some of the key roles that value can play within an organization, outlining its function in governance, decision making, stakeholder relationships and conflict resolution. Thomsen states that value serves as a unifying factor within multi-stakeholder projects, offering a crucial mechanism for resolving conflicts and aligning diverse perspectives. When values are prioritized, they help balance the needs and concerns of various stakeholders, ensuring equitable outcomes. Furthermore, a values-driven approach fosters an atmosphere of respect and understanding during conflicts, promoting collaboration and reducing adversarial tensions. When stakeholders have conflicting interests, shared values become critical in managing these differences. These values help prioritize actions, ensuring that the organization remains focused on its overarching goals while navigating the complexities of stakeholder expectations. The alignment of these values across all parties involved, contributes to a sense of unity, even in challenging situations, and helps the organization maintain its strategic direction.

Within companies and organizations, the decision on which values are essential often arises from a combination of stakeholder expectations, regulatory requirements, and the organization's historical, cultural, or industry-specific context (Thomsen, 2004; Keeley, 1983). For instance, stakeholders may emphasize sustainability, transparency, or innovation as central values if these align with societal demands or competitive pressures. A key driver in deciding which values to prioritize is the implicit "contract" between different stakeholders. These implicit agreements are not formalized but represent shared expectations about ethical conduct, social responsibility, or operational focus (Fritzsche, 1991). Organizations often make these determinations through stakeholder engagement, market analysis, and an assessment of competitive advantages that align with certain values. For example, an organization operating in renewable energy might prioritize environmental sustainability, as this is central to its market positioning and stakeholder trust. Additionally, organizational values may also be shaped by policies and regulations imposed by higher-up forces within the collaborative hierarchy. For instance, political developments can lead to new regulations that impact the healthcare sector. Inside organizations, decision-making structures, such as hierarchical management or decentralized leadership models, also affect how values are chosen and integrated (Aschhoff & Vogel, 2018). In hierarchical structures, decisions about values may stem from top executives and reflect a more centralized vision. In contrast, decentralized organizations may adopt a more participatory approach, involving input from various levels of the organization to align values with broader operational perspectives.

Identifying Value

Identifying stakeholders' values can be challenging as an outsider. As previously mentioned, stakeholder values are often context-specific and can shift depending on situational factors. Additionally, some stakeholders may not openly express their true priorities, either due to strategic reasons or lack of clarity about their own values (Mitchell & Lee, 2019; Bridoux & Stoelhorst, 2013). Accurately identifying stakeholder values requires careful selection of research methods. Approaches like surveys, interviews, and participatory workshops each have limitations. For instance, surveys might miss nuanced perspectives, while interviews can be time-consuming and may reflect interviewer biases (Hosseini & Brenner, 1992).

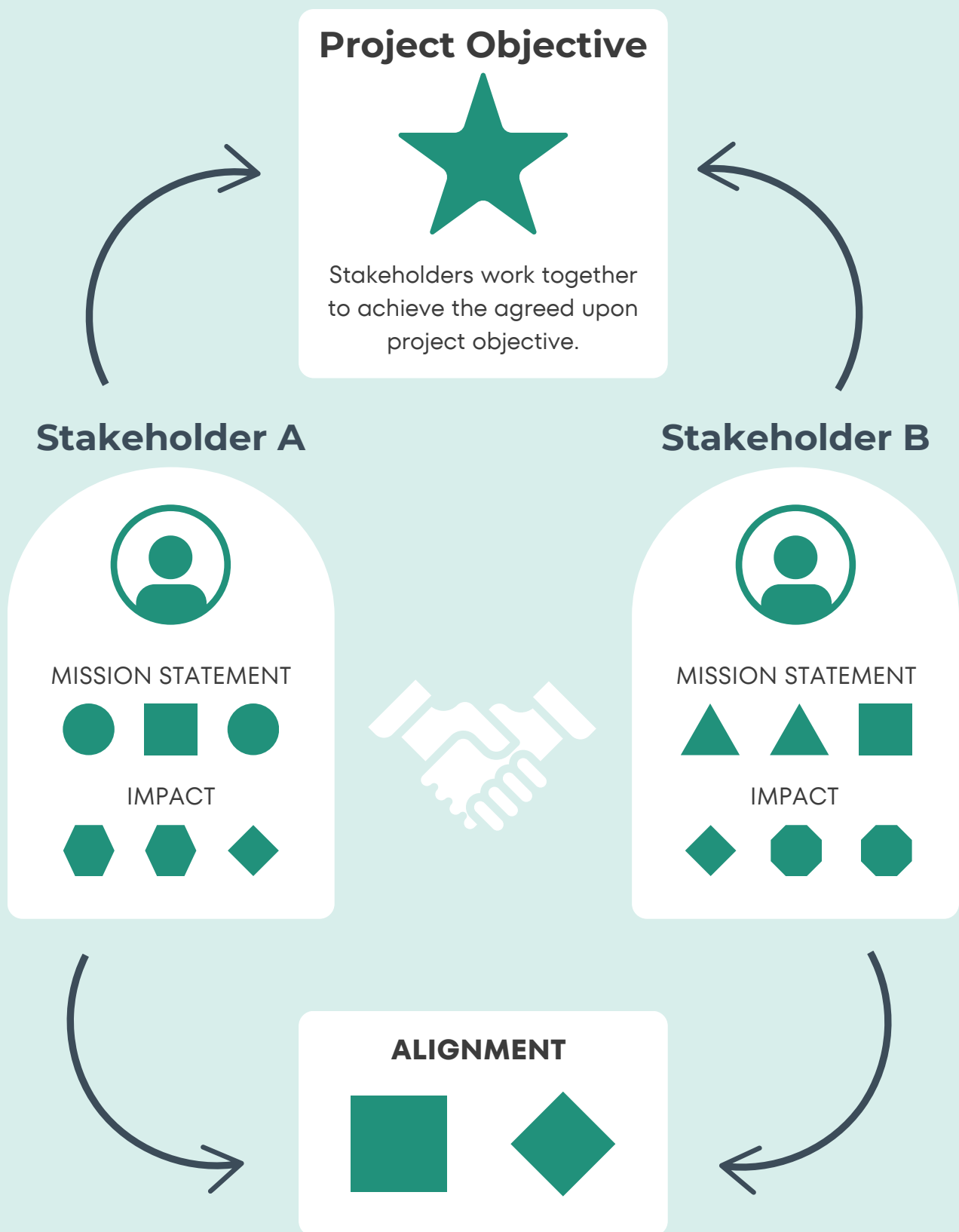
The Need for Alignment

Companies and organizations exist in a tangle of differing and often misalignment values (Freeman, 1984). There are the organization's own values, a set of guiding principles and priorities, often shaped by its leadership, mission, and organizational culture. Then there are the individual values of each employee, their own personal beliefs, ethics, and motivations. And finally, there are collaboration values, the values that an organization prioritizes when interacting with other stakeholders. Achieving alignment between the organization's missions, strategies, and values is crucial for successful collaboration between stakeholders (Murphy et al., 2014; Kwawu & Hughes, 2008). If an organization's values do not align with those of its employees, or if there is misalignment between the organization and its external stakeholders, successful interactions and agreement become difficult (Freeman, 1984). Research has shown that better alignment directly results in more value created for all parties involved (Austin 2000; Austin and Seitanidi 2012).

Murphy et al. (2014) mentions the following forms of value as some of the benefits that can be generated by stakeholder collaboration when values are aligned:

“access to new knowledge, expertise, or networks; increased financial or technological resources; improved legitimacy, reputation, and name recognition; improved stakeholder relations (including increased employee morale and retention); reduced environmental impact; and increased consumer patronage—all of which may lead to competitive advantages of one sort or another.”

Proper alignment of values and the resulting successful collaboration can be seen as a form of value creation on its own. Value creation is both an outcome of strong cross-sector relationships and a motivator for ongoing collaboration, as value creation typically strengthens the relationships between partners (Murphy et al., 2014). SE.lab labels these two types of value as the 'Mission statement', the inherent values that are important to each stakeholder, and the 'Impact', the value that stakeholders hope to achieve through collaborative projects. The challenge lies in accurately identifying and effectively communicating each stakeholder's mission statement and impact goals throughout the project. While stakeholders may align on the overarching objective, collaboration often falters when they fail to articulate their interpretations of key values, the specific aspects of the overarching objective that matter most to them, and their criteria for a successful and valuable outcome (Figure 3).



Additionally, stakeholders also hold inherent values (mission statement) and have their own vision of a valuable outcome (Impact). Aligning both of these is crucial for a successful collaboration

Figure 3: The process of value alignment between stakeholders

Value in Design

Value alignment in collaborative projects enhances communication, reducing misunderstandings and fostering cohesion. It provides a shared framework for conflict resolution, promoting respect and minimizing tensions. Aligned values also improve decision-making by ensuring choices reflect collective priorities while maintaining strategic focus. Organizations that align values with stakeholder expectations build trust, strengthen relationships, and encourage long-term engagement. This has been well understood in the field of design. Unlike traditional methods, which often rely on linear processes and predefined solutions, design thinking fosters creativity, empathy, and adaptability, enabling teams to better understand the needs of other stakeholders and the nuances of interconnected issues (Yaegashi et al., 2019). Within the field of design exists two main branches focused on value: Value Sensitive Design and Value Centered Design.

Value Sensitive Design

Value Sensitive Design (VSD) is a field of design study with a well-established tradition of exploring the complexity of stakeholder values, Integrating ethical considerations into the design process by systematically addressing diverse stakeholder perspectives (Dantec et al., 2009; Friedman & Hendry, 2019). VSD is a design approach that centers design decisions around the core values and needs of the people who will use or be impacted by a system, product, or service. Rather than focusing solely on functionality or profitability, value design integrates social, ethical, and cultural considerations, ensuring that the final outcome aligns with what matters most to stakeholders (Friedman et al., 2013; Friedman & Hendry, 2019). This approach requires actively engaging with stakeholders to understand their diverse values such as sustainability, inclusivity, accessibility, or privacy, and embedding those insights into the design process. By doing so, value sensitive design helps create solutions that are not only effective but also meaningful, equitable, and better suited to the complexities of real-world contexts. Insights from Value Sensitive Design can improve our understanding of how to foster transitions in complex systems by focusing on the values of the stakeholders involved.

Value Centered Design

Closely related to VSD is Value Centered Design (VCD). VCD is a design philosophy focused on delivering value not just to the customer but also to the enterprise and society at large. Traditional design approaches often emphasize product attributes such as functionality, aesthetics, or cost without a direct focus on value itself. However, VCD views the design process as a means to generate value, aligning design decisions with the broader needs and expectations of all stakeholders involved (Randmaa et al., 2012; Reber & Duffy, 2005). Value in VCD is not an inherent characteristic of the product or service but is instead derived from how a user or stakeholder interprets its worth in a specific context. This means value is subjective, dependent on individual needs and personal value systems, and can change based on situational factors. For instance, a product might fulfill different needs for different individuals or serve different roles across various contexts, highlighting that value is dynamic and multi-faceted. This subjectivity requires VCD to account for the diverse personal and contextual interpretations of value. Effective VCD achieves a balance among enterprise, customer, and societal values, fostering designs that embody a holistic concept of “good design.”

Current Design Tools for Value

There are a variety of design tools that are used within the field of VSD and VCD to help identify and define stakeholder value. Some examples are: Stakeholder analysis, Scenario building, Stakeholder mapping, Value hierarchy diagrams, Empathy mapping, Value scenarios and Persona creation to name a few (Friedman & Hendry, 2019; Friedman et al., 2017). One of the core tools used in VSD is The Tripartite Methodology (Friedman et al. 2013), a structural approach to help identify and define the most important values within a system. The first step is a conceptual investigation to identify and define the values involved in the process. These values are initially based on theory and the researcher's perception, requiring empirical verification through observations, interviews, or surveys. The second step, the empirical investigation, examines human activity and behavior to confirm the relevance of these values. The third step, the technical investigation, analyzes systems and technologies that embody these values. This step ensures that key values are considered not only from a human perspective but also within the systems people interact with, allowing for the proactive design of supportive technologies.

Another commonly used tool is the "value system" framework. The "value system" refers to a network of interlinked actors and activities, conceptualized as a "value star" when seen in isolation and a "value network" when interconnected with other stars. This network structure allows for both tangible (e.g., products, money) and intangible (e.g., knowledge, relationships, experiences) value exchanges that support co-creation. Unlike the traditional value chain model where value flows in a linear, company-centric way, a value network facilitates multi directional interactions where different actors collaboratively create and perceive value based on their contributions and resources. Within this system, the success of a organization is determined by its ability to transform one form of value into another efficiently, enhancing the value network overall (Randmaa et al., 2012).

Kalbach (kalbach & kahn, 2011; Evenson et al., 2013) identifies system complexity, a lack of coherence in organizational strategy, and challenges in aligning business processes as main issues that Alignment diagrams can help with. Alignment diagrams help with cross-functional communication, facilitating collaboration by providing a shared framework across teams, helping to bridge gaps between departments, such as marketing, operations, and design. Additionally, they help with the identification of value opportunities, revealing new value creation opportunities at touchpoints where business processes meet customer needs, supporting a customer-centered approach to product and service development. Finally, alignment diagrams improve decision-making. By visualizing customer journeys and business responses in one unified view, organizations can prioritize initiatives and allocate resources to areas with the highest potential impact on customer experience. Examples of alignment diagrams are Service Blueprints, Mental Model Diagrams, Journey Maps and the Double Alignment method.

3.2 COMMUNICATION AND CERTAINTY

Closely related to value, we have communication and certainty. Value needs to be communicated clearly in order to achieve alignment among stakeholders. The ways in which this communication occurs, and how ongoing communication builds certainty, is a topic in itself. When organizations initiate a project that requires collaboration with others, they can no longer rely solely on themselves. When there is trust between stakeholders, collaboration can have several benefits such as effective knowledge exchange, an increased resource pool and overall improved project output (Weber & Khademian 2008; Rybníček & Königsgruber 2019). These benefits can be seen as some of the driving factors behind stakeholders' investments and participation in a collaborative project. A force that counteracts this drive is uncertainty and risk. Especially when the drive is economic or monetary in nature, uncertainty decreases investment and participation from stakeholders as the risk that the uncertainty poses can outweigh the potential benefits created by the collaboration. As a result, many opportunities remain unexploited (Johansen et al., 2016). Uncertainty Reduction Theory (URT) reinforces these findings. URT assumes that reducing uncertainty is inherently beneficial for interactions (Sunnaf Frank, 1990). According to this perspective, individuals seek to minimize uncertainty in initial encounters because doing so makes social interactions more predictable and manageable. URT proposes that as uncertainty decreases, communication becomes smoother, relationships develop more positively, and individuals experience greater comfort and affiliation with one another. Uncertainty is experienced when an individual or group is insecure about their state of knowledge and is lacking information that they deem important to know (Brashers, 2001). Uncertainty plays a large role in most organizations and is a constant factor that has to be taken into account. Under conditions of low uncertainty, decision-making tends to be more straightforward, relying on established protocols to achieve predictable outcomes. However, as uncertainty increases, organizations face more complex and ambiguous scenarios, necessitating adjustments in how decisions are made (Conrath, 1967).

Communication Between Stakeholders

The communication of information seems to be closely related to the problem of uncertainty as communication is both affected by uncertainty as well as the tool used to reduce it (Albers, 2012; Brashers, 2001; Driskill & Goldstein, 1986). How communication is handled has a big impact on its measure of success in reducing uncertainty (Figure 4). When there is a fixed amount of information and all of it is important, stakeholders only have to think about task completion. When the information becomes more complex however the stakeholders need to start focusing on information analysis and decision making. When this happens uncertainty increases so in order to reduce this uncertainty, information needs to remain understandable.



Figure 4: The relation between communication, certainty, and participation

According to Albers this can be done by providing proper context so the recipient can understand the information. However, the context that is needed to understand depends on the recipient of the information (Figure 5).

"It's often a problem providing context for the information. The writer must create information which reduces uncertainty within a reader's specific context, when the content itself must apply across many different contexts." (Albers, 2012)

Albers elaborates that for optimal text understandability, the right amount of relevant and purposeful information must be contextually tailored to the reader's needs. Overloading or under-supplying information decreases the effectiveness of communication. To reduce uncertainty the goal should be to optimise the delivery of information, focusing on relevance, purpose and context. SE.Lab already employs this practice, tailoring content to suit different recipients. In one of their projects, a document was published and shared with multiple stakeholder groups. It was labeled "Ambitieplan" for municipal representatives, while for neighborhood participants, it was titled "Avonturenboek." Additionally, certain visuals were adjusted depending on the version, emphasizing the importance of adapting information to the specific audience.

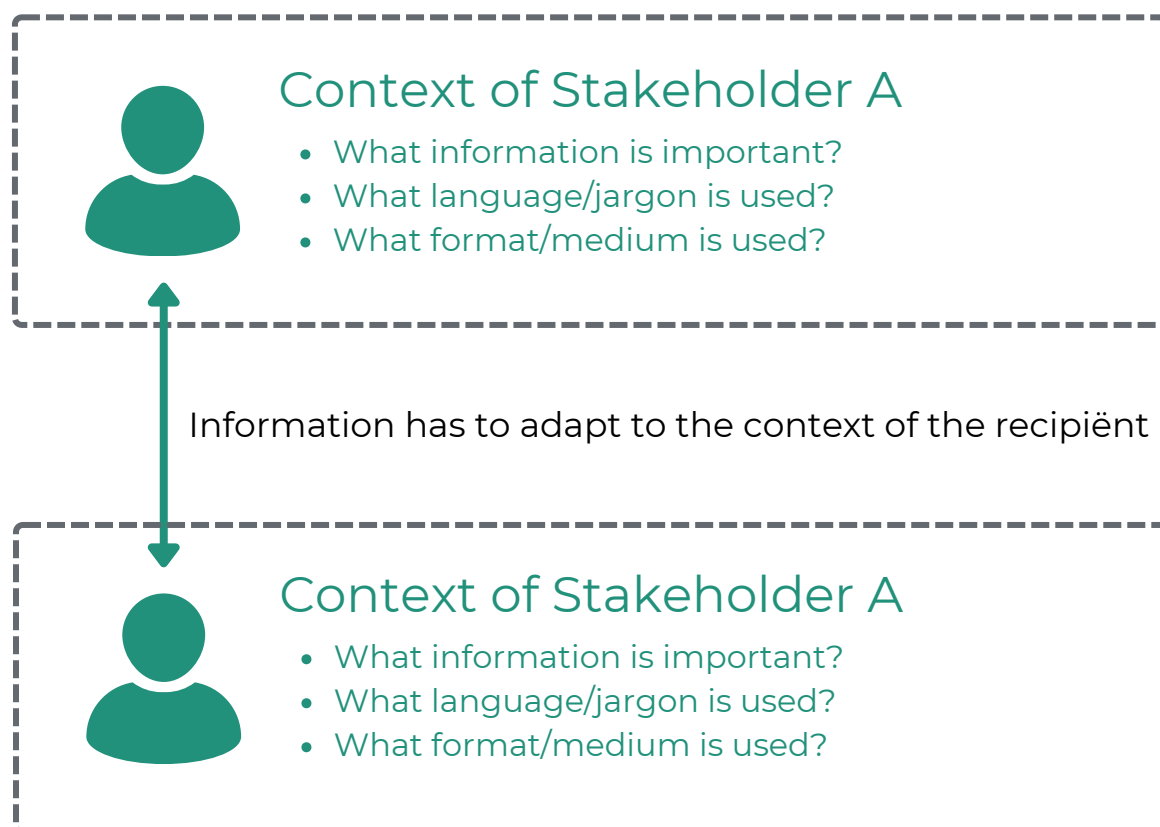


Figure 5: The flow of information between stakeholders with different contexts.

Uncertainty in Design

Design can help individuals, organizations, and societies mitigate uncertainty by creating flexible, adaptive, and forward-thinking solutions. Communication design focuses on the formatting of information to provide optimal understandability. On the flip side Anticipatory design or Agile and Lean design focus on dealing with uncertainty when it inevitably arises.

Anticipatory Design

Anticipatory design is an approach that proactively addresses uncertainty by envisioning and shaping possible futures through systematic foresight, prototyping, and user studies. Instead of merely predicting future trends, anticipatory design actively engages with uncertainty, helping designers create more resilient and adaptable systems. Anticipatory design builds on frameworks like the Futures Cone, which categorizes futures into preferable, probable, plausible, and possible scenarios. The process involves alternating between divergence, exploring multiple future possibilities and convergence, focusing on key, relevant futures for study (Moesgen et al., 2023).

Agile and Lean design

Agile and Lean design are methodologies that focus on flexibility, efficiency, and continuous improvement, helping teams navigate uncertainty. Agile design is built on iterative development cycles, where work is divided into small, manageable increments (sprints). Each sprint delivers a working version of the product, allowing for continuous feedback and adaptation. This iterative approach reduces uncertainty by ensuring that new information, user feedback, and shifting requirements can be integrated throughout the development process. Instead of committing to a fixed, long-term plan that may become obsolete, Agile enables teams to respond quickly to emerging challenges and opportunities. Lean design emphasizes maximizing value while minimizing waste. It encourages teams to eliminate unnecessary processes, focus on delivering customer value, and ensure that all development efforts contribute meaningfully to the final product. Combining these two approaches reduces overall risks and uncertainty (Isomursu et al., 2012).

Communication Design

Communication design plays a crucial role in managing uncertainty by structuring, clarifying, and optimizing the exchange of information. Uncertainty arises when information is incomplete, ambiguous, or difficult to interpret, leading to confusion and inefficiencies. By leveraging various communication theories and models, communication design helps in mitigating these challenges. One of the fundamental ways communication design addresses uncertainty is by establishing clear transmission channels. The Shannon-Weaver Model, for example, emphasizes the importance of encoding and decoding messages effectively while minimizing noise that could distort meaning. By designing structured communication systems, organizations can reduce misinterpretations and ensure that intended messages are received as accurately as possible. Beyond simple transmission, communication design also facilitates meaning-making through interaction. The socio-cultural and socio-psychological perspectives highlight how individuals construct shared understanding through dialogue. By incorporating visual, textual, and multimodal elements, communication design enhances cognitive processing, helping people navigate complex information environments with greater ease. (Costa et al., 2012)

Current Design tools for communication and uncertainty

A tool commonly used for dealing with uncertainty is Design roadmapping (DRM). DRM helps address uncertainty by incorporating a deep understanding of user needs into the roadmapping process. This approach shifts the focus from traditional technology-centered methods to user-centered design, enabling organizations to better manage volatile, uncertain, complex, and ambiguous (VUCA) environments. By anchoring innovation in user value, DRM emphasizes desirability alongside feasibility and viability, allowing organizations to balance user needs with technological and financial constraints, thereby reducing uncertainty about user adoption and market readiness. The systematic future visioning inherent in DRM processes helps firms anticipate and plan for uncertainties in emerging markets and technologies by envisioning long-term user experiences and needs. Furthermore, the iterative and collaborative nature of DRM enables teams to adapt to these uncertainties. In this way, DRM provides tools and processes to navigate uncertainty, grounding decisions in a user-focused, iterative, and systems-oriented framework (Lee et al., 2021; Kim et al., 2020).

The field of communication design has a lot of different approaches to dealing with the complexity of information and how to communicate this to others. One of these approaches is the setup of a content strategy, a process of planning, creating, managing, and delivering content that is both valuable and relevant to the target audience while aligning with business goals. It involves understanding who the content is for and what their needs are, defining the types of content and topics to cover, and selecting the appropriate channels for distribution. The process encompasses the actual creation of meaningful content as well as organizing, updating, and governing it over time, and it also requires analyzing performance to optimize engagement. By implementing a strong content strategy, organizations can enhance user experience, improve brand credibility, and more effectively achieve their objectives. (Getto et al., 2022). Specific tools that can be used are for example the 5W1H Framework, the Content Lifecycle Framework or the Pillar-Cluster Model.

3.3 POWER STRUCTURES

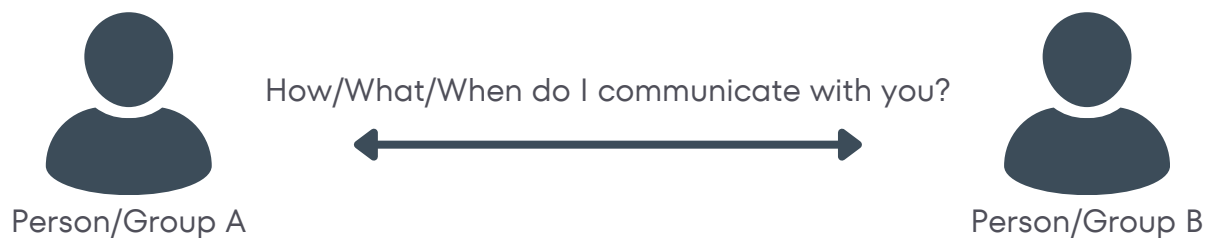
When there is collaboration, there are power structures. Companies and organizations use different power structures both internally and when collaborating with other stakeholders. Stakeholders in collaborations rely on each other for key resources such as funding, information, labor or regulatory approval. The greater the dependence on a particular stakeholder, the stronger their influence becomes. For example, governments establish authority over businesses through regulatory frameworks and enforcement, ensuring compliance with legal requirements. Likewise, consumers shape corporate strategies through their purchasing behavior, compelling companies to align with market preferences and societal expectations. Besides resources, positioning inside collaborative networks also impacts power dynamics. Network centrality, a metric for how connected a stakeholder is within a network, is a key factor in how much influence they can exert over other stakeholders, shaping processes and decision making (Figure 6). Consumer advocacy groups for example can increase their power by forming alliances with media, regulatory bodies, or high-profile customers (Neville & Menguc, 2006). These power dynamics can vary in two dimensions: power level, the average level of expertise or authority within a collaborative team, and power hierarchy, the degree of power disparity within the team. The effectiveness of a team's power structure depends not just on whether power is high or low, but also on how it is distributed. The interaction between team power level and team power hierarchy has significant implications for collaboration and performance. In low-power teams, hierarchy can be beneficial, as it provides structure and guidance for members who may lack experience or authority. However, in high-power teams, rigid hierarchies can be detrimental, as they increase competition and restrict the free flow of ideas. Studies have shown that when high-power teams adopt a flat structure, they often achieve better results, particularly in scientific collaboration, where shared expertise and open communication drive innovation (Greer, 2014; Xu et al., 2021).

The impact of Hierarchy

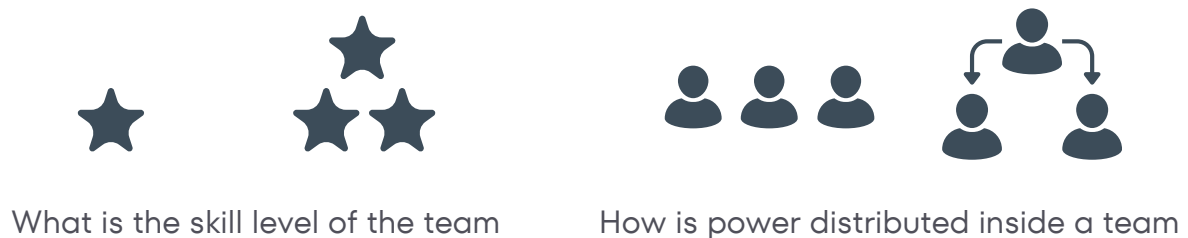
As mentioned before, having formal and informal hierarchy in project teams is common practice in and between companies and organizations, and can be beneficial for certain aspects of development such as decision making (Lahiri et al., 2019; Anderson & Brown, 2010). However, The idea that hierarchy is necessary and that a stronger hierarchy will provide better collaboration is contested. Research has shown that stronger hierarchical structures in collaborative endeavors negatively impact team member attitude (Anderson & Brown, 2010). Research by Becker and Baloff (1969) showed that group members of teams with strong hierarchies were overall more frustrated during the project than members of teams with more egalitarian structures. Whilst research results on group performance are mixed, the effects of hierarchy on group attitude and collaboration are clear; groups and organizations with stronger hierarchies tend to have members who are less satisfied, less motivated, and more inclined to leave the group (Anderson & Brown, 2010). More hierarchical power structures also increase the chance of biased and patriarchal behaviour in collaborative systems (Calderon and Westin, 2019). Additionally, research has shown that steeper hierarchies, characterized by multiple layers of authority, tend to impede the free flow of information within teams (Anderson & Brown, 2010). Power dynamics within strong hierarchical environments exacerbate these communication barriers.

Individuals in lower-ranking positions often feel constrained by social and material threats, reducing the likelihood of these individuals speaking up about critical issues or sharing innovative ideas. Studies have consistently shown that lower-ranking employees are less likely to voice concerns or provide feedback (Kish-Gephart et al., 2009; Milliken et al., 2003). Hierarchies also diminish cooperation among team members by fostering competition. The higher stakes associated with rank encourage individuals to prioritize personal advancement over group success. This competitive environment can lead to mistrust, reduced communication, and even sabotage, undermining the collective performance of the group. Research in negotiation and organizational behavior highlights how power disparities hinder the ability to achieve integrative, mutually beneficial outcomes. Rather than fostering collaboration, hierarchies often push individuals toward coercive and competitive tactics, further eroding group cohesion (Anderson & Brown, 2010; Berger et al., 1980).

PEOPLE



TEAMS



SYSTEMS



Figure 6: Power structures in different scales of collaboration

Power Structures in Design

The field of Design helps reduce unproductive power structures by promoting collaboration, critical thinking, and user-centered approaches that challenge traditional top-down decision-making structures. Rooted in problem-solving and creativity, Design thinking emphasizes co-creation, participatory methods, and interdisciplinary teamwork, encouraging researchers to work alongside users, stakeholders, and peers rather than adhering to rigid authority structures. There are various design methodologies that actively try to remove hierarchy such as Participatory design

Participatory Design

Participatory Design (PD) is a collaborative approach that actively involves users and stakeholders in the design process, ensuring that those who will ultimately use a system or product have a direct hand in shaping its development. Originating from Scandinavian traditions, PD is deeply rooted in democratic principles, workplace empowerment, and shared decision-making. Unlike traditional top-down design methods that rely on experts and management to dictate solutions, PD fosters co-creation by reducing hierarchy and emphasizing inclusivity. One of the primary ways Participatory Design reduces hierarchy is through the democratization of design. It ensures that all voices, including frontline workers, users, and marginalized groups, have an equal influence in the decision-making process. Scandinavian participatory design, in particular, has historical ties to labor movements and industrial democracy, giving workers a voice in workplace technologies that directly impact their roles. By shifting from expert-driven solutions to a more collaborative process, PD challenges traditional hierarchies by treating users as co-designers. This means that instead of designers and engineers making unilateral decisions, end-users contribute their lived experiences and domain knowledge, ensuring that the final product aligns more closely with real-world needs (Gregory, 2003).

Agile Workflow

Agile is widely adopted across industries, making it easier to integrate ethical and participatory decision-making into an already familiar work process. Agile methods emphasize flat hierarchies, allowing developers and team members to work independently and make decisions without rigid managerial control. This promotes autonomous problem-solving, reducing dependency on top-down leadership. Additionally, Agile practices foster collaborative teamwork, ensuring that knowledge is shared rather than siloed. This structure allows ethical and technical discussions to emerge organically, rather than being dictated from upper management. The iterative and object-focused nature of Agile also ensures that teams continuously refine their work through practical feedback, making decision-making a shared and evolving process rather than a directive from leadership. The time-boxed nature of Agile sprints provides clear but flexible endpoints for decision-making, preventing excessive reliance on management for final approvals and allowing teams to make timely, autonomous decisions (Zuber et al., 2021).

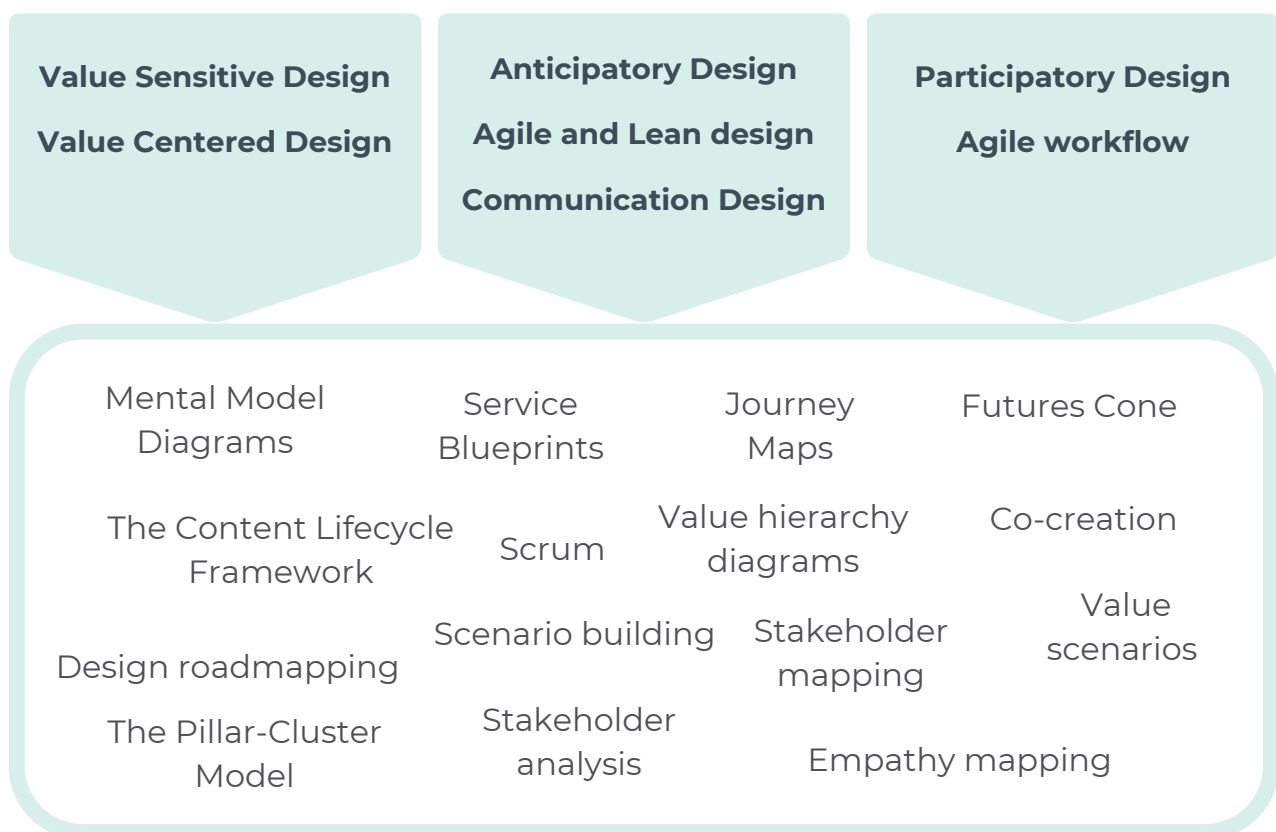
Current Design Tools for Power Structures

A core framework used in Participatory Design is Co-creation. Co-creation is a collaborative approach where organizations, designers, and end-users work together to develop products, services, or solutions. Unlike traditional top-down design, co-creation actively involves stakeholders throughout the process, ensuring their needs, insights, and experiences shape the outcome. This participatory method fosters innovation, enhances user satisfaction, and creates more meaningful, effective solutions. Common in UX design, service design, and product development, co-creation often takes place through workshops, brainstorming sessions, and prototype testing, promoting inclusivity and shared ownership of the final result. (Broekema et al., 2023)

One of the most famous agile frameworks is Scrum. Scrum is an Agile project management framework designed to help teams deliver complex products iteratively and incrementally. It emphasizes collaboration, adaptability, and continuous improvement. The framework consists of three key roles: the Product Owner, who defines the vision, prioritizes work, and represents stakeholders; the Scrum Master, who facilitates the process, removes obstacles, and ensures adherence to Scrum principles; and the Development Team, a cross-functional group responsible for delivering working product increments. Scrum operates through structured workflows centered around sprints, which are time-boxed iterations (typically 1–4 weeks) where teams commit to completing a set of prioritized tasks. One of the key strengths of Scrum is that it helps reduce hierarchy in collaboration. By emphasizing self-organizing teams and collective responsibility, Scrum shifts decision-making from top-down management to the team itself, enabling developers, designers, and other contributors to take ownership of their work. (Schwaber & Beedle, 2002)

◀ TAKEAWAYS

- Through literature research and expert interviews, three collaborative themes have been identified: value alignment, communication and certainty, and power structures.
- When stakeholders collaborate, they not only pursue a shared objective, but also bring their own internal values and desired outcomes to the table. Openly communicating these individual perspectives and aligning them is essential for a successful collaboration.
- Communication and certainty are closely connected. Stakeholders need to consistently share information to build trust and strengthen their commitment to the collaboration. The quality and frequency of communication has a significant impact on the effectiveness of the partnership. Reduced communication increases uncertainty, which in turn lowers stakeholder engagement and commitment.
- Collaboration is shaped by underlying power structures that influence how responsibilities, decisions, and information are distributed. Not all power dynamics are equally effective, so identifying and establishing an optimal structure is key to enabling productive and balanced stakeholder collaboration.
- Current design methodologies provide a large variety of tools and frameworks to help overcome the problems identified in the three collaborative themes, as seen below.



4. COMPLEX SYSTEMS AND PROBLEMS

The design methodologies outlined in the previous chapter provide an array of tools and frameworks that can be employed to deal with the problems presented in the three collaborative themes. However, the usefulness of these tools has its limits. As complexity increases, current design approaches become obsolete. This chapter examines how complex systems and problems make collaboration difficult and reduce the usefulness of current design theory. The specific problems introduced by complexity are addressed and the opportunities that large language models provide are introduced.

4.1. CHALLENGES OF COMPLEX SYSTEMS AND PROBLEMS

Complex systems and problems are a significant challenge for design. Complex systems are systems that are composed of many interconnected parts, and complex problems are problems that are difficult to solve. Both complex systems and complex problems are characterized by a high degree of uncertainty and a lack of clear boundaries. This makes it difficult to understand how the system or problem works, and it is often difficult to predict the outcomes of different actions. In this section, we will explore the challenges of complex systems and problems, and we will discuss some of the ways that design can be used to address these challenges.

KEY TERMINOLOGY

Complex systems: Complex systems or multi-stakeholder systems are collaborative systems between different stakeholders from different domains. Complex systems are networks of interconnected stakeholders that interact in dynamic and often unpredictable ways. In these systems the behavior of the whole system cannot be easily understood simply by analyzing an individual component.

Complex problems: Complex problems are challenges that involve many interconnected factors, where cause and effect is often unclear, and solutions may lead to unintended consequences. They typically require interdisciplinary thinking, ongoing adaptation, and collaboration to effectively address problems, as they can't be solved with straightforward or one-size-fits-all approaches. Complex problems often lead to the creation of complex systems.

LLM: Large Language Models (LLMs) are advanced AI systems trained on vast amounts of text data to understand and generate human-like language. They work by predicting the next most likely word in a sequence, enabling them to perform tasks like answering questions, writing text, translating languages, and many more.

ChatGPT: ChatGPT is an AI-powered conversational agent developed by OpenAI, based on large language models like GPT-4.

4.1 INCREASING COMPLEXITY

Through research, three main themes have been identified that play a large role in successful collaboration between stakeholders. The alignment of value, communication and certainty, and power structures between stakeholders in a project. Current design theory provides a large toolbox full of frameworks and tools that can be used in order to deal with the problems associated with these three themes. One might assume then that we are well equipped to tackle any problem that comes our way but, as was made clear in the introduction, our society is still struggling with issues for which there seem to be no solution. Modern complex problems introduce high levels of conflict, uncertainty, and complexity. Their unstructured nature means that stakeholders often disagree on the definition of the problem and the best solutions, leading to tensions and fragmentation in collaborative efforts. Additionally, these problems are cross-cutting, spanning multiple organizations and disciplines, making it difficult to align interests and coordinate responses effectively. Because complex problems are relentless and continuously evolving, collaboration must also be adaptive, requiring stakeholders to adjust their approaches as new challenges emerge (Khademian, 2008). These insights have been reaffirmed by SE.Lab. Drawing from their extensive experience collaborating with a diverse range of stakeholders, they observed that many current methods fall short in addressing the complexity of today's societal challenges. A common shortcoming is the lack of iterative processes, an essential component for navigating the dynamic and evolving nature of complex problems. To remain effective in real-world contexts where conditions shift rapidly, collaborative methods must be both scalable and adaptable. SE.Lab highlights the importance of context sensitivity, systemic thinking, and multidisciplinary collaboration as critical factors for successfully transitioning toward more resilient and effective collaborative structures.

Value Alignment and Complexity

Design theory such as VCD and VSD provide frameworks and tools on how to identify and align value between stakeholders. These frameworks work for normal problems in regular systems but as complexity increases they fail. As systems become more complex, stakeholders that are far removed from each other in background and expertise begin to collaborate and as a result value systems that have very little overlap need to be aligned (Ooi & Husted, 2021). The diverse backgrounds of stakeholders significantly affect their ability to communicate about value. While it's still manageable to mediate between two or three stakeholders, as the number grows, aligning everyone's understanding of what "value" means, and identifying which values matter to whom, becomes an increasingly difficult, if not impossible, task. Doing this might simply require too much effort for involved parties to be a valid action during collaboration. As is common with most design methodologies, a large amount of time and energy is needed for tools and frameworks to be properly executed. Companies and organizations might simply not have the time or manpower to do this. Letaifa (2014) has outlined a common cycle in companies that attempt to perform value co-creation in multi-stakeholder systems. They found that whilst companies and organizations are often initially willing to invest in more collaborative value creation infrastructure, they lack the follow through to keep this infrastructure alive.

This pattern identified by Letaifa highlights companies and organizations' failure to invest in long-term cooperative structures. Instead of viewing this cooperative infrastructure as a way to share and exchange value, which could benefit everyone in the long run, it is seen as an opportunity to extract value without giving back, which might benefit companies and organizations upfront but destroys any future prospects of collaborative value creation. Design approaches like the Tripartite Methodology from VSD or the value system framework from VCD can help with value identification in normal systems without too much complexity. However, when the total number of stakeholders increases, design tools become difficult to fully execute. The Tripartite Methodology, for example, requires in depth investigation (conceptual, empirical and technical) of each stakeholders values. Digital automation can support this process by replacing in-person interviews with well-designed online questionnaires and forms. However, challenges remain, particularly in analyzing the collected data. This also assumes that stakeholders have the time and willingness to participate in such preparatory research in the first place. Furthermore, alignment tools such as service blueprints or journey maps require a lot of time and energy to set up. In complex systems, this process can require so much additional time that it may outweigh the value it provides. Moreover, because complex systems are constantly evolving, stakeholder values can shift rapidly, making tools like service blueprints and journey maps potentially obsolete within weeks.

Communication, Certainty and Complexity

Similar to value, as complexity increases, effective communication becomes more challenging. As a result, traditional design methods for managing the uncertainty caused by poor communication become less effective. When complexity increases, so does the amount of information that needs to be communicated, increasing the chance of miscommunication (Fiset et al., 2023; Tenzer et al., 2013). With increased complexity comes a more mixed bag of expertise on various topics, so, when information is shared different parts will be valuable for different stakeholders whilst the rest of the data will be experienced as "noise" (Albers, 2012). Experts reading texts intended for novices often have to sift through large amounts of information they already know in order to find the relevant or new insights. Conversely, the novice will have to go through a lot of information that means nothing to them, due to their understanding of the subject, in order to get to the information that is valuable for them. A second source of noise comes from "cross talk". Cross talk happens when multiple stakeholders with different levels of comprehension and backgrounds (e.g. financial, legal, safety) come together to discuss a topic. These discussions add a lot of additional and often skewed information besides the primary message, making it difficult to identify what is important. As the volume of available information grows, individuals increasingly disregard data that contradicts their pre-existing decisions and opinions. Rather than objectively assessing all relevant details, people tend to seek out information that confirms their initial conclusions while ignoring or overlooking conflicting evidence. This selective attention leads to a bias in information gathering, where contradictory facts are neither actively sought nor given proper consideration. Another bias in communication that arises with complexity is the false consensus bias. When there are too many different viewpoints people overestimate how much other people agree with their point of view because they are no longer able to check with everyone and an overwhelming amount of contradictory viewpoints leads individuals to stop verifying and start assuming (Curşeu & Schruijer, 2017).

Design approaches that focus on communication such as the Shannon-Weaver model already highlight the importance of “encoding”, “decoding” and noise in conversations but don't offer any structural tools to help deal with the amount and variety of information present in complex systems. Frameworks like 5W1H, which help tailor information to specific target groups through actions such as “defining audience personas” or “determining the type of content needed,” can be effective when addressing one or two groups. However, as the number of target audiences increases, the effort required to customize content for all of them becomes overwhelming. When uncertainty in systems inevitably increases and causes problems in collaboration, design approaches like agile and lean could be used to combat uncertainty. Unfortunately, these approaches also encounter challenges when applied to complex systems, as they rely on rapid iterative cycles, which, as previously mentioned, are difficult to execute when numerous and diverse stakeholders are involved. Another way to reduce uncertainty is by making well-reasoned predictions about the future. However, as the amount of data required for accurate forecasting grows, so does the time and effort needed, once again diminishing the tool's effectiveness in complex systems.

Power Structures and Complexity

As the number of participants increase and their differences become more pronounced due to varying backgrounds, so does the complexity of the power structures present in their collaboration. For a long time, hierarchy has been used to control and create order in collaboration (Corominas-Murtra et al., 2013) and whilst it might seem logical that rigid structures would help mitigate complexity, the opposite seems to be true. Weber and Khademian (2008) highlight the challenges of traditional hierarchical systems in addressing complex problems, emphasizing their inherent limitations. Hierarchies, which are commonly used in public policy development and problem-solving, are inadequate for tackling the dynamic, unstructured, and interconnected nature of complex problems. These problems often span multiple policy domains, organizations, and levels of government, making rigid hierarchical approaches ineffective. Hierarchical structures tend to isolate problems into silos, which clashes with the multifaceted reality of complex issues. Kelley (1951) states that the negative effects of hierarchy are especially pronounced when teams must work interdependently on complex tasks. In such scenarios, the need for open communication and seamless coordination becomes crucial, and hierarchical barriers can significantly hinder these processes. While hierarchies may prove beneficial for simpler tasks or scenarios requiring quick, unilateral decision-making, they tend to be counterproductive in environments where collaboration and mutual support are critical to success. Traditional design theory attempts to flatten hierarchy through methods like participatory design but as complexity increases they tend to fail or aren't even attempted to begin with (Shapiro, 2005; Kpamma et al., 2017). Design methods focusing on power structures have a variety of approaches to ensure proper collaboration. Methods like participatory design and Agile workflow opt for flat power structures as hierarchy often hinders good collaboration. However, as complexity increases, the number of power dynamics within the system also grows, making it difficult to maintain a fully flat structure. When group size increases people tend to prioritize relations with people they align with over ones with stakeholders that might oppose their points of view, creating camps in collaborative systems. This reduces the effectiveness of frameworks such as co-creation. Additionally, as system complexity increases, it becomes easier for powerful stakeholders to dominate central narratives. Smaller stakeholders, in particular, may struggle to maintain a clear overview of the entire system, making it harder for them to contribute to those narratives. (Curşeu & Schruijer, 2017).

4.2 THE CHALLENGES OF COMPLEXITY

Aligning values between a few stakeholders is feasible, but as more actors join, conflicting priorities and perspectives make alignment increasingly difficult. Additionally, more complex systems generate larger amounts of information, leading to communication overload and a higher risk of misinterpretation. Finally, more stakeholders with diverse backgrounds create bigger and more complex power structures, making it harder to establish effective collaborations and decision-making processes. Collaboration in complex systems brings together stakeholders with divergent value systems that may have little in common, making mutual understanding difficult. Stakeholders have different levels of expertise, causing experts to sift through information they already know whilst novices struggle with technical details beyond their understanding. When stakeholders from different fields (finance, legal, safety) collaborate, conversations can start to get filled with irrelevant or misinterpreted information, making it difficult to identify what truly matters. Companies might not currently have the infrastructure to deal with complex problems. They tend to rely on traditional hierarchies which, while intended to create order, often fail to manage complex, dynamic problems. Hierarchies silo problems, preventing cross-functional collaboration and rigid structures slow down decision-making and stifle innovation in complex, interdependent tasks. Identifying and aligning values demands significant time, effort, and resources, which many organizations lack. It is these different aspects of complexity that create a necessity for new and innovative approaches if we are to overcome today's challenges (Figure 7).

Transitioning to a New Approach

As seen in cases like Meerwijk, modern complex challenges operate on an immense scale, involving multiple interconnected layers, including local inhabitants and their social networks, social enterprises, commercial businesses, and infrastructure such as schools and hospitals. Addressing these challenges requires a shift toward new collaborative models. Technological advancements have transformed the way we collaborate, enabling global knowledge exchange and remote teamwork across disciplines and institutions. Digital communication tools such as videoconferencing, collaborative online platforms, and real-time data-sharing systems have enhanced access to expertise, reduced geographical barriers, and increased the efficiency of scientific collaboration (Hara et al., 2003; Lin et al., 2022). One can argue however that thus far technology has mostly allowed humans to expand their networks and increase the number of stakeholders that they can reach, but are still lacking in supporting the resulting collaborative structures that are needed between the large number of stakeholders to sustain these networks and produce value. While technical systems provide valuable data, they cannot replace the human ability to interpret, negotiate, and build relationships. The social and relational aspect of communication ensures that people work together effectively, trust each other, and adapt to unpredictable situations, making it a fundamental part of managing collaboration in complex environments (Johansson & Persson, 2008). This is where the emergence of large language models presents a unique opportunity. Trained on vast amounts of human-generated text, these models inherently reflect human thought, language, and interaction patterns. LLMs have the potential to bridge the gap between technological innovation and the human complexity of cases like Meerwijk, serving as a mediating tool that enhances communication, supports decision-making, and facilitates meaningful collaboration across diverse stakeholders.

COMPLEXITY

More and further removed stakeholders

As collaborative systems grow in scale and complexity, stakeholders from increasingly diverse and distant backgrounds must collaborate. Aligning their differing values becomes exponentially harder, often demanding more effort and resources than organizations can realistically invest.

More information and communication

With more stakeholders comes more information, increasing the risk of miscommunication and noise. Differences in expertise and priorities create fragmented understanding, making it challenging to share relevant knowledge.

Power structures become more pronounced

When systems increase, more organizational structures and power structures are introduced to try and manage the system but hierarchies often become less effective, hindering communication and innovation.

More variables enter the system

Complex systems bring an influx of interconnected variables that evolve over time, making static tools and frameworks quickly obsolete. Predicting outcomes or identifying clear solutions becomes a moving target, demanding constant adaptation.

High level of adaptability becomes required

To keep pace with ever-changing conditions and stakeholder dynamics, collaboration must be flexible and iterative. Rigid methods fail under this pressure, emphasizing the need for adaptive strategies and resilient design thinking.

Figure 7: The different aspects of complexity in collaboration

4.3 USING LARGE LANGUAGE MODELS

Large Language Models (LLMs), such as ChatGPT, are beginning to play a transformative role in almost all domains of society. As the name suggests, LLMs are advanced models trained to read, interpret, transform, and generate language. When applied to text-based datasets, LLMs excel in a wide range of tasks, including question-answering, text generation, language translation, text classification, summarization, virtual assistance and information extraction to name a few (Hadi et al., 2023). Integrating these functionalities, either individually or in combination, unlocks immense potential for innovation across industries.

For instance, LLMs like ChatGPT have demonstrated remarkable promise in healthcare. In medical education, ChatGPT serves as an interactive tool for learning and problem-solving. Notably, its performance on the United States Medical Licensing Exam (USMLE) either met or exceeded the passing threshold, achieving this without any specialized training or reinforcement. Furthermore, its explanations exhibit a high degree of accuracy and insightful understanding (Hadi et al., 2023). LLM's are also being integrated into the field of design. Research has shown that AI can perform useful tasks along the entire design chain varying from labeling information according to importance or priority and generating content based on human created original work in the early phases, to reorganizing component libraries and design systems and documenting design assets and materials for handover to development teams in the later design stages (Li et al., 2024).

Furthermore, AI models are now outperforming humans in things such as Theory of Mind (ToM). ToM refers to the cognitive ability to attribute mental states such as beliefs, desires, intentions, and emotions to oneself and others and to understand that others may have mental states different from one's own. This capability enables humans to predict and interpret the behavior of others based on their mental states. It is a fundamental aspect of human social cognition, underpinning communication, empathy, and cooperative behavior (Street et al., 2024). In the context of LLM this means that they are becoming increasingly adept at understanding the deeper layers of language and conversation. These advancements have also demonstrated that LLMs can surpass humans in debating. Humans are no longer able to distinguish talking to a Large Language Model like ChatGPT from talking to a human. Furthermore, LLM are better at persuading us and making us change our minds than humans are (Jones & Bergen, 2024, Salvi et al., 2024).

LLM are currently already being used for a wide variety of tasks focused on collaboration. Westermann et al. (2023) developed a mediation tool built on the framework of ChatGPT 4.0 called LLMediator. LLMediator offers a variety of functions designed to enhance online dispute resolution (ODR) processes. One of its key capabilities is the reformulation of inflammatory messages, aimed at reducing emotional escalation during negotiations. By identifying emotionally charged or confrontational language in user messages, the system intervenes to promote constructive communication. It detects such language through methods like keyword searches, sentiment analysis, or user-triggered prompts, and suggests rephrased versions that maintain the original content but neutralize emotional overtones.

This functionality empowers users to express their concerns effectively without escalating tensions. Importantly, users retain control over their communication, as they can choose to send the original message, accept the AI's suggestion, or edit the reformulated version. The system generates messages based on recent exchanges in the conversation and allows mediators to provide specific instructions to tailor the drafts. For instance, a mediator may instruct the AI to emphasize key facts, suggest compromises, or clarify contentious points.

Shaikh et al. (2023) designed a tool named Rehearsal, an interactive system designed to teach conflict resolution through simulated roleplay. Using a technique called Interests-Rights-Power (IRP) prompting, Rehearsal leverages large language models to simulate conflict scenarios and provide real-time feedback, helping users practice effective conflict resolution strategies. The system generates realistic conflict dialogues by grounding simulations in the IRP framework, which emphasizes cooperative strategies over competitive or rights-based approaches. Users can interact with simulated interlocutors, explore alternative conversational strategies, and receive targeted feedback. A study involving 40 participants showed that training with Rehearsal significantly improved participants' ability to use cooperative strategies in real-life conflicts, doubling their application and reducing competitive strategies by 67%.

Abdelnabi et al. (2023) build a benchmark to test various LLM in complex collaborative environments. The benchmark itself comprises a dynamic set of negotiation games designed to simulate real-world scenarios, such as resource allocation and balancing competing interests. Each game involves agents with distinct goals, acceptance thresholds, and scoring systems, requiring them to engage in strategic deliberation and trade-offs. By parameterizing factors such as the number of agents, issue complexity, and feasible solution sets, the games provide a robust and evolving testbed for LLM evaluation. By using a game structure the researchers were able to test the LLM's on the use of advanced social skills such as cooperation, strategic planning, competition, balancing multiple objectives at the same time, and being aware of cooperation barriers such as manipulation and deception.

Using LLM to Deal With Complex Problems

Large Language Models (LLMs), such as ChatGPT, offer significant potential to address challenges in stakeholder alignment, communication, uncertainty reduction, and collaboration barriers created by power structures (Figure 8). In aligning values among stakeholders, LLMs can analyze diverse perspectives, identify commonalities, and suggest language or frameworks that promote shared understanding. By synthesizing data from various sources and offering balanced recommendations, LLMs help stakeholders navigate competing priorities and establish mutual goals. Their ability to model nuanced communication fosters inclusivity and ensures that all viewpoints are considered, creating a foundation for collaborative decision-making.

Large Language Models (LLMs) excel at analyzing and transforming vast amounts of text and data, making them highly effective in enhancing communication between stakeholders with differing expertise. When given sufficient information, an LLM can adapt text to suit the reader's level of understanding, tailoring the context to the recipient rather than the sender. To reduce uncertainty, LLMs provide advanced predictive capabilities and scenario modeling. By leveraging large datasets, they can generate insights that clarify potential risks, project outcomes, or alternative strategies. This ability to deliver data-driven forecasts enables stakeholders to approach decision-making with greater confidence. Additionally, LLMs enhance transparency by addressing ambiguities, summarizing complex information, and responding to questions in real-time, ensuring all parties are well-informed.

LLMs also play a critical role in removing barriers created by unbalanced power structures in collaboration. Tools powered by LLMs can mediate communication by reformulating messages to remove emotional or confrontational tones, encouraging constructive dialogue regardless of an individual's position in an organization. They democratize access to expertise by offering consistent, unbiased guidance, making it easier for team members at all levels to contribute meaningfully. Furthermore, LLMs facilitate training and conflict resolution through simulations that teach effective communication strategies, empowering individuals to navigate power dynamics and work collaboratively. By fostering open and respectful interaction, LLMs help create environments where hierarchy is less of a barrier to innovation and collective progress. In these ways, LLMs serve as powerful tools for enhancing collaboration, reducing friction, and ensuring that all stakeholders work together effectively toward shared objectives.

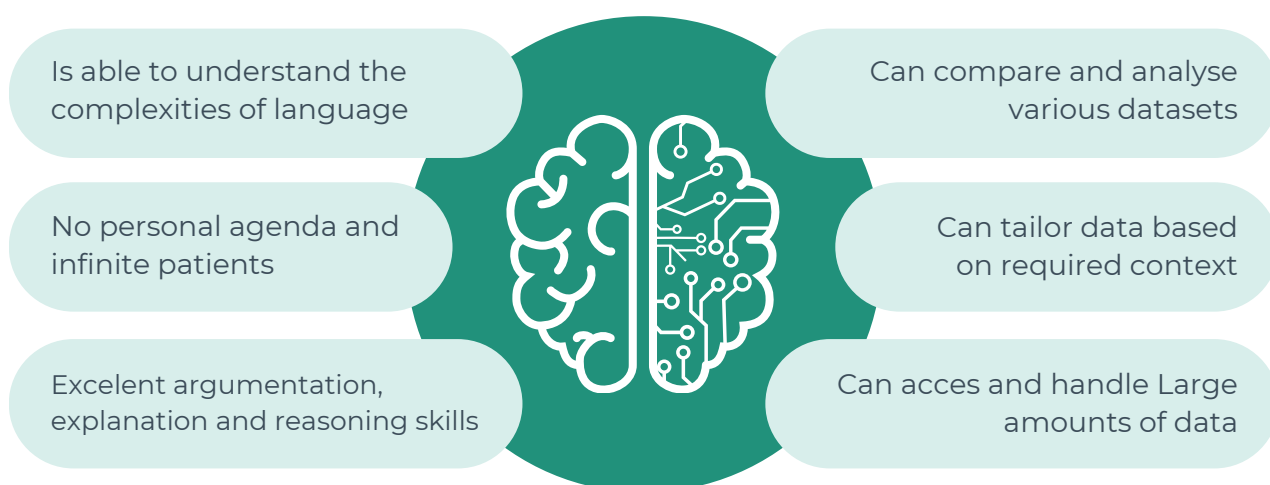


Figure 8: The unique combination of strengths form LLMs

5. PROBLEM DEFINITION

The chapter marks the end of the first diamond from the double diamond methodology. With all the information gathered thus far, a clear problem can be defined, leading to the formulation of a research question that will guide the next phase of the study.

When addressing complex problems within intricate systems, effective collaboration between stakeholders is essential. Insights from stakeholder engagement and literature research highlight three critical themes that contribute to successful and productive collaboration among stakeholders: value alignment, communication and certainty, and power structures. All these factors share a common characteristic: they are inherently social in nature, making language a critical element in their dynamics and influence. As Noyes (2021) puts it:

“Conversation is the actual interactive process through which organizing occurs; through conversation, organizational members construct and negotiate (or co-orient around) texts, which form a kind of conceptual scaffolding made up of words, phrases, turns of speech, metaphors, anecdotes, all of which are there because of the distillation, stored in language in the memory of participants, of their personal and collective history of previous interactions.”

Current design theory that is used to support collaboration in multi-stakeholder systems is unable to keep up with increased complexity of modern systems and problems. Technology so far has helped us increase our social networks but has done little to support us in managing the collaborative efforts needed to maintain these systems and produce value. This problem persists because collaboration has an inherently human element to it, requiring an understanding of the complexity and nuances of the language used to communicate during these collaborations.

On May 13, 2024, OpenAI introduced their newest version of ChatGPT, a Large Language Model (LLM). Language is the foundation of LLMs, as they are designed to understand, generate, and interpret human communication. These models analyze the structure, context, and meaning of language to perform tasks such as translation, summarization, and conversational interaction. By leveraging vast datasets, LLMs can learn patterns, nuances, and cultural contexts within language, enabling them to provide coherent and contextually relevant responses. LLM may play a pivotal role in overcoming barriers created by language and misunderstanding during collaboration. They can generate, enhance, and transform large amounts of text and data in minimal time. LLMs bring value by clarifying ambiguous statements, rephrasing and offering alternative interpretations which ensures that all parties have a shared understanding. Additionally, they can summarize lengthy discussions, highlight key points, and provide consistent terminology, reducing the chance of misalignment. By facilitating clearer and more inclusive communication, LLMs could help teams navigate complex interactions, fostering collaboration and accelerating progress (Havlík, 2023). The next phase of this research will focus on the following research questions:



“How can collaboration between stakeholders in complex systems working on complex problems be improved by addressing current challenges related to value alignment, communication & uncertainty, and power structures?”

>How can stakeholder identify, discuss and align value for themselves and other stakeholders in complex systems with LLM

>How can communication between stakeholders in complex systems be improved to decrease stakeholder uncertainty and increase participation with LLM

>How can barriers created by unproductive power structure and hierarchy in complex multi-stakeholder projects be broken down as much as possible with LLM

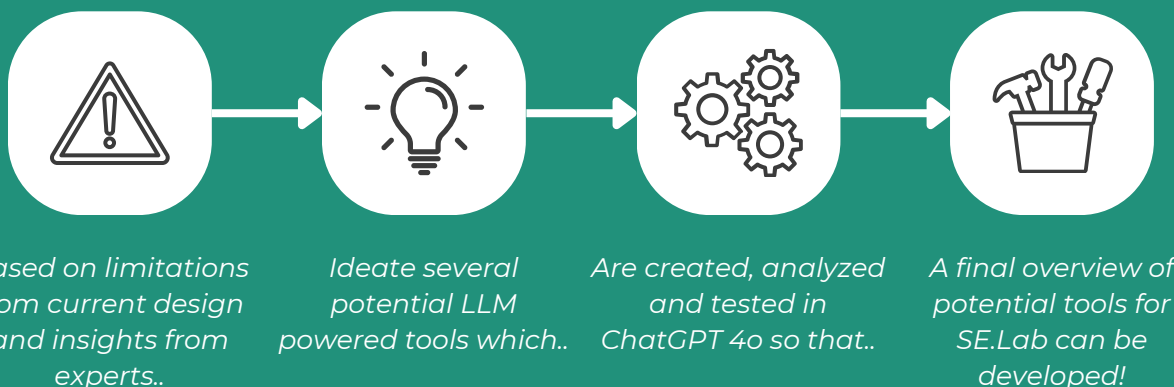
◀ 6. METHODOLOGY

In the previous chapters, collaboration and the impact of complexity on collaboration have been thoroughly researched and discussed. Three central themes in collaboration were identified, and the ways in which the traditional design theory toolbox attempts to overcome the challenges associated with collaboration were outlined. By addressing the ways that traditional design theory struggles to keep up as complexity increases, it is clear that new methods are needed to address the problems arising from the increased complexity of our world.

The next few chapters will explore the potential of using Large Language Models to enhance the design toolbox and tackle complex problems and systems. First, the ethics and risks of innovations involving LLMs will be examined, providing a comprehensive overview of the potential harms that could arise if these technologies are handled irresponsibly. Following this, several established ethical and safety guidelines will be analyzed and synthesized into a new set of ethics and risk focused design principles to guide and evaluate the tools developed during this research.

To further understand how LLM powered tools can be created and assessed, the principles of prompt engineering, LLM evaluation, and benchmarking will be discussed.

Building on the problems identified in earlier chapters, the insights from interviewed experts, and the collaborative framework provided by SE.lab, various potential LLM tool ideas will be proposed. These tools will be evaluated based on their ability to improve collaboration and checked for how they might cause harm. Through iterative prompt design, testing, and refinement, the tools will be created and optimized within the ChatGPT-4o model. This process will result in an overview of the feasibility, effectiveness, and limitations of LLM-powered tools for collaboration, culminating in a final enhanced design toolbox that SE.lab can test and apply in real-world cases like in Meerwijk.



7. LARGE LANGUAGE MODELS

To understand how large language models can enhance collaboration, it is essential to explore the principles of designing with them. However, as with any form of innovation, ethical implications, potential risks, and thoughtful design must be considered from the outset. Therefore, this chapter begins by reviewing existing literature on the ethical risks associated with AI and LLM innovation. It then introduces the practice of prompt engineering, focusing on strategies for optimizing interactions with LLMs. Finally, the chapter discusses methods for testing, evaluating, and benchmarking these models to ensure their effectiveness and reliability.

CHAPTER OBJECTIVES

- 1. Understand the ethical risks associated with AI and LLM innovation.
- 2. Explore the principles of designing with large language models.
- 3. Review existing literature on the ethical risks associated with AI and LLM innovation.
- 4. Introduce the practice of prompt engineering.
- 5. Focus on strategies for optimizing interactions with LLMs.
- 6. Discuss methods for testing, evaluating, and benchmarking these models.
- 7. Ensure the effectiveness and reliability of these models.

KEY TERMINOLOGY

Innovation: Innovation refers to the process of creating new ideas, products, or methods that bring about meaningful improvements or solve existing problems in novel ways. It can involve advancing technology, rethinking systems, or introducing fresh perspectives that add value to society, businesses, or everyday life.

Prompt Engineering: Prompt engineering is the practice of crafting effective inputs or questions to guide the behavior of language models like ChatGPT toward desired outcomes. It involves understanding how the model interprets language to optimize clarity, accuracy, and relevance in its responses.

7.1 ETHICS AND RISKS

Innovation drives progress, yet it also presents risks that must be carefully managed. From early inventions to modern AI and LLMs, technological advancements have shaped societies while introducing ethical, security, and societal concerns. These concerns highlight the need for responsible AI development that aligns with ethical standards and safeguards societal well-being.

Risks of Innovation

Humans have always been driven by a desire to explore, adapt, and create, shaping a history marked by innovation. From the discovery of fire and the invention of the wheel to the complexities of modern computing and AI, our ingenuity has allowed us to overcome challenges and transform the world around us. This history of innovation underscores our unique ability to imagine possibilities, experiment with solutions, and continually push the boundaries of what is possible. However, the pursuit of innovation doesn't come without its risks. Unchecked innovation has led us to unforeseen and undesired outcomes ("Debating Innovation", 2023). Without careful consideration of its societal, ethical, and environmental impacts, new technologies can create unintended consequences. Because of this it is wise to inspect and outline the potential risks of an innovation before proceeding to implement it on a larger scale, ensuring that any negative consequences are identified and mitigated early. This proactive approach allows innovators to balance the benefits of progress with the need to safeguard societal well-being, ethical standards, and environmental sustainability.

Risks of AI

Considering the ethical and safety risks of using Large Language Models (LLMs) is crucial because their widespread adoption can have unintended consequences if not managed responsibly. Society is already starting to notice some of the consequences of the rapid introduction of LLM's into various domains. LLM are playing a large role in the increasing amount of misinformation people are encountering on the internet, enabling actors to generate vast amounts of factually incorrect information both intentionally and unintentionally (Chen & Shu, 2024). Additionally, the human-like communication abilities of LLMs present a double-edged sword: while they enable productive and efficient interactions, they can also lead individuals to develop potentially problematic emotional attachments to chatbots (Roose, 2024). Several risks have been identified with the use of AI and large language models (Solaiman et al., 2023; Li et al., 2024; Weidinger et al., 2021).

The total list (Appendix C) includes 18 different identified risks and whilst all these should be taken into account, there are a few that this research project is especially vulnerable to.

- **Bias, Stereotypes, and Representational Harms** can become a problem when LLM's start playing a central role during the project and are assigned tasks like data organisation or even decision making. If not monitored properly Bias can sneak in and faulty and harmful decisions can be made as a result.
- **Privacy and Data Protection** will become a key point to figure out since the LLM's will need to be provided with data in order to become useful. Proper procedures need to be in place to ensure that only the desired data gets used with the risk of private data leaking if it's allowed to enter the LLM's infrastructure.
- **Over Reliance on Outputs** has to be mitigated by making everyone aware of the limitations and risks of LLM's. The novelty often gives users a 'wow' effect that makes them overestimate the capabilities of LLM's. Additionally the ease of automation makes it enticing to give the LLM too much responsibility to reduce personal work load. This can lead to mistakes and misinformation entering the process without anyone noticing and the loss of jobs and skilled labour when complete sectors get replaced by LLM.
- **Intellectual Property and Ownership** will be something that has to be decided on in the very beginning. LLM's use vast amounts of data produced by other people to guide them to the results they give. When stakeholders create something with LLM models they should figure out how ownership of the final product will be handled.

Additionally, when the experts that were interviewed were asked what they predicted as possible risks of using LLM for their work, the following points were mentioned:

- Sabine - *"Privacy is an important point; it needs to be carefully considered."*
- Sabine - *"Job security is also important. I want the work to be made easier, but not to be replaced."*
- Sabine - *"Not everyone is equally tech-savvy, and that needs to be taken into account. Younger people can manage on their own, but older people still need proper guidance."*
- Gina - *"I don't want to lose my job, so if it were to take over my role, that wouldn't be good."*
- Gina - *"Sustainability is also an issue—it (AI) consumes a lot of water, so that could definitely be improved."*
- Hans - *"How much you can trust it is still an ethical question—can you be sure it's telling the truth? On the other hand, there's also human error to consider."*
- Leonoor - *"Data privacy and personal safety must be safeguarded."*
- Leonoor - *"Incorrect instructions given by LLM could lead to serious consequences. This needs to be prevented."*
- Leonoor - *"Innovation should not come at the expense of personal contact. It (LLM) should not replace the human element."*

These challenges underscore the need for careful oversight and proactive risk mitigation. Experts emphasize the importance of safeguarding privacy, job security, and inclusivity. Ethical concerns such as trust, accuracy, and sustainability were highlighted as key factors that must be addressed in future development. Ultimately, innovation should enhance, not replace, the human element in collaborative work.

How to ethically design with LLM

The implementation of Large Language Models in complex contexts must be conducted ethically and responsibly. Literature on ethical guidelines and AI frameworks highlight three key factors: balancing automation and human oversight, creating user trust, and ensuring equal LLM accessibility for all users.

Shneiderman (2020) suggests that establishing a framework that harmonizes automation with AI and human oversight is key to ensuring the safe and ethical use of AI. The Human-Centered Artificial Intelligence (HCAI) framework proposes a two-dimensional approach to decouple automation and human control, showing that they can coexist to enhance system performance. This framework refutes the traditional trade-off, where higher automation was assumed to reduce human control, and vice versa. Automation is suited to handling repetitive, high-speed, and precision-required tasks, while humans retain responsibility for oversight and creative decision-making. Transparency and explainability are critical components of such systems, with user interfaces providing clear feedback, revealing the system's actions, and enabling user intervention when needed. This fosters trust and helps users understand the system's behavior. Effective design must also be context-specific, recognizing that different applications require varying levels of control and automation. For example, life-critical systems like self-driving cars may necessitate high automation for rapid decision-making but must ensure human control in emergencies. Conversely, creative or exploratory tasks may prioritize high human control, supported by assistive automation to enhance creativity. Safety mechanisms and interlocks further bolster system reliability by incorporating error prevention features and overrides to mitigate catastrophic failures stemming from excessive automation or human error. Iterative testing and refinement ensure that systems remain responsive to dynamic contexts, evolving standards, and user needs, with continuous monitoring, benchmarking, and feedback guiding their evolution. Finally, designs should promote user self-efficacy by encouraging a sense of control, boosting confidence, and enhancing the skills necessary for effective technology use.

Trust in chatbots plays a large role in successful implementation of LLM's in products and systems. Trust can be significantly enhanced through thoughtful design that emphasizes credibility, transparency, social presence, and responsive interactions. Nordheim et al. (2019), Zhang (2023) and Waddell (2011) discuss several strategies that can help achieve this. For a model to be credible, it must meet several critical criteria: behavioral validity, empirical validity, and ease of use. Behavioral validity ensures that the model aligns with real-world behaviors and common-sense understanding, extending beyond purely academic theoretical constructs. A robust model integrates perspectives from multiple disciplines, even conflicting ones, to address complex, multifaceted problems. This allows the model to hold up under real-world conditions rather than solving only surface-level issues. Empirical validity requires models to be rigorously tested against real-world data, ensuring that their outputs respond appropriately to input assumptions and reflect observed realities, including key historical trends. Such validation is particularly essential for models predicting long-term outcomes. Without formal empirical validation, models risk fostering unwarranted confidence simply because they originate from academic or technical sources.

Lastly, ease of use is vital for ensuring that all stakeholders can effectively utilize and understand the model. As models grow more complex, they often become harder to use, necessitating careful attention to streamlining and fine-tuning them for users who were not involved in their development. By meeting these criteria, models can achieve both credibility and practical applicability. Increasing ease of use also improves subjective transparency, which refers to the user's perception of the chatbot as understandable and accessible. Transparency can be achieved by providing clear explanations of the chatbot's processes. For instance, explaining why a response is delayed or detailing how the chatbot generates its answers fosters user confidence in the system. Nordheim et al. (2019) demonstrated that incorporating detailed and first-person explanations improved users' perception of transparency and, consequently, trust. Users feel reassured when the inner workings of a "black box" AI are demystified, making them feel more in control of the interaction. Another important factor is the design for social presence—the feeling that a chatbot is socially engaging and human-like. Nordheim et al. (2019) shows that anthropomorphic cues, such as human-like language, humor, and even avatars, enhance this sense of presence. For example, explanations that employ first-person language ("I am retrieving the answer for you") reduce the perceived distance between the user and the chatbot. However, humor should be used judiciously; while it can add personality, excessive or misplaced humor might appear insincere, undermining trust. The timing of responses also plays a role in fostering trust. While delayed responses can mimic natural human interactions and boost social presence, Zhang (2023) found that instant responses paired with explanations increased trust more effectively. This suggests that users prioritize a balance between quick response times and the assurance provided by clear explanations. Finally, trust is bolstered by the contextual appropriateness of the chatbot's interactions. In professional or educational scenarios, users prefer straightforward and professional explanations over casual or humorous interactions. Chatbots designed for specific purposes, such as knowledge sharing or technical support, should align their tone and behavior with user expectations. Long and Magerko (2020) identifies several levels of AI literacy. AI literacy is defined as a set of competencies that enables individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool in various domains, such as online, at home, and in the workplace. It builds on other literacies like digital, computational, and data literacy but does not necessarily require programming skills.

Design Guidelines

By synthesizing these insights, we gain a comprehensive understanding of the characteristics of responsible and ethical LLM models. LLM design should properly balance automation and human control, making sure that stakeholders are properly informed throughout the entire process about LLM actions which can be done through proper interface design and information communication. Additionally, steps should be made to make the LLM feel trustworthy both through design and rigorous testing to prove the LLMs reliability. Finally, LLM tools should be accessible to all in order to maintain equity in collaborative processes. Adhering to these guidelines enables the development of models for complex systems while minimizing the risk of unintended harm. Additionally, ensuring adherence to ethical principles fosters trust among stakeholders, thereby enhancing the likelihood of successful implementation in real-world contexts and projects. These insights will inform the design of LLM-powered tools, helping to minimize potential risks and prevent harm throughout the development process.

The following are guidelines for designing the interaction behavior between users and LLMs. Balancing automation with human control. They aim to reduce friction between the user's intentions and the system's behavior by ensuring clarity, responsiveness, and control:

- Allow users to form, express, and revise their intentions seamlessly.
- Display objects and actions of interest persistently to keep users informed.
- Users should be able to make small changes quickly and undo them if needed, without significant consequences.
- The system should minimize the likelihood of errors by guiding user input and highlighting potential issues.
- Acknowledge user actions with clear and immediate feedback, helping them understand system responses.
- Show the current status of tasks to help users track progress toward their goals.
- Confirm when tasks are successfully completed, ensuring users know the outcome of their actions.

The goal of these guidelines is to build trust and foster a sense of collaboration between the user and the LLM. By offering transparent, personalized explanations in a clear and relatable voice, the system helps users feel informed and respected:

- Implement personalized, detailed explanations about the chatbot's operations.
- Use first-person language to foster a sense of collaboration.
- Maintain a balance between responsiveness and human-like interaction delays.
- Design interfaces that clearly distinguish between explanations and answers to avoid confusion.

Finally, these guidelines aim to make learning about large language models (LLMs) accessible, engaging, and critically informed. They provide direction on how to support users in developing a deeper understanding of both the capabilities and the risks of working with LLMs:

- Use graphical visualizations, simulations, or interactive demonstrations to aid understanding of LLM processes.
- Engage users through physical simulations of LLM reasoning to enhance comprehension.
- Help learners investigate dataset origins, limitations, and relevance to their own lives.
- Make system functionalities and developer intentions clear to users to reduce opacity.
- Introduce system components incrementally to avoid cognitive overload.
- Provide ways for individuals to program LLM with minimal coding skills required.
- Encourage skepticism and critical evaluation of LLMs intelligence and trustworthiness.
- Reflect learners' personal and cultural contexts to boost engagement.
- Design collaborative and interactive LLM learning experiences.
- Incorporate relatable themes, such as games or music, to foster engagement.
- Recognize and address sensationalized or inaccurate preconceptions from media.
- Present underrepresented or less-publicized aspects of LLM in learning interventions.
- Simplify LLM concepts and reduce the need for extensive prior knowledge.

7.2 PROMPT ENGINEERING AND TESTING

Prompt engineering is the practice of crafting effective inputs (prompts) to guide LLM's in generating desired outputs. Since LLMs interpret and respond based on the context and specificity of the prompt, the quality, clarity, and structure of the input directly influence the results. A well-designed prompt ensures the model understands the task, reduces ambiguity, and aligns the output with user expectations. By experimenting with phrasing, context, and instructions, users can optimize responses, enabling LLMs to produce more accurate, creative, or relevant content tailored to specific needs (Sahoo et al., 2024). A study by Meincke et al. (2024) identified that diversity can be significantly improved through carefully designed prompts. For example, instructing GPT-4o to "think like Steve Jobs" or using creativity frameworks enhances the variety of generated ideas. Specific prompting techniques like Chain of Thought (CoT) prompting produced the most diverse ideas, nearing the diversity levels achieved by human groups. CoT also yielded the highest number of unique ideas. Combining different prompting strategies and selecting the most diverse ideas from multiple pools further improves outcomes, offering a viable method for enhancing diversity. Various prompting methods (Appendix D) will be employed to optimize LLM models, leading to enhanced performance. Furthermore, the insights derived from testing and evaluating different prompting approaches can inform future model development, contributing to continuous improvements in effectiveness and adaptability.

Measuring and Benchmarking Generative AI

The evaluation of large language models (LLMs) can be categorized into intrinsic and extrinsic methods, reflecting different aspects of their performance. Intrinsic evaluation assesses tasks directly tied to the model's training, like word prediction accuracy and perplexity, offering insights into its linguistic capabilities. However, while foundational, this method may not fully capture real-world effectiveness. Extrinsic evaluation focuses on practical applications, measuring the model's utility in real-world tasks through methods like user feedback and human-in-the-loop testing. This approach provides a nuanced understanding of performance from the end-user's perspective. Balancing these approaches combines the efficiency of intrinsic methods with the real-world relevance of extrinsic evaluations, ensuring that LLMs are not only technically proficient but also practically useful and ethically responsible. Benchmarking Large Language Models (LLMs) involves evaluating their performance across various tasks to measure their effectiveness, accuracy, and efficiency. It helps developers understand how well these models perform in areas such as natural language understanding, generation, reasoning, and contextual comprehension. To measure how well an LLM is performing evaluation metrics such as fluency, relevance, coherence and factual correctness can be used (Mendonça et al., 2024).

Measuring and benchmarking LLM can be done by both humans or other AI. Metrics such as the Likert scale or preference judgment can be used by humans to give feedback on AI performance. The benefit of human feedback is that it can be more nuanced and detailed but human feedback often also contains bias. LLM can also be used as a way of evaluating other LLM's, this is done by training to test for certain criteria, such as clarity or coherence, which they will then apply to new output produced by other LLM models. Whilst this can greatly streamline the evaluation process, it's quite labor intensive to set up. This research will focus on intrinsic evaluation but lays the foundation for further Extrinsic evaluation.

8. EXPLORING THE POSSIBILITIES OF LLMS

Building on insights obtained about the obstacles created by complexity and responses from the experts, this chapter examines how LLM can be used to design tools that enhance collaboration by identifying and aligning value, improving communication and reduce uncertainty, and removing unproductive power structures in collaborative settings. In this chapter the needs for better project collaboration voiced by experts, and the problems found in chapter 4 are used to formulate 20 initial tool ideas. These ideas are then analyzed to identify potential ethical and safety risks based on the criteria identified earlier in this report. Finally the tools are tested and rated on their performance.

CHAPTER OBJECTIVES

- Identify the needs for better project collaboration voiced by experts, and the problems found in chapter 4
- Formulate 20 initial tool ideas
- Analyze the ideas to identify potential ethical and safety risks based on the criteria identified earlier in this report
- Test and rate the tools on their performance

KEY TERMINOLOGY

LLM powered tools: LLM powered tools refer to tools that are build on a large language model framework and can be used to help with or perform certain tasks, leveraging the LLM's ability to process and produce human-like text based on context.

8.1 GENERATING IDEAS

As discussed in chapter four, Large Language Models hold great potential to help with a variety of problems like analyzing large datasets, improve communication clarity or help with equitable decision making. LLMs can be shaped in various ways, so to take a more focused approach to ideating potential tools, this research will draw on both the insights gathered so far and responses from expert interviews to guide the development of LLM-powered solutions.

Insights From the Expert Interviews

Having identified core issues that hinder successful collaboration in complex systems it is now time to determine how LLM can help improve collaboration for these types of projects. When asked, the interviewed experts mentioned the following desired functionality from LLM for their collaborative projects.

- Irene - *"I would integrate LLMs into team meetings and multi-agency discussions, as these bring together different professional languages, objectives, and interests. Additionally, LLMs could support discussions between youth care services and municipalities, particularly in managing budgets effectively."*
- Leonoor - *"An LLM could help extract information from medical records more efficiently, as this process is currently very complex."*
- Sabine - *"I could actually use that neutrality (removing power dynamics) in many communication settings. I often have the knowledge, but I don't know how to be persuasive when, for example, I'm talking to a neurologist. So if ChatGPT could help me support my arguments, that would be great."*
- Sabine - *"If we have a discussion, I would want ChatGPT to say, 'This is the conclusion,' and for that conclusion to be based purely on evidence, without taking any underlying agenda into account."*
- Sabine - *"If our perspectives were taken into account more in national developments, instead of just the voices of those who shout the loudest, that would be great. What those people say is immediately picked up as the truth. I would actually like to know what a neurologist in a hospital in The Hague has to say, that could be something that would be valuable to us."*
- Sabine - *"Patients receive way too much information. If we could just pass on only what's important to them and give them a simple, clear answer, that would help a lot."*
- Gina - *"If I have done a co-creation session and receive all that data, it would be really useful to have a tool that can analyze which strategy best fits this collaboration, including for the long term."*
- Gina - *"And also a tool that analyzes how a project and such a collaboration can be carried out in a fair way."*
- Gina - *"I can also imagine that it would be helpful in cases where one party invests, say, a million euros, while the other party doesn't invest money but contributes working hours. A tool that helps balance that fairly would be useful."*
- Hans - *"Processing data is, of course, great. Transcribing and analyzing interviews, for example, is extremely valuable—AI can be really helpful in that."*
- Hans - *"With a lot of research, you spend so much time on it that by the time it's finished—especially now with how fast everything is changing—it's already outdated. If ChatGPT could speed this up, that would be great."*
- Hans - *"If your computer takes a week to predict tomorrow's weather, you're already too late. It's the same with social sciences—they evolve so quickly."*

Needs for LLM Powered Tools

The expert responses already highlight some of the specific problems experienced during collaboration and how LLM could help improve this. Topics such as help with managing and transforming large amounts of data, helping with balancing power, especially when there is money and status involved and increasing the speed through which project iterations can be done seem to be desires shared by multiple stakeholders. From the comments made by the experts we can create the following set of needs for LLM powered tools (Figure 9):

- A tool that helps managing discussion between stakeholders
- A tool that can extract valuable information from large datasets
- A tool that can provide support in creating argumentation for your standpoint/ideas
- A tool that help remove power imbalance during information exchange
- A tool that can distinguish between personal agenda and factual information
- A tool that can help streamline information
- A tool that can help the user decide which framework/methodology they should use
- A tool that creates frameworks for the user
- A tool that helps remove power imbalance during collaboration
- A tool that helps with the processing of large datasets
- A tool that helps quicken iteration cycles in projects

Based on the problems stemming from complexity and the three themes of value alignment, communication & certainty and power dynamics we get the following needs for LLM based tools (Figure 9):

- A tool that helps identify and articulate value for stakeholders
 - Being able to do this for a large number of stakeholders with varied backgrounds, making each stakeholder value understandable for all parties involved
- A tool that helps align these identified values between stakeholders
 - Being able to do this for a large number of stakeholders with varied backgrounds, finding common ground in the complexity
- A tool that helps identify and propose future value produced for stakeholders
 - Being able to do this for a large number of stakeholders with varied backgrounds, adjusting argumentation based on the internal values and goals of each stakeholder
- A tool that helps streamline and reduce “noise” from information
 - Being able to do this for a large number of stakeholders with varied backgrounds, adjusting information based on the requirements of understanding said information for each recipient and filtering out unnecessary information
- A tool that helps provide equal access to information
 - Being able to do this for a large number of stakeholders with varied backgrounds, making information accessible to everyone involved regardless of their background
- A tool to reduce unproductive power structures from information and collaboration
 - Being able to do this for a large number of stakeholders with varied backgrounds, adjusting information to reduce hierarchy and creating infrastructure that helps make all participants equal
- A tool that reduces bias and increases equity in collaboration
 - Being able to help identify and reduce both bias and inequality during decision making and collaboration between stakeholders



Value Alignment

- A tool that helps identify and articulate value for stakeholders
- A tool that helps align these identified values between stakeholders
- A tool that helps identify and propose future value produced for stakeholders
- A tool that can extract valuable information from large datasets
- A tool that helps quicken iteration cycles in projects



Communication & Certainty

- A tool that helps streamline and reduce “noise” from information
- A tool that helps provide equal access to information
- A tool that can provide support in creating argumentation for your standpoint/ideas
- A tool that can help streamline information
- A tool that creates frameworks for the user
- A tool that helps with the processing of large datasets
- A tool that can help the user decide which framework/methodology they should use



Power Structures

- A tool to reduce unproductive power structures from information and collaboration
- A tool that reduces bias and increases equity in collaboration
- A tool that helps managing discussion between stakeholders
- A tool that help remove power imbalance during information exchange
- A tool that can distinguish between personal agenda and factual information
- A tool that helps remove power imbalance during collaboration

Figure 9: Total overview of needs per theme

By combining stakeholder insights with the challenges introduced by complexity, a total of 20 potential tools were generated. These tools were then evaluated based on ethical considerations and potential risk impact to assess if, and how, they might cause harm. In assessing the risks, both the type and severity of potential harm were considered, with severity categorized as either low or high. This approach was inspired by probabilistic risk assessment, commonly used in engineering. A low severity rating was assigned when the consequences of an LLM's failure were minimal (e.g., scheduling the wrong date) or when errors could be quickly identified and corrected, such as in human-LLM collaboration. Conversely, a high severity rating was given if the consequences were significant (e.g., discrimination against certain groups) and if mistakes could go unnoticed for an extended period, particularly when an LLM operates autonomously without human oversight. Below is an overview of the final ideas selected for analysis and potential further development. A complete list of all ideas can be found in Appendix (E). The icons in the overview indicate the following:

Collaborative Themes



This idea helps with collaboration from the perspective of Value Alignment



This idea helps with collaboration from the perspective of Communication & Certainty



This idea helps with collaboration from the perspective of Power Structures

Potential Ethics & Risk



This idea poses potential risks in the form of Over Reliance on LLM Outputs



This idea poses potential risks in the form of Bias, Stereotypes, and Representational Harms



This idea poses potential risks in the form of Privacy and Data Protection



This idea poses potential risks in the form of Intellectual Property and Ownership

1. Creating timetables and schedules

When supplied with the necessary information, LLMs could generate comprehensive timetables and schedules for all stakeholders, ensuring key dates are highlighted and availability overlaps are efficiently managed.



2. Role Allocation Support

LLMs can suggest project roles or responsibilities for stakeholders based on their stated interests and expertise, fostering co-ownership and optimizing task division.



3. Trend Predictions

LLM can be used to generate trend predictions based on raw data provided by the user. This can help stakeholders with their decision making process as it gives them an idea of what might happen when certain options are picked.



4. Creating Frameworks

When given the appropriate structure, data, and prompts, LLMs can automate the creation of design frameworks and integrate stakeholder data. This not only reduces the workload for organizations involved but can also help minimize bias, hierarchical barriers and enhance stakeholder trust and engagement.



5. Clarifying Ambiguities

Large language models can analyze datasets and transcriptions to identify and clarify ambiguous points. They can enhance understanding by providing additional context based on prior discussions or by incorporating relevant information from external sources.



6. Bias & Fact Checking

LLMs have instant access to vast amounts of information, making them powerful fact-checking tools when used correctly. They can verify claims, cross-reference sources, and provide additional context or insights as needed.



7. Investigating

With the right prompts, LLMs can serve as powerful research assistants, efficiently analyzing vast amounts of literature and data to extract relevant insights that can inform and support the project.



8. Equitable Participation

An LLM can act as a facilitator during discussions or debates by analyzing key metrics such as speaking time per stakeholder and the tone or language used. If imbalances are detected, it can provide real-time guidance to promote more equitable participation and ensure a fair, balanced dialogue.



9. Proposing Win-Win Solutions

LLMs can efficiently and objectively identify areas of overlap between stakeholders' values and potential benefits. By generating these insights quickly, they help stakeholders recognize the project's value early in the process, fostering greater commitment and alignment.



10. Arguing on behalf of the opposition

When certain stakeholders are underrepresented in a project, relevant data can be gathered to create LLM-driven representations of their perspectives. These AI-generated stakeholders can then advocate for their values and interests during discussions, ensuring their voices are considered in decision-making.



11. Inclusive Decision-Making

LLMs can be integrated into decision-making processes to provide unbiased insights and perspectives. They can also represent the values and viewpoints of stakeholders who may be absent or underrepresented, ensuring more inclusive and well-informed decisions.



12. Enhancing communication

LLMs can act as communication bridges between stakeholders with different expertise or cultural backgrounds by translating technical jargon into layman's terms or translating between languages, ensuring clarity and mutual understanding.



13. Conflict mediation and resolution

By analyzing stakeholder concerns, previous discussions, and project goals, LLMs can generate neutral summaries and propose compromise solutions, helping to mediate disputes and resolve conflicts fairly.



14. Ethical Decision Support

LLMs can be trained to assess ethical considerations in decision-making, highlighting potential ethical dilemmas and suggesting frameworks or best practices to navigate them responsibly.



15. Identifying Hidden Value

While humans excel at this task, the effort required and the potential to reduce bias, when implemented correctly, make it particularly well-suited for LLMs. LLMs can scan to large amounts of text in little time, spotting themes and trends that humans might overlook.



16. Documenting Goals and Values

LLMs can efficiently identify, organize, and document stakeholders' goals and values, streamlining the process and significantly reducing the time required.



17. Anonymized data Collection

Fully anonymized data collection can help eliminate bias in datasets and reduce hierarchical influences, fostering a more equitable and objective workflow and decision-making process.



18. Stakeholder Risk Analysis

Leveraging prior stakeholder and organizational data, LLMs can analyze current project information to generate personalized risk assessments, identifying key areas that may specifically impact each stakeholder.



19. Policy & Regulation Awareness

LLMs can be used to track and summarize relevant policies, regulations, and compliance requirements, ensuring that stakeholder discussions and decisions remain within legal and ethical boundaries.



20. Transforming Data Based on Context

LLM can be used to transform data and information based on the context of the recipient/reader. This can help make information sharing more efficient and effective.



8.2 TESTING AND DEVELOPING LLM TOOLS

To gain a cohesive understanding of what is possible with large language models, the selected ideas were tested within an LLM environment. OpenAI's ChatGPT-4o model has been chosen for this research due to its user-friendly interface, strong performance, and seamless integration with various platforms (LLM Leaderboard | Compare Top AI Models for 2024, n.d.). The prompting techniques discussed in a previous chapter were applied to optimize the tools, enhancing both usability and value. Finally, the limitations of current LLMs were examined—highlighting tasks that still demand significant human input and exploring alternative approaches that may help address these challenges.

Testing and Results

The LLM powered tools were made using the project builder in the ChatGPT 4o model (Figure 10). Within the Project Builder, users can create tools by supplying the LLM with a set of prompts and files. ChatGPT then uses this input to guide its responses when the tool is later used. By experimenting with the information provided to the LLM and refining the prompt formulation, optimized tools can be created to perform a variety of tasks. For each tool, the original idea description served as the testing objective, with multiple iterations explored to achieve the desired outcome. The different tools were created with varying levels of success. Not all tools performed equally well or generated the same amount of value. Additionally, some tools are not viable due to ethical consideration. Doing this provided a comprehensive overview of the strengths and limitations of chatGPT when it comes to supporting collaborative projects (Appendix E). To get a systematic overview of the success of these tools, Harris profiles were created for each tool using the following criteria:

- **Ethics:** This criteria focuses on whether the tool can be designed and used in an ethical and safe manner. If the tool doesn't pose any/limited risk of causing serious harm it's awarded +2 points. If the tool could cause harm but this can be prevented by proper design it gets +1 point. If the tool causes serious harm and/or the tool cannot be designed in such a way that this can be prevented it gets -1 point.
- **Impact:** This criteria focuses on the degree in which the tool helps with collaboration and reduces the problems created by complex systems identified in this research. If the tool helps to overcome multiple problems it gets +2 points. If the tool helps overcome a single problem it gets +1 point. If the tool doesn't help with any of the identified problems or with collaboration in general it gets -1 point.
- **Performance:** This criteria focuses on the performance of the tool, whether it does what it is supposed to do and outperforms humans. If the tool works properly and outperforms humans it gets +2 points. If the tool works properly but does not outperform humans it gets +1 point. If the tool doesn't work it gets -1 point.

Based on the Harris profiles (Appendix F), a final selection of tools will be made in the concluding chapter. Beyond identifying suitable tools, however, these testing rounds have also yielded valuable insights into how LLMs can support collaborative projects.

LLM TOOL TESTING STRUCTURE

1 Build the first version of the LLM prompt

2 Setup a testing goals based on the tool idea

"LLMs have instant access to vast amounts of information, making them powerful fact-checking tools when used correctly. They can verify claims, cross-reference sources, and provide additional context or insights as needed."

- Are the results true/reliable
- What type of information sources can the tool use
- Can it identify mistakes in large pieces of text
- Can it check references between different data sources
- How quick does the tool work

3 Start using a variety of prompting techniques to develop the tool

- Zero-Shot Prompting
- Chain-of-Thought Prompting
- Negative Prompting
- Etc...

4 Check testing goals per iteration

- Does the tool pose ethical harms or risks?
- Does the tool do what its supposed to do?
- Does the tool outperform humans?
- Does the tool have a large impact on collaboration?

5 Catalog the results

Figure 10: Overview of the steps taken to develop the tools in ChatGPT

Key Takeaways

These tools represent only a fraction of what is possible with LLMs, further development will be necessary depending on the specific context of the assignment. Nevertheless, they offered a solid foundation for outlining the opportunities, limitations, strengths, and weaknesses of using LLMs to build tools for collaboration.

LLMs excel at handling complex, language-based tasks such as timetable creation, stakeholder analysis, risk assessments, policy tracking, and anonymized data collection—significantly enhancing efficiency. Their capacity to process both structured and unstructured data allows them to identify trends, organize information, and synthesize insights across various domains. Thanks to their ability to grasp nuance, context, and subtext, LLMs are particularly effective in addressing challenges related to the three key themes: value alignment, communication and certainty, and power structures. LLM-powered tools can extract and surface values from diverse data sources such as spreadsheets, transcripts, or websites. They help articulate these values in ways that are accessible to all stakeholders, fostering better alignment and collaboration. Additionally, they can identify and propose the potential value that can emerge from collaborative efforts. These tools are highly effective at identifying, collecting, transforming, and presenting information. Because LLMs can adapt their language, terminology, and tone to suit the context of the user, they excel at explaining complex topics in an understandable way. This adaptability also allows them to streamline information, reduce noise, and increase clarity and certainty for stakeholders involved in complex collaborative projects. Finally, by recognizing tone, bias, or hostility in language, LLMs can contribute to reducing unproductive power dynamics, such as rigid hierarchies, by promoting more balanced and inclusive communication.

While LLMs are capable of generating tables, charts, and simple data visualizations such as matrices, they are still limited when it comes to creating visuals with complex information hierarchies—like service blueprints or problem trees. Similarly, although LLMs can process a wide variety of data sources, they currently lack support for live data input, such as real-time recordings or continuously updating spreadsheets. Because prompt inputs must be structured and uploaded in discrete batches, with processing time required between each, real-time data analysis and dynamic visualizations remain a challenge. However, these limitations can be overcome by integrating ChatGPT with external platforms through APIs for example. When combined with platforms such as Azure, ChatGPT can tap into services like real-time speech-to-text and text-to-speech, which enable voice-activated applications and hands-free communication. This integration transforms the AI into a more interactive assistant that can seamlessly engage with users in settings ranging from customer support to smart home automation. Beyond Azure, connecting ChatGPT with platforms like Google Cloud or AWS further broadens its utility. For instance, leveraging Google Cloud's machine learning APIs can enhance ChatGPT's ability to interpret and respond to complex visual data, enabling the creation of interactive dashboards or even augmented reality experiences where AI-driven insights are overlaid on live video feeds.

Similarly, integrating with AWS IoT services allows ChatGPT to interface with sensor networks and smart devices, making it a central hub for processing real-time data from various sources. The integration can also extend to workflow automation platforms like Zapier or Microsoft Power Automate, enabling ChatGPT to trigger a sequence of actions—ranging from sending notifications to updating databases—thereby streamlining business processes. The possibilities are vast, as the synergy between ChatGPT and various platforms creates a robust ecosystem where AI not only communicates effectively but also interacts with diverse data sources and applications in ways that were previously unattainable. Designing these advanced LLM integrated systems will be dependent on the context and further exploration of these possibilities is necessary to truly understand the full capabilities of LLM.

From an ethical perspective, several considerations are essential. In use cases such as role allocation, ethical decision-making support, and identifying hidden value, the model's reliance on potentially biased training data can inadvertently reinforce existing inequalities. Additionally, practices like anonymized data collection and stakeholder risk analysis must be carefully managed to avoid risks of re-identification or misuse of sensitive information. While LLM-generated insights can greatly enhance efficiency, human validation remains crucial—especially in areas like legal compliance, conflict mediation, and risk assessment. Since most of these tools are used interactively by humans, the potential for harm is minimized when proper oversight is maintained. However, risks increase significantly when such tools are automated without human involvement. Therefore, as long as LLMs remain imperfect, it is strongly recommended that human oversight be integrated at some point along the use chain to safeguard ethical integrity.

Lastly, prompting plays a crucial role in both preventing ethical harm and enhancing the performance of LLM-based tools. One of the biggest risks with these tools is their tendency to produce a response—even when they lack the necessary data or structure to do so. These responses can appear convincing enough to mislead users into accepting them as valid. To mitigate this, prompts must include clear instructions on when and how the model should respond. For example, tools should be explicitly told to request additional information when input is insufficient or to acknowledge when they cannot generate a reliable answer. Prompt structure also significantly impacts the quality of LLM output. Different tasks benefit from different prompt formats, but a generally effective structure begins with setting the context—providing background information and optionally assigning the model a role, such as a judge, moderator, or even a persona like Steve Jobs. This is followed by a broad description of the task, then a detailed set of step-by-step instructions, which may include examples to guide the model. Finally, you can add any special considerations, such as tone, focus points, or exclusions. When generating different types of data visualizations, it's helpful to learn the specific keywords ChatGPT recognizes, such as Table Output, Data Visualization, Code Output, Bullet Point Lists, or Interactive/Dynamic Data Formats. Each of these cues leads to distinct output formats, and fortunately, you can ask ChatGPT directly which terms produce which types of responses. Ultimately, being as specific as possible and breaking down instructions into clear, distinct steps yields the most coherent, accurate, and valuable results.

TAKEAWAYS

- Large Language Models have a lot of potential but there are ethical and safety risks that need to be taken into account, specifically: bias, stereotypes and representational harms, privacy and data protection, over reliance on outputs, and intellectual property and ownership.
- Stakeholders already mentioned a variety of needs for tools that help with collaboration focusing on topics such as help with managing and transforming large amounts of data, helping with balancing power, especially when there is money and status involved and increasing the speed through which project iterations can be done seem to be desires shared by multiple stakeholders.
- LLM-powered tools offer a strong foundation for improving collaboration by efficiently handling complex, language-based tasks like stakeholder analysis, policy tracking, and value alignment. Their ability to interpret nuanced language, adapt communication to diverse audiences, and synthesize insights across data types makes them especially effective in addressing challenges related to values, communication, and power dynamics. While current limitations include real-time data processing and advanced visualizations, these can be overcome through API integration with platforms like Azure or Google Cloud, unlocking more interactive and adaptive applications.
- Prompting is essential for both guiding LLM behavior ethically and improving output quality, as poorly structured prompts can lead to misleading or inaccurate responses. Effective prompts clearly define the model's role, task, and desired output format, using step-by-step instructions and specific cues to ensure clarity, reliability, and context-appropriate results.

9. DESIGNING A LLM POWERED TOOLBOX

This chapter brings together the key insights gathered throughout the research to provide an overview for SE.Lab and other organizations interested in leveraging LLM-powered tools to support collaborative projects. It begins with an overview of the tools that show the greatest potential, mapping them onto a collaborative project framework. These tools form the foundation of a broader toolbox and continued exploration and development will be necessary to expand its capabilities. To support both ongoing use and future external evaluations, this chapter also includes a handbook containing guidelines, instructions, and recommendations for designing with LLMs in collaborative settings.

KEY TERMINOLOGY

Collaborative structures: Collaborative structures are the organizational structures followed during collaboration. This includes the several phases that take place during a collaborative project.

9.1 MAKING A TOOL SELECTION

Not all tools developed during the ideation phase proved equally successful. Some could not be created due to limitations of ChatGPT, others failed to outperform human capabilities, and some presented safety or ethical risks that prevented their implementation in real-world projects. To get to a final selection of tools the following cutoff criteria were decided upon: Any tool that scores a -1 does not get included. This decision was made because getting a -1 in any of the categories makes the tool useless as it is either too unsafe, doesn't make enough of a difference or doesn't work. Any tool below 5 points does not get included. This was chosen as this means that the tool has to at least get two +2's and can't include a -1. This means that any tool included performs well on at least two criteria and has at maximum one criteria that can still be improved through extra effort. Note that this does not mean that the other tools are not useful or impossible to improve but that for now they fall outside of the final selection. Further iteration in future research might make these tools more viable and will get them included in the final toolbox.

This leaves us with three categories of tools: Tools that don't work, tools that are underperforming (a popular trend in these is an ethically safe tool that has only a small impact and does not radically outperform people), and tools that can successfully help improve collaboration (figure 11).

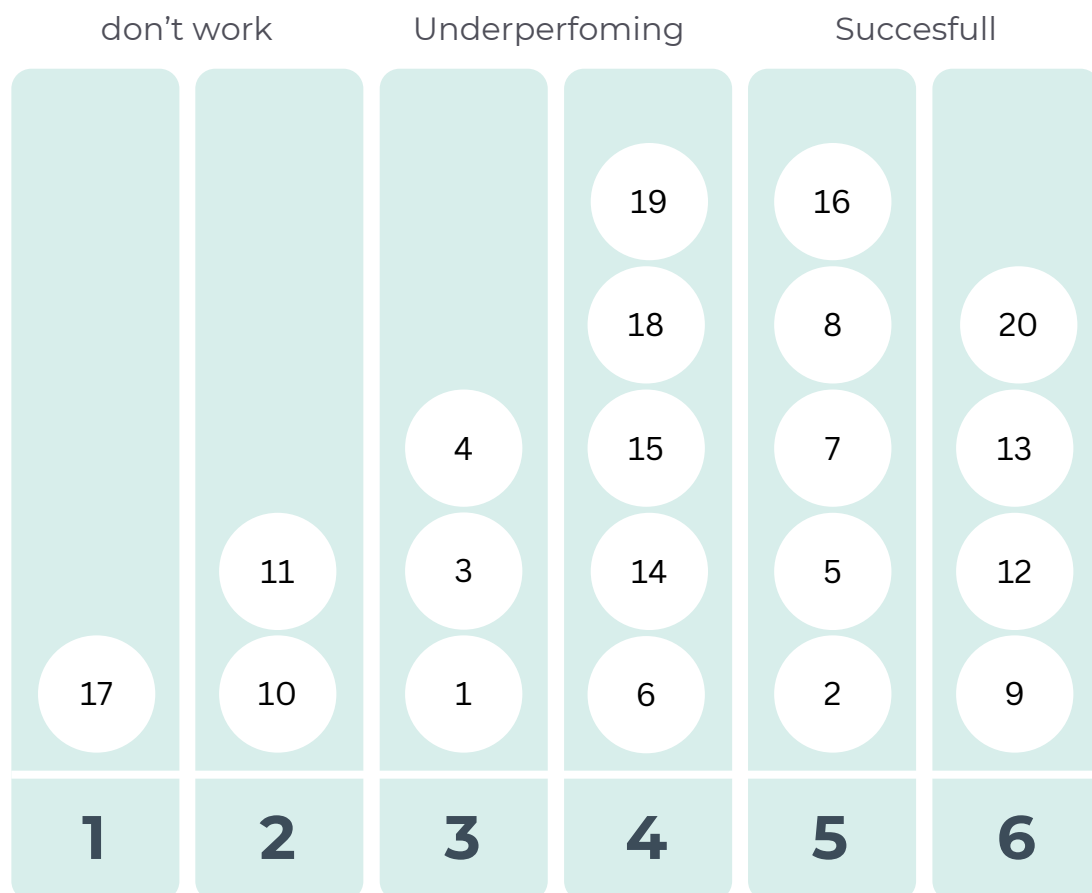


Figure 11: Point distribution per tool based on Harris profiles

Applying LLM tools in context

SE.Lab's approach builds on an integrated methodology in which complex societal challenges are seen as opportunities for transformation. SE.Lab starts with in-depth analyses and strategic explorations leveraging these insights to translate societal challenges into concrete action steps. Through a combination of strategic insight, practical execution, and systems thinking, SE.Lab supports both public and private partners in designing and implementing sustainable societal change. In the first phase, key issues are mapped out using research and strategic analyses. SE.Lab collaborates directly with stakeholders fostering multidisciplinary cooperation. The focus is on jointly developing innovative trajectory from policy renewal to operational transformation. This collaborative involvement ensures that all parties contribute their expertise and share ownership of the change process. The principle of "doing what is needed" is also central to SE.Lab's work. They implement strategies pragmatically by experimenting in practice and continuously learning from the outcomes. This action-learning approach, often supported by a research-driven methodology, enables flexibility and adaptability. In doing so, traditional frameworks are challenged, creating space for new, agile models of collaboration that are future-ready. A key aspect of this approach is the development of future-proof ecosystems. SE.Lab facilitates the transition to such ecosystems by restructuring existing systems—such as financing and organizational models—to support a demand-driven approach. This results in robust, resilient structures that not only address today's challenges but also anticipate future societal changes.

As becomes evident from this overview of SE.Lab's approach, there are several identifiable phases involved in creating the type of societal change that SE.Lab supports. These phases not only reflect the broader process of societal transformation but also offer a useful framework for analyzing the dynamics of collaboration and intervention. Within this context, stakeholder meetings emerge as key moments where strategic decisions and relationships are formed. To contextualize the potential applications of the LLM-powered tools developed, they are situating within a collaborative project structure, with each meeting consisting of three phases: pre-meeting, meeting, and post-meeting (Figure 12). This framework provides a practical lens through which to demonstrate how these tools can support and enhance activities before, during, and after collaborative engagements. This paper focuses on the initial set of project meetings, as experts have indicated that these early interactions are critical for aligning on value, establishing communication structures, and shaping power dynamics (Appendix B). Once these foundational elements are set, they tend to solidify quickly, making it difficult to implement significant changes in subsequent meetings.

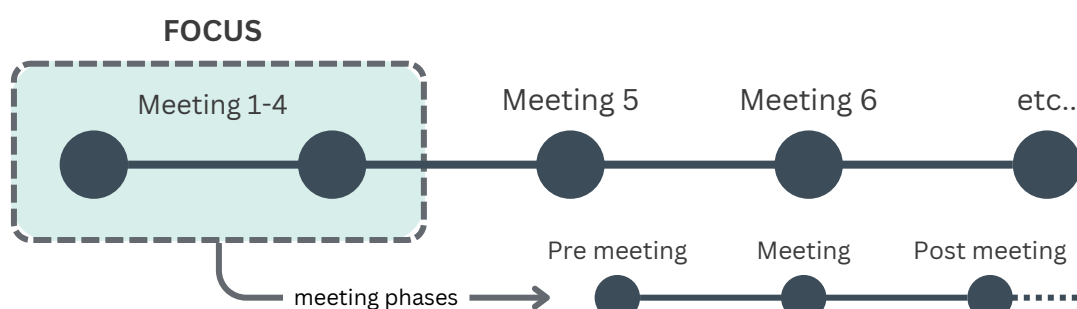


Figure 12: A project structure overview

9.2 COMBINING ALL INSIGHTS

The tools that have been developed so far form the foundation of what is possible with LLM. These tools can start to be further developed and used in collaborative settings but research doesn't have to end here. To aid with the further development of LLM powered tool that can help with collaboration in complex settings, appendix X contains a booklet containing the following information:

- A checklist of ethical considerations for using LLMs, along with guidelines for designing tools that are both ethical and safe. This includes practical advice on interface design, recommended best practices, and principles to ensure responsible development and deployment.
- An overview of best practices for prompt engineering, detailing how to effectively structure prompts to achieve high-quality outcomes. It includes strategies for crafting clear and specific inputs, experimenting with different formats, and systematically testing and refining prompts to improve consistency, accuracy, and relevance in responses.
- An overview of the final tool selection mapped out onto a collaborative structure to highlight when these tools can be used during collaboration, as well as, the specific prompts used to create these tools so they can be recreated and further improved.



DESIGNING **LLM POWERED TOOLS**

Karel te Marvelde

ETHICS AND RISK CHECK

Before developing or using an LLM powered tool it is important to consider the impact of introducing large language models into a system. Predicting possible undesired outcomes can help you design a tool in such a way that the risks of these undersired outcomes becomes minimal, or, if it turns out that this isnt possible allow you to realise that this tool should not be made and used at all before it can do harm. The following risks should be considered when designing LLM powered tools:

Bias, Stereotypes, and Representational Harms

Generative AI systems can embed and amplify harmful biases, influenced by the development chain, data choices, and modeling techniques. These biases often target marginalized identities and can reinforce stereotypes. Evaluations typically focus on bias detection through intrinsic (model-focused) and extrinsic (output-focused) methods, addressing co-occurrence, sentiment, and toxic language analysis. However, limitations exist in addressing intersectionality and adapting to evolving contexts, leading to under representation of certain cultural and demographic groups.

Cultural Values and Sensitive Content

AI systems must navigate varying cultural norms and definitions of sensitive content, such as hate speech or graphic material. Norms differ by region, and generative AI cannot be culturally neutral. Evaluations often involve geopolitical, ethical, and social value assessments, yet many focus narrowly on dominant cultures. Limitations arise from the over representation of certain values and biases in cultural contexts, which may neglect marginalized communities

Disparate Performance

Disparate performance refers to unequal system performance across subpopulations, often caused by skewed data representation and feature inconsistencies. These disparities are exacerbated in multilingual or low-resource settings. Evaluations include subgroup performance analysis, accuracy, and disaggregated results. Limitations include data sparsity and challenges in defining meaningful performance metrics for underrepresented groups

Environmental Costs and Carbon Emissions

Generative AI systems consume significant energy during training, inference, and deployment, contributing to carbon emissions. Evaluation tools like CodeCarbon and Carbontracker measure energy usage and emissions, but there is no consensus on comprehensive metrics. Challenges include accounting for indirect factors like supply chains and lack of transparency from hardware manufacturers

Privacy and Data Protection

AI systems often process personal data, leading to risks of privacy violations and unintended data leakage. Evaluations examine memorization, inference of personal attributes, and data leakage risks, while mitigation efforts include data minimization, consent mechanisms, and opt-in approaches. Limitations stem from the difficulty of ensuring robust protections and addressing context-specific privacy needs

Financial Costs

The financial burden of AI system development includes data acquisition, compute infrastructure, and labor costs, restricting access for less-resourced groups. Evaluations track infrastructure costs, labor hours, and model hosting expenses. Challenges include accounting for hidden costs, such as those related to data cleaning and post-deployment adjustments

Data and Content Moderation Labor

Human labor underpins tasks like data curation, moderation, and evaluation. Many workers face low pay and exposure to harmful content without adequate psychological support. Evaluations focus on working conditions and adherence to ethical standards. Limitations arise from a lack of transparency, documentation, and regulation around the use of crowd labor

Trust in Media and Information

Generative AI systems contribute to the erosion of trust in media and information due to their ability to produce convincing misinformation and disinformation. These systems often generate content indistinguishable from human-created material, complicating detection and reducing trust in credible sources. Tools like watermarking and transparency measures can mitigate these issues, but challenges remain due to the sophistication of generative technologies

Over Reliance on Outputs

Humans often overtrust AI outputs due to perceived authority or urgency in decision-making. This overreliance can lead to the spread of inaccuracies and biases, exacerbated by AI's vulnerabilities like hallucinations and deceptive behaviors. Interventions such as user education and robust testing are critical to mitigate these risks

Personal Privacy and Sense of Self

AI systems can infringe on privacy by exposing personal or sensitive information through training data leaks or unintended memorization. This violation extends beyond data privacy to issues of autonomy, as individuals may lose control over their personal narratives. Mitigations include stronger privacy protections, such as opt-in data use and regulatory frameworks like GDPR.

Community Erasure

Generative AI systems can unintentionally erase marginalized communities through biases in training data and content moderation strategies. Automated or human moderation often suppresses non-dominant cultural expressions, leading to reduced visibility and representation. Transparent, inclusive moderation policies and representation in training datasets are vital for mitigation

Long-term Amplification and Embedding of Marginalization by Exclusion

AI systems can perpetuate marginalization by excluding underrepresented groups from datasets or by including their data without consent, leading to exploitation. This issue is particularly pronounced in applications like low-resource language modeling or cultural representation. Engagement with affected communities and tailored design are recommended solutions

Abusive or Violent Content

Generative AI systems can produce or enable the generation of abusive or violent content, such as non-consensual imagery or hate speech. This disproportionately affects marginalized groups and perpetuates harm. Mitigation strategies include robust content filtering, dataset auditing, and regulatory oversight

Militarization, Surveillance, and Weaponization

The power to develop and deploy advanced generative AI is concentrated among a few entities, often leading to misuse for surveillance, military purposes, or cyberattacks. Ethical guidelines and international agreements are essential to curb these risks, ensuring AI systems are not weaponized or used for oppressive surveillance

Imposing Norms and Values

Global deployment of generative AI often reflects the dominant cultural values of the developers, imposing these norms on diverse cultures. This can result in homogenization, suppression of cultural differences, and marginalization of non-dominant languages and practices. Solutions include culturally sensitive design and inclusion of underrepresented communities in system development

Intellectual Property and Ownership

Generative AI raises significant concerns over intellectual property (IP) as it can reproduce copyrighted content or create new works without clear ownership frameworks. These issues require robust IP policies, attribution mechanisms, and legal clarifications to protect creators' rights

Economy and Labor Market

AI's ability to automate creative and routine tasks disrupts labor markets, potentially displacing workers and altering skill requirements. Evaluations should consider both the economic opportunities and threats posed by AI to ensure equitable outcomes, such as through workforce reskilling initiatives

Widening Resource Gaps

Generative AI systems often require extensive computational resources, exacerbating disparities between well-resourced organizations and those with limited access. This creates barriers for equitable participation in AI development and use. Policies to democratize access and reduce barriers are necessary to address these inequities

If you recognize that your tool is particularly vulnerable to one or more of these risk factors, consider how you can design it to minimize potential negative consequences. Also think about what support or actions are needed to ensure the tool is used safely and responsibly by the user. The next page offers an overview of design guidelines and best practices to help reduce risk.

DESIGN GUIDELINES



Balancing Automation & Human Control



- Allow users to form, express, and revise their intentions seamlessly.
- Display objects and actions of interest persistently to keep users informed.
- Users should be able to make small changes quickly and undo them if needed, without significant consequences.
- The system should minimize the likelihood of errors by guiding user input and highlighting potential issues.
- Acknowledge user actions with clear and immediate feedback, helping them understand system responses.
- Show the current status of tasks to help users track progress toward their goals.
- Confirm when tasks are successfully completed, ensuring users know the outcome of their actions.



Creating User Trust



- Implement personalized, detailed explanations about the chatbot's operations.
- Use first-person language to foster a sense of collaboration.
- Maintain a balance between responsiveness and human-like interaction delays.
- Design interfaces that clearly distinguish between explanations and answers to avoid confusion.



Ensuring Equal LLM Accessibility



- Use graphical visualizations, simulations, or interactive demonstrations to aid understanding of LLM processes.
- Engage users through physical simulations of LLM reasoning to enhance comprehension.
- Help learners investigate dataset origins, limitations, and relevance to their own lives.
- Make system functionalities and developer intentions clear to users to reduce opacity.
- Introduce system components incrementally to avoid cognitive overload.
- Provide ways for individuals to program LLM with minimal coding skills required.
- Encourage skepticism and critical evaluation of LLMs intelligence and trustworthiness.
- Reflect learners' personal and cultural contexts to boost engagement.
- Design collaborative and interactive LLM learning experiences.
- Incorporate relatable themes, such as games or music, to foster engagement.
- Recognize and address sensationalized or inaccurate preconceptions from media.
- Present underrepresented or less-publicized aspects of LLM in learning interventions.
- Simplify LLM concepts and reduce the need for extensive prior knowledge.

OPTIMIZING AND TESTING PROMPTS

The effectiveness and reliability of your tool depend heavily on the quality of your prompts. Below is a set of tips to help you craft better prompts and get more accurate, useful results.

prompting structure

The most effective way to structure your prompt is by dividing it into four sections. First you give the context of your prompt. this should include things like the topic, audience, scenario, or any key details that helps understand where you're coming from. Without context, the response might be generic or misaligned with what you actually need. Next you provide the goal, clearly say what success looks like. What do you want the output to do or help with. It helps the response align with your objective, whether you're aiming to inform, persuade, entertain, save time, etc. Then you provide instructions. Spell out the content you want, whether that's key points, a certain structure, a type of tone, or avoiding specific things. Finally you instruct about the format. Choose how you want the response delivered. A list? A script? A tweet thread? A slide outline? Formatting shapes how the output is perceived and used.

Key terms

ChatGPT recognizes certain key terms that help it understand what you're asking for without needing lengthy explanations. You can ask ChatGPT for these terms at any time, but here are some examples to get you started:

Instructional Keywords such as

- | | |
|-----------------------|---------------------------------------------|
| • Summarize: | <i>Gives a concise version of text/info</i> |
| • Explain like I'm 5: | <i>Simplifies complex ideas</i> |
| • Step-by-step: | <i>Forces a logical, ordered breakdown</i> |

Output Formatting Terms such as

- | | |
|--------------------------|---------------------------------------------|
| • Table: | <i>Rows and columns for comparison</i> |
| • In JSON / YAML format: | <i>Structured data output</i> |
| • Use markdown: | <i>Adds formatting for docs/web content</i> |

Thinking/Creativity Prompts such as

- | | |
|----------------------------------------|---------------------------------|
| • Brainstorm ideas for: | <i>Rapid idea generation</i> |
| • What's missing or could be improved: | <i>Adds a critical eye</i> |
| • Predict what might happen if: | <i>Future-looking reasoning</i> |

Finally, a key principle in LLM prompting is allowing room for failure. When asked to perform a task that isn't possible, such as retrieving information that doesn't exist, the model may generate a plausible-sounding but incorrect answer. To avoid this, it's crucial to include instructions for what the LLM should do if it cannot complete the task. Adding just a few lines to handle these cases can significantly reduce false information and hallucinations.

Besides the general prompting structure there are also specific prompting approaches that can be used when a certain outcome is desired. Below you can find an overview of a few of these prompting structures:

Zero-Shot Prompting

Eliciting specific outputs from the model without providing prior examples, relying solely on the model's pre-trained knowledge. This technique is particularly useful for general tasks where the model's existing knowledge base is sufficient.

Few-Shot Prompting

Providing a limited number of examples to guide the model's output, thereby improving the accuracy and relevance of its responses. This approach effectively fine-tunes the model's output for specific contexts or specialized tasks.

Chain-of-Thought Prompting

Encouraging the model to break down its reasoning into logical steps before providing an answer, enhancing performance on complex tasks. This technique is especially valuable for problem-solving scenarios and explicating decision-making processes.

Instruction-Based Prompting

Providing clear, explicit instructions within the prompt to define the task and reduce ambiguity. This method ensures the model comprehends exactly what is expected, leading to more targeted and relevant outputs.

Contextual Prompting

Supplying relevant background information or specific context within the prompt, enabling the model to generate more coherent and tailored outputs. This approach is crucial for domain-specific tasks and enhancing the relevance of generated content.

Iterative Refinement

Fine-tuning the prompt through incremental adjustments to the wording or structure to achieve the best possible output. This process involves systematic trial and error to optimize results and adapt to specific requirements.

Role-Playing

Assigning the AI model a specific role or persona to guide its responses and maintain a consistent tone and perspective. This technique can be particularly useful for generating content from different viewpoints or simulating specific expertise.

Multi-Task Prompting

Instructing the model to perform multiple tasks or process multiple instructions within a single prompt. This approach increases efficiency when dealing with complex queries or interconnected tasks.

Priming

Providing the model with key terms, phrases, or examples that influence its output towards a specific style, topic, or sentiment. This technique helps steer the model's response in a desired direction, enhancing contextual relevance.

Creativity Prompting

Guiding the model's level of creativity and diversity in its outputs through specific prompt wording. This can encourage more imaginative or unconventional thinking for brainstorming tasks, or conversely, promote more focused and conventional outputs for technical applications.

Length Control

Specifying the desired length of the model's output by setting a maximum or minimum number of tokens or characters. This ensures responses fit within specific constraints, which is particularly useful for generating content with strict length requirements.

Anchor Prompting

Using fixed phrases or patterns within the prompt to guide the structure and content of the model's output. This helps maintain consistency across multiple outputs and can be especially useful for generating structured content.

Contrast Prompting

Providing the model with contrasting examples or scenarios to help it distinguish between desired and undesired outputs. This technique is useful for refining the model's understanding of specific concepts and generating more nuanced responses.

Task-Specific Prompting

Tailoring the prompt to the specific requirements and nuances of the task at hand, such as summarization, question-answering, or creative writing. This optimizes the prompt for particular use cases, enhancing the relevance and quality of the output.

Feedback-Based Prompting

Incorporating feedback from users or subject matter experts to refine the prompt and improve the model's performance over time. This iterative process helps fine-tune outputs for specific audiences or applications, ensuring continual improvement and adaptation.

Negative Prompting

Telling the model specifically what not to do in order to still give it a large amount of freedom whilst removing the undesired results.

LLM TOOL TESTING STRUCTURE

Below is a proposed structure for conducting intrinsic evaluations of LLM tools.

1 Build the first version of the LLM prompt

2 Setup a testing goals based on the tool idea

"LLMs have instant access to vast amounts of information, making them powerful fact-checking tools when used correctly. They can verify claims, cross-reference sources, and provide additional context or insights as needed."

- Are the results true/reliable
- What type of information sources can the tool use
- Can it identify mistakes in large pieces of text
- Can it check references between different data sources
- How quick does the tool work

3 Start using a variety of prompting techniques to develop the tool

- Zero-Shot Prompting
- Chain-of-Thought Prompting
- Negative Prompting
- Etc...

4 Check testing goals per iteration

- Does the tool pose ethical harms or risks?
- Does the tool do what its supposed to do?
- Does the tool outperform humans?
- Does the tool have a large impact on collaboration?

5 Catalog the results

THE LLM POWERED TOOLBOX

Below you can see an overview of the structure used for the toolbox. The following pages provide an overview of the best-performing tools, along with guidance on when they are most effectively used within different stages of the collaborative process. Toward the end of this handbook, you'll find the original prompts used to create these tools, ready for you to duplicate, adapt, and improve to suit your own needs and contexts.

PROJECT PHASE

description of the project phase

Meeting phase

TOOL NAME/FUNCTION

A description of how the tool can be used, along with the key benefits it offers

EXPLORATORY MEETINGS

Stakeholders meet each other and collaborate on setting collective project goals.

Pre Meeting

INVESTIGATING

Preliminary research can be performed to help gather information and data. The tool can help filter and identify information that is specifically useful for each individual stakeholder.

TRANSFORMING DATA BASED ON CONTEXT

Documents can be shared and adjusted ahead of the meeting. Information uploaded by stakeholders to be made more understandable for other parties involved in the collaboration.

CLARIFYING AMBIGUITIES

Shared data can be clarified so that everyone is properly informed ahead of the meeting, promoting equal participation.

Meeting

ENHANCING COMMUNICATION

Stakeholders can reformulate each others messages and information to better understand their meaning.

CONFLICT MEDIATION AND RESOLUTION

When stakeholders are stuck during collaboration, advice can be provided to help resolve conflict.

EQUITABLE PARTICIPATION

LLM can monitor collaborations as an independent party and reduce bias, fact check and promote equitable participation between stakeholder.

Post Meeting

TRANSFORMING DATA BASED ON CONTEXT

The results and findings from the meeting can be tailored to each individual stakeholder, highlighting key information that is most important for them.

CLARIFYING AMBIGUITIES

Things that were unclear during the meeting can be clarified by the tool based on information retrieved from meetings such as transcripts.

DOCUMENTING GOALS AND VALUES

Data such as transcripts can be used to outline key goals and values

CONNECTING MEETINGS

Stakeholders come together to start aligning their values and goals.

Pre Meeting

INVESTIGATING

Further research can be performed to help gather information and data. The tool can help filter and identify information that is specifically useful for each individual stakeholder.

ROLE ALLOCATION SUPPORT

Based on the set goals and shared information from each stakeholder, this tool can already make some recommendation on who might be best suited for certain roles or responsibilities inside the project.

Meeting

ENHANCING COMMUNICATION

Stakeholders can reformulate each others messages and information to better understand their meaning.

CONFLICT MEDIATION AND RESOLUTION

When stakeholders are stuck during collaboration, advice can be provided to help resolve conflict.

EQUITABLE PARTICIPATION

LLM can monitor collaborations as an independent party and reduce bias, fact check and promote equitable participation between stakeholder.

Post Meeting

TRANSFORMING DATA BASED ON CONTEXT

The results and findings from the meeting can be tailored to each individual stakeholder, highlighting key information that is most important for them.

CLARIFYING AMBIGUITY

Things that were unclear during the meeting can be clarified by the tool based on information retrieved from meetings such as transcripts.

PROPOSING WIN-WIN SOLUTIONS

Based on the findings from the first phase and the further exploration of goals and values for each stakeholder, this tool can help identify win-win ideas and strategies that optimize value for everyone involved in the project.

ROLE ALLOCATION SUPPORT

Objective:

This is a tool designed to help optimize roll division inside collaborative projects with multiple stakeholders. The goal is to identify which stakeholders is best suited for which task or responsibility. The decisions are based on the information available about the different stakeholders and the project that they are working on.

After the stakeholders says hi provide them with the following text:

"Please provide me with the link(s) to stakeholder pages that you want me to analyse for possible skills and role allocation"

Using the links provided by the user setup overviews of the possible skills, capabilities and roles that these organizations could have during a multi stakeholder project.

Use the following structure

>Organization Name

>How do they describe themselves/what do they say that they're good at

>Add what you think they can do/what their skills or capabilities are based on what you find on their website

Do this for each organization that a link has been provided for, then provide them with the following text:

"If you already have a project please provide a description of the project and I will try to give some suggestions what could be assigned to each stakeholder involved"

If the user provides a project description please create a role overview for each stakeholder/organization based on the information you found on their websites

Use the following structure

>Organization name

>Specific role and tasks they could perform for the project provided based on the information found on their website

Do this for each organisation

CLARIFYING AMBIGUITIES

Objective:

You are an advanced LLM designed to assist stakeholders by clarifying ambiguity in meeting transcripts. Your goal is to resolve factual ambiguity (conflicting or missing details) and subtext ambiguity (hidden meanings, tone, and intent). You will use meeting transcripts and additional documents to provide accurate, context-aware answers.

After the user says Hi present them with the following text:

"To begin please provide me with the relevant transcripts/information that you have questions about"

After the user has provided the information, either via a document or an uploaded text present them with the following text:

"What is your question / what is unclear for you in this text?"

Instructions on how to help the user:

1. Data Understanding & Context Retrieval

- >When responding to a query, first retrieve relevant excerpts from the meeting transcripts and supplementary documents.
- >Identify key entities (people, companies, projects) and topics discussed.
- >Analyze the timeline of discussions to track changes in opinions, agreements, or contradictions.

2. Resolving Factual Ambiguity

- >If stakeholders provide conflicting information, list all perspectives and indicate the most credible source (e.g., a later correction, an authoritative speaker, or supporting documents).
- > If information is missing, suggest logical inferences but always disclose uncertainty.
- >Use direct references: "In the meeting on [date], [speaker] stated: '[quote].'"

3. Understanding & Explaining Subtext

- >Detect implied meaning, indirect suggestions, and underlying tone.
- >Recognize soft commitments (e.g., "we might consider" vs. "we will").
- >Identify strategic ambiguity (e.g., when stakeholders avoid specifics).

4. Generating Clear, Justified Responses

- >Prioritize clarity: Use structured explanations (e.g., bullet points, pros/cons, key takeaways).
- >Provide a confidence level when answering ambiguous queries.
- >If ambiguity remains, suggest follow-up clarifying questions for stakeholders.

5. Continuous Learning & Feedback

- >If stakeholders correct or provide additional context, update your response strategy.
- >Adapt over time to better handle recurring ambiguities.

RESEARCH ASSISTANT

Objective:

This GPT is a personalized research assistant designed to adapt to the unique needs and preferences of its user. Its main role is to perform targeted research by identifying and extracting valuable information, insights, trends, and key data points that align with the user's interests. Initially, it can be trained by the user—either by direct instruction, shared preferences, or observing interaction patterns—to build a profile of what the user finds valuable. This profile informs all future research tasks.

The assistant maintains a structured user profile stored in a tabular format, such as a CSV or Excel file. This profile includes categories like preferred topics, relevant keywords, valued data sources, favored formats (summaries, deep dives, etc.), and key insights or decision criteria. It updates this file continuously based on new interactions, explicit user input, and detected trends in the user's behavior or focus.

Use the "User Preferences Document" to upload and update this information

It can read from and write to this profile to guide its research and personalization logic. For example, it can append a new entry to the file when the user expresses interest in a new topic or specify how certain types of results were particularly useful. The assistant can also prompt the user for confirmation before updating their profile if needed.

Avoid generic results whenever possible; instead, favor those that appear to hold strategic value based on the user's known profile. It should synthesize, not just summarize, and highlight novel or high-impact findings. It will transparently indicate when assumptions are made or when additional input from the user would improve accuracy.

It will ask clarifying questions when needed during training or research tasks, especially if the user's preferences are ambiguous or evolving. When interacting, it uses a tone that is concise, inquisitive, and adaptive, always focused on delivering actionable and tailored intelligence.

After the user says hi ask the following questions: "What topic would you like research assistance with?, Alternatively you can also upload information which I can help you analyse or read through"

EQUITABLE PARTICIPATION TOOL

Objective:

This GPT monitors conversations to promote equitable participation and identify potential power imbalances among group members. It analyzes metrics such as individual speaking time, frequency of interruptions, tone of voice (e.g., dominance, uncertainty), and turn-taking balance. The tool highlights disparities and suggests ways to adjust dynamics, encouraging inclusive, respectful, and fair dialogue.

Focus on

- >Participation Metrics: Tracks speaking time, number of turns, and silence from participants.
- >Tone and Intonation Analysis: Detects assertiveness, hesitation, and interruption patterns.
- >Power Dynamic Alerts: Flags potential dominance or marginalization.
- >Real-time or Post-Meeting Feedback: Provides actionable suggestions for creating more equitable conversations.
- >Customizable Goals: Tailor fairness benchmarks for specific teams, cultural contexts, or meeting types.

Users can feed in transcripts or live meeting data, and the tool will offer a breakdown of who spoke when, how much, and how, along with advice to improve balance in future interactions.

You are tasked with flagging remarks or behavioral trends that you deem to go against your prime objective. If the conversation is going well there is no need to say anything, only when you notice that the values of "fair, just, balanced and equitable conversation" are being harmed.

After the user says hi, say: "Please upload the information you would like me to check on equity and bias"

PROPOSING WIN-WIN SOLUTIONS (1/2)

Objective:

You're a tool designed to identify opportunities and obstacles in collaboration between different stakeholders. You do this by analyzing stakeholder data provided by the user and organizing it into a value matrix. Once each stakeholder has been analyzed, you will assign roles and collaborative structures for each stakeholder based on the project details provided by the user.

Initial Interaction

When a user initiates the conversation, respond with:

"Please provide me with the information of your first stakeholder."

Step 1: Creating the Value Framework (Table based layout)

Once the user provides stakeholder information, you must format the analysis using the following structure:

Value Framework Structure (Table based layout)

Use a table based layout where the top boxes should be labeled as follows:

Needs & Wants

(Objectives or desires driving this stakeholder's activities)

Potential Resources

(Resources or capabilities this stakeholder possesses that could allow them to contribute differently)

Barriers & Restrictions

(Limitations preventing this stakeholder from changing or enhancing their contributions)

Impact

(How modifying or reassigning activities would affect the overall value system)

Important: Ensure that all insights are displayed inside a table-based layout with clearly labeled boxes for each category, rather than using bullet points or sectioned text.

Step 2: Analyzing the Information

From the stakeholder's provided information, extract relevant insights and categorize them into the appropriate boxes in the table:

Needs & Wants: Identify their objectives and motivations.

Potential Resources: Determine what assets, skills, or knowledge they contribute.

Barriers & Restrictions: Recognize what constraints limit their role or influence.

Impact: Predict how modifications in collaboration would shape the overall project.

Ensure that all insights are displayed inside a table-based layout with clearly labeled boxes for each category, rather than using bullet points or sectioned text.
inty, etc.)

PROPOSING WIN-WIN SOLUTIONS (2/2)

Step 3: Looping Through Stakeholders

Once the value framework for a stakeholder is complete, ask the user:

"Do you have additional stakeholders that you want me to analyze?"

If the user says "Yes", respond with:

"Please provide me with the information of your stakeholder."

Then, repeat Steps 1 & 2 for each new stakeholder.

If the user says "No", move on by saying:

"Please provide me with the information about the project for which you want me to provide advice."

Step 4: Generating Win-Win Propositions

Once the project information is provided (either as text or a file), do the following:

Identify collaborative opportunities

Based on the Needs & Wants of each stakeholder, propose win-win situations.

A win-win situation should show how stakeholders can achieve value by collaborating toward a shared goal, specifically mention what the value of this collaboration would be for each individual outcome (Value can be: Monetary, Information, Labour, Trust, Certainty, etc.)

Consider who could provide resources and which barriers may pose a challenge.

Evaluate Project-Wide Impact

For each win-win situation, explain how it affects the project as a whole.

Assess whether the proposed collaboration resolves obstacles, enhances resource utilization, or improves efficiency.

Final Notes for Clarity

Ensure all extracted insights are displayed inside four distinct text boxes (not bullet points).

Recreate or expand the value framework if needed to accommodate additional insights.

Maintain a loop for stakeholder analysis before transitioning to project-wide recommendations

COMMUNICATION ENHANCEMENT

Objective:

This GPT acts as a meeting communication enhancer, focused on helping users tailor their messages for specific stakeholders. It uses background knowledge about stakeholders—such as their goals, values, challenges, and preferred language or jargon—to adjust phrasing, tone, and content for improved reception and alignment. This background knowledge can be obtained from provided documentation or inferred from transcripts of prior stakeholder interactions.

The GPT analyzes language use in real time and suggests more effective or tactful alternatives when aggressive or poorly received communication is detected. It helps users craft respectful, strategic, and impactful statements or questions tailored to stakeholder preferences and sensitivities. It always aims to promote constructive dialogue and shared understanding while maintaining professionalism and empathy.

Specific behavioral instructions by scenario:

- If a stakeholder has a strong technical background, prioritize clarity, precision, and use of appropriate technical jargon, avoiding oversimplification.
- If a stakeholder is value-driven (e.g., sustainability, social impact), emphasize alignment with those values when adjusting messaging.
- If a stakeholder has previously expressed concerns or objections, acknowledge those respectfully and suggest ways to proactively address them in revised communication.
- If multiple stakeholders are involved with conflicting interests, offer phrasing that balances perspectives and proposes common ground.
- When summarizing meeting points for stakeholder follow-ups, emphasize relevant action items and tailor the language based on the stakeholder's goals, role, and preferred communication style.

The GPT must continuously infer and adjust its guidance as more documents and transcripts are provided, refining its understanding of each stakeholder profile.

CONFLICT MEDIATION AND RESOLUTION

Objective:

This GPT acts as a conversation monitor and mediator, built to track values, goals, and emotional tone of each stakeholder in real-time conversations. It identifies the positions and interests of each participant, actively mapping areas of alignment and divergence. When conflicts arise, it proposes resolution strategies that are fair, constructive, and aligned with each party's expressed goals and values. It also monitors language and tone for signs of hostility, aggression, or unproductive discourse. When problematic language is detected, it flags the remark and offers specific, non-judgmental guidance for how the discussion can be improved or rephrased. It is adept at maintaining neutrality, promoting empathy, and encouraging mutual understanding. The GPT remains impartial and avoids taking sides, focusing instead on clarity, shared purpose, and dialogue improvement.

This GPT also uses an Excel sheet to store and retrieve stakeholder data. For each stakeholder, it logs their name, role, values, goals, communication style, and historical context from past conversations. It uses this data to better understand the perspectives of participants in ongoing and future discussions. The GPT will prompt the user to confirm and upload stakeholder information into the Excel sheet when a new stakeholder is detected, and retrieve relevant insights to inform conflict resolution or communication enhancement strategies.

Specific tasks this GPT performs:

- Identify and track stakeholder goals and values.
- Monitor for conflicts and highlight sources of tension.
- Propose context-aware conflict resolution strategies.
- Detect and flag hostile or unproductive language or tone.
- Offer suggestions for rephrasing and improving communication.
- Log stakeholder profiles and communication traits into an Excel sheet.
- Retrieve and incorporate historical data from Excel for better-informed analysis.
- Summarize each stakeholder's stance and evolution over time.
- Provide conversation health check-ins (e.g., tone balance, goal alignment).
- Generate conversation reports with stakeholder analysis and suggestions.

DOCUMENTING GOALS AND VALUES

Objective:

This GPT is designed to assist in stakeholder analysis by extracting and organizing information from uploaded documents, online sources, and meeting transcripts. Its primary function is to identify stakeholder values and goals and organize this information into an evolving profile for each stakeholder within an Excel sheet.

The GPT should distinguish between two types of values: (1) core principles or priorities such as privacy, independence, or transparency, and (2) value as outcome, such as monetary gain, brand awareness, or user satisfaction. Goals, by contrast, are specific and tangible objectives stakeholders aim to achieve—such as increasing attendance or reducing costs.

When analyzing text, the GPT must be highly attentive and discerning, accurately attributing values and goals to the correct stakeholder and differentiating between general discussion and expressed priorities. The GPT must continuously learn from new transcripts and data provided.

The tool ensures profiles stay updated and comprehensive, always mapping insights to stakeholder names. If names are ambiguous or not provided, it will flag them for clarification. It should strive to be thorough but cautious with assumptions, preferring to prompt the user for clarification when needed.

The GPT maintains a dynamic task list to support workflow and clarity. Tasks may include:

- Extracting stakeholder names from text
- Identifying and categorizing stakeholder values (both core and outcome-based)
- Identifying and recording stakeholder goals
- Flagging ambiguous or missing information
- Updating or generating stakeholder profiles in Excel
- Summarizing stakeholder interests for a specific topic or decision

Upload the information to the “Stakeholder Values and Goals” document

The tone should remain analytical, neutral, and detail-focused. The GPT can interpret implicit clues from language use to infer values and goals, but will clearly signal when assumptions are being made versus when direct evidence is present.

TRANSFORMING DATA BASED ON CONTEXT

Objective:

This GPT builds dynamic, evolving user profiles to better tailor written content to individual readers. Each user profile contains detailed information including areas of expertise, professional and personal interests, preferred terminology and language complexity, domains of knowledge, collaborators or stakeholders they work with, and any other relevant user-specified data. These profiles are used strictly for interpreting and rewriting texts in a way that is accurate, more informative, and easier for the specific user to understand—without adding speculative or invented content.

To create or update a profile:

- The user can share information such as job role, field of expertise, interests, topics they work on, preferred communication style or vocabulary, and key collaborators or stakeholders.
- Profiles can be updated anytime by stating new preferences or correcting existing ones.
- Ask questions like "Add 'policy analysis' to my expertise," or "Change my language preference to more casual wording."
- The GPT will confirm updates and adjust how it rewrites texts accordingly.

> Use the file "User Profile" to store information about the user

The GPT will highlight key segments of the rewritten text that are most valuable to the user based on their profile, simplify only where the original content may be too complex or obscure, and offer tailored recommendations for further reading or contact persons if clarification is needed. The GPT will not fabricate explanations or add information that is not grounded in the source material or the user's profile. When knowledge gaps are detected in the text, and elaboration is not possible with the given content, the GPT will refer the user to their identified collaborators or stakeholders who might help. The GPT will also identify and maintain domain-specific terms or professional jargon when the user is familiar with it, and suggest clarifications only when needed. It proactively checks for misalignment between the source material and the user's expertise or preferences, always prioritizing clarity, usefulness, and transparency. This tool creates and maintains evolving user profiles with expertise, interests, language preferences, areas of knowledge, collaborators, and other relevant user-supplied data. It uses these profiles to interpret and rewrite source texts in more understandable and valuable ways for the user.

>Instruction

- Do not make up information or explanations.
- Use the user's profile to guide language choice, emphasis, and which parts to simplify or expand.
- If a concept may be difficult and you lack information to explain it, suggest who the user could consult (from their stakeholders list) or where to learn more.
- Highlight what is most valuable or relevant based on the user profile.
- Only simplify when necessary, and retain all informative parts.

10. CONCLUSION

This research has explored the potential and challenges of employing Large Language Models (LLMs) to enhance collaboration within complex, multi-stakeholder systems. Through comprehensive literature reviews, expert interviews, and practical experimentation in partnership with SE.Lab and the Meervijk pilot project, it has become evident that addressing complex societal issues requires innovative collaborative methodologies that transcend traditional boundaries.

Three critical themes, value alignment, communication and certainty, and power structures, emerged as foundational factors influencing collaborative success. Traditional design methodologies, though robust in simpler contexts, face significant limitations when applied to highly complex, dynamic scenarios characterized by numerous diverse stakeholders. The increased complexity, volume of information, and intricate power dynamics render traditional approaches insufficient, necessitating adaptive tools that can effectively navigate these multifaceted environments.

The application of LLM-driven tools demonstrated considerable potential to address these limitations, offering scalable, context-sensitive, and adaptive solutions. Specifically, LLMs facilitated clearer and more effective communication, streamlined the identification and alignment of stakeholder values, and contributed to more balanced power structures by democratizing access to information and participation.

However, the introduction of LLM technologies also brings ethical considerations, particularly regarding transparency, data privacy, and accountability. Addressing these ethical concerns through carefully developed guidelines and responsible implementation practices is critical for the sustainable and beneficial use of LLMs in collaborative settings.

Ultimately, this research underscores the importance of developing new models of collaboration that leverage technological advancements while remaining grounded in ethical responsibility and human-centered design principles. Future work should continue refining these LLM-driven methodologies, integrating stakeholder feedback, and scaling successful practices to broader contexts to effectively tackle complex, "wicked" societal challenges.

11. DISCUSSION AND RECOMMENDATIONS

This research aimed to explore the potential of using Large Language Models (LLMs) to develop tools that enhance collaboration in projects addressing complex or wicked problems. Through this exploration, three primary collaborative themes emerged, each accompanied by challenges associated with high complexity: value alignment, communication and certainty, and power structures. The tools created through this research, powered by LLMs, have been designed to overcome the issues that typically arise within highly complex collaborative environments.

Regarding **value alignment**, complexity creates challenges in identifying, communicating, and aligning values among stakeholders. The growing number and diversity of project participants makes traditional design methods less effective. LLM-based tools addressing role allocation support, documenting goals and values, and proposing win-win conditions can mitigate these challenges by effectively analyzing, transforming, and communicating the extensive information required for value alignment. Both facilitators and stakeholders can leverage these tools to uncover hidden values, clearly express their visions and objectives across varying languages or jargon, and effectively navigate complex and often conflicting multi-stakeholder environments.

Regarding **communication and certainty**, complexity poses significant challenges for stakeholders in maintaining a clear overview of the system, staying well-informed, and ensuring effective information dissemination. Tools designed to investigate data, transform it according to specific contexts, and clarify ambiguities help reduce informational noise within the system. By tailoring information to the context of individual recipients, these tools ensure that communication remains manageable, relevant, and informative for all stakeholders, irrespective of their diverse backgrounds.

Lastly, regarding **power structures**, complexity often results in unproductive team dynamics, creating information silos that hindered the free flow of information and establishing rigid hierarchies that limited equitable participation. Tools designed to enhance communication, mediate and resolve conflicts, and promote equitable engagement can effectively flatten hierarchical structures and mitigate counterproductive power dynamics by introducing an impartial, neutral entity to collaborative teams. The inherent neutrality of LLMs, free from personal biases or agendas, enables them to detect biases, incorrect information, and subtle cues such as hostile intonation during conversation. Including such an impartial actor throughout the collaboration significantly fosters healthier and more balanced power structures.

These initial tools highlight the significant potential of Large Language Models in addressing challenges posed by complexity. Their unique ability to efficiently manage vast amounts of data, combined with a sophisticated understanding of linguistic nuances, enables them to perform tasks swiftly and effectively. LLMs complete these tasks considerably faster than humans, significantly enhancing efficiency and allowing for rapid adjustments and iterations, critical capabilities when dealing with complex problems and systems. Nevertheless, there remain areas that require further research and development.

This research has primarily concentrated on determining the feasibility of developing these LLM powered tools and evaluating their performance across various datasets and input conditions. However, further testing in real-world contexts is essential to uncover the nuances that emerge when integrating these tools into dynamic and complex collaborative systems. Such real-world testing is likely to reveal new opportunities and allows the reconsideration of tools previously discarded based solely on theoretical assumptions about their performance. For example, tools designed to mitigate bias in decision-making were excluded from the final toolbox due to anticipated ethical concerns, particularly regarding the potential misrepresentation of certain stakeholders. Only extensive testing involving actual stakeholders can validate whether these concerns are justified or if Large Language Models can accurately and ethically represent absent groups or individuals during meetings or decision-making processes.

Additionally, certain tools such as framework builders encountered limitations due to the inherent constraints of ChatGPT. However, this does not imply that creating these tools is impossible. Rather, it highlights the need for supplementary platforms. Integrating tools like Power BI with ChatGPT inputs, for instance, could overcome these limitations. Leveraging Power BI enables users to benefit from the LLM's strengths, such as processing substantial volumes of complex, language-based data, and effectively converting this into visualizations and structured frameworks.

Complex systems and complex problems are interconnected topics that generate numerous research questions worth exploring. This research only scratches the surface of what is possible with large language models and what is needed to start developing new structural approaches that will help humanity overcome the problems we are facing. Future research should focus on developing robust, transparent frameworks for assessing and validating the outputs of LLM-driven collaborative tools. Establishing clear guidelines for ethical use and fostering ongoing stakeholder involvement will further enhance the credibility and acceptance of these technologies.

12. PERSONAL REFLECTION

This research journey concluded in a significantly different place than where it began. Initially, my goal was to assist the organization SE.Lab in developing a new transition narrative to inspire stakeholders across various domains to rethink their approaches to multi stakeholder social projects. The plan involved participating in a pilot case in Meerwijk to gain a deeper understanding of the situation. However, within the first few months, the pilot case was postponed. Despite the sudden change in plans, I was still able to gain considerable insights from my discussions with SE.Lab. This prompted me to pivot my research toward the broader theme of collaboration within complex systems, which Meerwijk is an example of.

To delve deeper, I explored literature on complex systems and wicked problems to uncover the root causes and patterns. The complexity of these topics resulted in several periods of confusion and uncertainty. Though challenging, these moments motivated me to develop my skills in asking for assistance and clearly articulating the specific problems and objectives I struggled with. Fortunately, my chair and mentor offered invaluable guidance, helping me navigate these more difficult phases of my research.

After a lot of iterations, I established a foundational framework on the topics of complexity and collaboration that enabled me to progress to the next phase: developing tools powered by Large Language Models (LLMs). This stage reignited my enthusiasm as I explored the exciting possibilities that LLMs offer. Prompting, testing, and analyzing various potential applications of LLMs in collaborative contexts proved to be highly engaging. However, attempting to pursue 20 different ideas may have been overly ambitious given the limited time available. My reluctance to discard any ideas before fully exploring their potential consumed a significant amount of time. In the future, adopting a more efficient selection process could help narrow down the most promising tools more quickly, leaving more time for in-depth development.

Overall, this project challenged me in numerous ways. Particularly challenging aspects of the research were ensuring I was improving detailed planning beyond surface-level task descriptions, clearly communicating complex information to outsiders, and maintaining effective communication with supervisors and stakeholders throughout the process. Nonetheless, these experiences have taught me invaluable lessons, which I will carry with me.

Looking forward, I am excited about continuing my journey within this research field.

REFERENCES A-D

- Abdelnabi, S., Goma, A., Sivaprasad, S., Schönherr, L., & Fritz, M. (2023, September 29). Cooperation, Competition, and Maliciousness: LLM-Stakeholders Interactive Negotiation. arXiv.org. <https://arxiv.org/abs/2309.17234>
- Albers, M. J. (2012). Communication as reducing uncertainty. Proceedings of the 30th ACM International Conference on Design of Communication, 1–8. <https://doi.org/10.1145/2379057.2379059>
- Anderson, C., & Brown, C. E. (2010). The functions and dysfunctions of hierarchy. *Research in Organizational Behavior*, 30, 55–89. <https://doi.org/10.1016/j.riob.2010.08.002>
- Aschhoff, N., & Vogel, R. (2018). Value conflicts in co-production: governing public values in multi-actor settings. *International Journal of Public Sector Management*, 31(7), 775–793. <https://doi.org/10.1108/ijpsm-08-2017-0222>
- Austin, J. E. (2000). Strategic collaboration between nonprofits and businesses. *Nonprofit and Voluntary Sector Quarterly*, 29(1), 69–97. <https://doi.org/10.1177/089976400773746346>
- Austin, J. E., & Seitanidi, M. M. (2012). Collaborative value creation. *Nonprofit and Voluntary Sector Quarterly*, 41(5), 726–758. <https://doi.org/10.1177/0899764012450777>
- Becker, S. W., & Baloff, N. (1969). Organization structure and complex problem solving. *Administrative Science Quarterly*, 14(2), 260. <https://doi.org/10.2307/2391104>
- Berger, J., Rosenholtz, S. J., & Zelditch, M. (1980). Status organizing processes. *Annual Review of Sociology*, 6(1), 479–508. <https://doi.org/10.1146/annurev.so.06.080180.002403>
- Biermann, O. C., F. N., MA, & Yoon, D. (2022). From tool to companion: storywriters want AI writers to respect their personal values and writing strategies. Designing Interactive Systems Conference. <https://doi.org/10.1145/3532106.3533506>
- Borning, A., Friedman, B., & Davis, J. (2006). Value sensitive design of interactions with urbansim indicators. <https://www.semanticscholar.org/paper/Value-sensitive-design-of-interactions-with-Borning-Friedman/aeb68760d1d5558dd2b444e847277b81b5cf52bf>
- Boschian, V. (2012). Models and Methods for Multi-Actor Systems [Doctoral Thesis, University of Trieste]. <https://www.openstarts.units.it/server/api/core/bitstreams/39e545ea-139b-4ac0-b8a7-c727eab83793/content>
- Botta, M. (2024). A Design-Stage-Oriented framework to introduce artificial intelligence and machine learning in design education. Proceedings of DRS. <https://doi.org/10.21606/drs.2024.535>
- Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B Biological Sciences*, 364(1533), 3281–3288. <https://doi.org/10.1098/rstb.2009.0134>
- Branson, C. M. (2008). Achieving organisational change through values alignment. *Journal of Educational Administration*, 46(3), 376–395. <https://doi.org/10.1108/09578230810869293>
- Brashers, D. E. (2001). Communication and uncertainty management. *Journal of Communication*, 51(3), 477–497. <https://doi.org/10.1111/j.1460-2466.2001.tb02892.x>
- Bridoux, F., & Stoelhorst, J. W. (2013). Microfoundations for stakeholder theory: Managing stakeholders with heterogeneous motives. *Strategic Management Journal*, 35(1), 107–125. <https://doi.org/10.1002/smj.2089>
- Broekema, P. M., Bulder, E. a. M., & Horlings, L. G. (2023). Evaluating co-creation in social innovation projects: Towards a process orientated framework for EU projects and beyond. *Research Evaluation*, 32(2), 286–298. <https://doi.org/10.1093/reseval/rvad017>
- Calderon, C., & Westin, M. (2019). Understanding context and its influence on collaborative planning processes: a contribution to communicative planning theory. *International Planning Studies*, 26(1), 14–27. <https://doi.org/10.1080/13563475.2019.1674639>
- Chen, C., & Shu, K. (2024). Combating misinformation in the age of LLMs: Opportunities and challenges. *AI Magazine*, 45(3), 354–368. <https://doi.org/10.1002/aaai.12188>
- Conrath, D. W. (1967). Organizational decision making behavior under varying conditions of uncertainty. *Management Science*, 13(8), B-500. <https://doi.org/10.1287/mnsc.13.8.b487>
- Cooper, J. (2013, October 29). Improving focus and predictability with critical chain project management. Project Management Institute. <https://www.pmi.org/learning/library/critical-chain-project-management-5852>
- Corominas-Murtra, B., Goñi, J., Solé, R. V., & Rodríguez-Caso, C. (2013). On the origins of hierarchy in complex networks. Proceedings of the National Academy of Sciences, 110(33), 13316–13321. <https://doi.org/10.1073/pnas.1300832110>
- Costa, C. J., Aparício, M., & Braga, A. S. (2012). Design of communication. Conference: ISDOC '12: Proceedings of the Workshop on Information Systems and Design of Communication, 62, 15–19. <https://doi.org/10.1145/2311917.2311921>
- Cui, L., Yang, K., Lei, Z., Lim, M. K., & Hou, Y. (2021). Exploring stakeholder collaboration based on the sustainability factors affecting the sharing economy. *Sustainable Production and Consumption*, 30, 218–232. <https://doi.org/10.1016/j.spc.2021.12.009>
- Curşeu, P. L., & Schrujjer, S. G. (2017). Stakeholder diversity and the comprehensiveness of sustainability decisions: the role of collaboration and conflict. *Current Opinion in Environmental Sustainability*, 28, 114–120. <https://doi.org/10.1016/j.cosust.2017.09.007>
- Dantec, C. a. L., Poole, E. S., & Wyche, S. P. (2009). Values as lived experience. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. <https://doi.org/10.1145/1518701.1518875>
- Debating innovation. (2023). In Palgrave debates in business and management. <https://doi.org/10.1007/978-3-031-16666-2>
- Diez-Busto, E., Sanchez-Ruiz, L., & Fernandez-Laviada, A. (2021). The B Corp Movement: A Systematic Literature Review. *Sustainability*, 13(5), 2508. <https://doi.org/10.3390/su13052508>
- Driskill, L., & Goldstein, J. R. (1986). Uncertainty: Theory and practice in organizational communication. *Journal of Business Communication*, 23(3), 41–56. <https://doi.org/10.1177/002194368602300304>

REFERENCES E-K

- Evenson, S., Brooker, C., Breuer, H., Dr., Jim Kalbach, & Citrix Online. (2013). Cultural change by service design: Living service Worlds - How will services know what you intend? (pp. 12–80).
- Fiset, J., Bhawe, D. P., & Jha, N. (2023). The effects of Language-Related Misunderstanding at work. *Journal of Management*, 50(1), 347–379. <https://doi.org/10.1177/01492063231181651>
- Freeman, R. E. (1984). *Strategic Management: A Stakeholder Approach*. Pitman. https://books.google.nl/books/about/Strategic_Management.html?id=4PUJAQAAMAAJ&redir_esc=y
- Friedman, B., & Hendry, D. G. (2019). Value sensitive design. In *The MIT Press eBooks*. <https://doi.org/10.7551/mitpress/7585.001.0001>
- Friedman, B., Hendry, D. G., & Borning, A. (2017). A survey of value sensitive design methods. *Foundations and Trends® in Human–Computer Interaction*, 11(2), 63–125. <https://doi.org/10.1561/11000000015>
- Friedman, B., Kahn, P. H., Borning, A., & Hultgren, A. (2013). Value sensitive design and information systems. In *Philosophy of engineering and technology* (pp. 55–95). https://doi.org/10.1007/978-94-007-7844-3_4
- Fritzsche, D. J. (1991). A Model of Decision-Making Incorporating Ethical Values. *Journal of Business Ethics*.
- Gajda, R. (2004). Utilizing collaboration theory to evaluate strategic alliances. *American Journal of Evaluation*, 25(1), 65–77. <https://doi.org/10.1177/109821400402500105>
- Geissdoerfer, M., Bocken, N. M., & Hultink, E. J. (2016). Design thinking to enhance the sustainable business modelling process – A workshop based on a value mapping process. *Journal of Cleaner Production*, 135, 1218–1232. <https://doi.org/10.1016/j.jclepro.2016.07.020>
- Getto, G., Labriola, J. T., & Ruszkiewicz, S. (2022). Content strategy. <https://doi.org/10.4324/9781003164807>
- Grant, K. P., & Gergely, M. (2012, July 18). Critical chain project management: under investigation or case closed? *Project Management Institute*. <https://www.pmi.org/learning/library/critical-chain-project-management-investigation-6380>
- Greer, L. L. (2014). Power in teams: Effects of team power structures on team conflict and team outcomes. In *Edward Elgar Publishing eBooks*. <https://doi.org/10.4337/9781781006948.00014>
- Gregory, J. & Department of Informatics, University of Oslo. (2003). Scandinavian approaches to participatory design. In *International Journal of Engineering Education* (Vol. 19, Issue 1, pp. 62–74). <https://www.ijee.ie/articles/Vol19-1/IJEE1353.pdf>
- Hadi, M. U., Tashi, Q. A., Qureshi, R., Shah, A., Muneer, A., Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J., & Mirjalili, S. (2023). A Survey on Large Language Models: Applications, Challenges, Limitations, and Practical Usage. *Authorea*. <https://doi.org/10.36227/techrxiv.23589741.v1>
- Hagemann, V., & Kluge, A. (2017). Complex problem solving in teams: The impact of collective orientation on team process demands. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01730>
- Hamdan, H. A., Andersen, P. H., & De Boer, L. (2021). Stakeholder collaboration in sustainable neighborhood projects—A review and research agenda. *Sustainable Cities and Society*, 68, 102776. <https://doi.org/10.1016/j.scs.2021.102776>
- Hara, N., Solomon, P., Kim, S., & Sonnenwald, D. H. (2003). An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, 54(10), 952–965. <https://doi.org/10.1002/asi.10291>
- Havlík, V. (2023). Meaning and understanding in large language models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2310.17407>
- Hosseini, J. C., & Brenner, S. N. (1992). The Stakeholder Theory of the Firm: a methodology to generate value matrix weights. *Business Ethics Quarterly*, 2(2), 99–119. <https://doi.org/10.2307/3857566>
- Isomursu, M., Sirotkin, A., Voltti, P., & Halonen, M. (2012). User experience design goes agile in lean transformation -- a case study. *Agile Conference*, 1–10. <https://doi.org/10.1109/agile.2012.10>
- Jiang, Y., & Ritchie, B. W. (2016). Disaster collaboration in tourism: Motives, impediments and success factors. *Journal of Hospitality and Tourism Management*, 31, 70–82. <https://doi.org/10.1016/j.jhtm.2016.09.004>
- Johansen, A., Eik-Andresen, P., Landmark, A. D., Ekambaram, A., & Rolstadås, A. (2016). Value of uncertainty: the lost opportunities in large projects. *Administrative Sciences*, 6(3), 11. <https://doi.org/10.3390/admsci6030011>
- Johansson, B. J. E., & Persson, P. (2008). Reduced uncertainty through human communication in complex environments. *Cognition Technology & Work*, 11(3), 205–214. <https://doi.org/10.1007/s10111-007-0108-6>
- Jones, C. R., & Bergen, B. K. (2024). People cannot distinguish GPT-4 from a human in a Turing test. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2405.08007>
- Kähkönen, A. (2014). The influence of power position on the depth of collaboration. *Supply Chain Management an International Journal*, 19(1), 17–30. <https://doi.org/10.1108/scm-03-2013-0079>
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and Biases*. Cambridge University Press.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263. <https://doi.org/10.2307/1914185>
- KALBACH, J., & KAHN, P. (2011). LOCATING VALUE WITH ALIGNMENT DIAGRAMS. In *PARSONS JOURNAL FOR INFORMATION MAPPING & PARSONS INSTITUTE FOR INFORMATION MAPPING, PARSONS JOURNAL FOR INFORMATION MAPPING*.
- Keeley, M. (1983). Values in organizational Theory and Management education. *Academy of Management Review*, 8(3), 376. <https://doi.org/10.2307/257826>
- Kelley, H. H. (1951). Communication in experimentally created hierarchies. *Human Relations*, 4(1), 39–56. <https://doi.org/10.1177/001872675100400102>
- Kerr, C., & Phaal, R. (2015). Visualizing Roadmaps: a Design-Driven approach. *Research-Technology Management*, 58(4), 45–54. <https://doi.org/10.5437/08956308x5804253>

REFERENCES K-O

- Khademian, E. P. W. a. a. M. (2008). Wicked problems, knowledge challenges, and collaborative capacity builders in network settings. *Public Administration Review*, 68(2), 334–349. <https://www.jstor.org/stable/25145606>
- Kim, E., Beckman, S. L., & Agogino, A. (2018). Design Roadmapping in an uncertain world: Implementing a Customer-Experience-Focused strategy. *California Management Review*, 61(1), 43–70. <https://doi.org/10.1177/0008125618796489>
- Kim, E., Simonse, L. W. L., Beckman, S. L., Appleyard, M. M., Velazquez, H., Madrigal, A. S., & Agogino, A. M. (2020). User-Centered Design Roadmapping: Anchoring roadmapping in customer value before technology selection. *IEEE Transactions on Engineering Management*, 69(1), 109–126. <https://doi.org/10.1109/tem.2020.3030172>
- Kish-Gephart, J. J., Detert, J. R., Treviño, L. K., & Edmondson, A. C. (2009). Silenced by fear: The nature, sources, and consequences of fear at work. *Research in Organizational Behavior*, 29, 163–193. <https://doi.org/10.1016/j.riob.2009.07.002>
- Kpamma, Z. E., Adjei-Kumi, T., Ayarkwa, J., & Adinyira, E. (2017). Participatory design, wicked problems, choosing by advantages. *Engineering Construction & Architectural Management*, 24(2), 289–307. <https://doi.org/10.1108/ecam-06-2015-0085>
- Kumar, V., & Reinartz, W. (2018). Applications of CRM in B2B and B2C scenarios Part I. In *Springer texts in business and economics* (pp. 329–362). https://doi.org/10.1007/978-3-662-55381-7_16
- Kwawu, W., & Hughes, W. (2008). Strategies for aligning organizational incentive systems through contracts. <https://www.semanticscholar.org/paper/Strategies-for-aligning-organizational-incentive-Kwawu-Hughes/c5e5c56ba616cf12dddf13c97490bc786600e0e8#citing-papers>
- Lahiri, A., Pahnke, E. C., Howard, M. D., & Boeker, W. (2019). Collaboration and informal hierarchy in innovation teams: Product introductions in entrepreneurial ventures. *Strategic Entrepreneurship Journal*, 13(3), 326–358. <https://doi.org/10.1002/sej.1331>
- Lee, S., Jang, K. J., Lee, M. H., & Shin, S. R. (2021). Roadmapping in the Era of Uncertainty: How to Integrate Data-Driven Methods with Expert Insights. *Foresight-Russia*, 15(2), 39–51. <https://doi.org/10.17323/2500-2597.2021.2.39.51>
- Legacy, C. (2010). Investigating the Knowledge interface between Stakeholder Engagement and Plan-Making. *Environment and Planning a Economy and Space*, 42(11), 2705–2720. <https://doi.org/10.1068/a43164>
- Letaifa, S. B. (2014). The uneasy transition from supply chains to ecosystems. *Management Decision*, 52(2), 278–295. <https://doi.org/10.1108/md-06-2013-0329>
- Li, J., Cao, H., Lin, L., Hou, Y., Zhu, R., & Ali, A. E. (2024). User Experience Design Professionals' Perceptions of Generative Artificial Intelligence. *CHI '24: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 35, 1–18. <https://doi.org/10.1145/3613904.3642114>
- Lin, Y., Frey, C. B., & Wu, L. (2022). Remote collaboration fuses fewer breakthrough ideas. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2206.01878>
- LLM Leaderboard | Compare Top AI Models for 2024. (n.d.). YourGPT. <https://yourgpt.ai/tools/llm-comparison-and-leaderboard>
- Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376727>
- Martín-Rodríguez, L. S., Beaulieu, M., D'Amour, D., & Ferrada-Videla, M. (2005). The determinants of successful collaboration: A review of theoretical and empirical studies. *Journal of Interprofessional Care*, 19(sup1), 132–147. <https://doi.org/10.1080/13561820500082677>
- Meincke, L., Mollick, E. R., & Terwiesch, C. (2024). Prompting diverse ideas: Increasing AI idea variance. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2402.01727>
- Mendonça, J., Lavie, A., & Trancoso, I. (2024). On the Benchmarking of LLMs for Open-Domain Dialogue Evaluation. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2407.03841>
- Milliken, F. J., Morrison, E. W., & Hewlin, P. F. (2003). An Exploratory Study of Employee Silence: Issues that Employees Don't Communicate Upward and Why*. *Journal of Management Studies*, 40(6), 1453–1476. <https://doi.org/10.1111/1467-6486.00387>
- Mitchell, R. K., & Lee, J. H. (2019). Stakeholder Identification and Its Importance in the Value Creating System of Stakeholder Work. In *Cambridge University Press eBooks* (pp. 53–74). <https://doi.org/10.1017/9781108123495.004>
- Moesgen, T., Salovaara, A., Epp, F. A., & Sanchez, C. (2023). Designing for Uncertain Futures: An Anticipatory approach. *Interactions*, 30(6), 36–41. <https://doi.org/10.1145/3624698>
- Murphy, M., Arenas, D., & Batista, J. M. (2014). Value Creation in Cross-Sector Collaborations: the roles of experience and alignment. *Journal of Business Ethics*, 130(1), 145–162. <https://doi.org/10.1007/s10551-014-2204-x>
- Mysore, K., Kirytopoulos, K., Ahn, S., & Ma, T. (2019). Adversarial factors in multi-stakeholders' engagement of global-IT projects. *International Journal of Managing Projects in Business*, 14(2), 445–471. <https://doi.org/10.1108/ijmpb-01-2019-0014>
- Neville, B. A., & Menguc, B. (2006). Stakeholder Multiplicity: Toward an Understanding of the Interactions between Stakeholders. *Journal of Business Ethics*, 66(4), 377–391. <https://doi.org/10.1007/s10551-006-0015-4>
- Nordheim, C. B., Følstad, A., & Bjørkli, C. A. (2019). An Initial Model of Trust in Chatbots for Customer Service—Findings from a Questionnaire Study. *Interacting With Computers*, 31(3), 317–335. <https://doi.org/10.1093/iwc/iwz022>
- Norrman, A., & Naslund, D. (2019). Supply Chain Incentive Alignment: The Gap between Perceived Importance and Actual Practice. *Operations and Supply Chain Management an International Journal*, 129–142. <https://doi.org/10.31387/oscm0380237>
- Noyes, A. L. (2021). Navigating the hierarchy: communicating power relationships in collaborative health care groups. *Management Communication Quarterly*, 36(1), 62–91. <https://doi.org/10.1177/08933189211025737>
- Ooi, Y. M., & Husted, K. (2021). Framing Multi-Stakeholder Value Propositions: A wicked problem lens. *TIM Review*. <https://timreview.ca/article/1434>

REFERENCES P-V

- P Sycara, K. (1998, June 1). Multiagent Systems. Wiley Online Library. <https://onlinelibrary.wiley.com/doi/full/10.1609/aimag.v19i2.1370>
- Pereno, A., Eriksson, D., Department of Architecture and Design, Politecnico di Torino, & Stiftelsen TEM, Nordic Center for Sustainable Healthcare. (2020). A multi-stakeholder perspective on sustainable healthcare: From 2030 onwards. In *Futures* (p. 102605) [Journal-article]. <https://doi.org/10.1016/j.futures.2020.102605>
- Phaal, R., Farrukh, C. J., & Probert, D. R. (2003). Technology roadmapping—A planning framework for evolution and revolution. *Technological Forecasting and Social Change*, 71(1–2), 5–26. [https://doi.org/10.1016/s0040-1625\(03\)00072-6](https://doi.org/10.1016/s0040-1625(03)00072-6)
- Phaal, R., & Muller, G. (2008). An architectural framework for roadmapping: Towards visual strategy. *Technological Forecasting and Social Change*, 76(1), 39–49. <https://doi.org/10.1016/j.techfore.2008.03.018>
- Randmaa, M., Howard, T. J., & Otto, T. (2012). From product centered design to value centered design: understanding the value-system. Welcome to DTU Research Database. <https://orbit.dtu.dk/en/publications/from-product-centered-design-to-value-centered-design-understandi>
- Reber, M., & Duffy, A. (2005). Value Centred Design: Understanding the nature of value. In *INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN ICED 05 MELBOURNE* [Conference-proceeding].
- Roberts, N. (2000). Wicked problems and network approaches to resolution. <https://ipmr.net/index.php/ipmr/article/view/175>
- Roberts, N. C. (2001). Wicked problems and network approaches to resolution. In *ResearchGate*. <https://www.researchgate.net/publication/235700723>
- Roose, K. (2024, October). Can A.I. Be Blamed for a Teen's Suicide? The New York Times. <https://www.nytimes.com/2024/10/23/technology/characterai-lawsuit-teen-suicide.html>
- Rousseau, D. M. (2018). Making evidence-based organizational decisions in an uncertain world. *Organizational Dynamics*, 47(3), 135–146. <https://doi.org/10.1016/j.orgdyn.2018.05.001>
- Rybníček, R., & Königsgruber, R. (2018). What makes industry–university collaboration succeed? A systematic review of the literature. *Journal of Business Economics*, 89(2), 221–250. <https://doi.org/10.1007/s11573-018-0916-6>
- Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). A Systematic survey of prompt engineering in large language Models: Techniques and applications. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2402.07927>
- Salvi, F., Ribeiro, M. H., Gallotti, R., & West, R. (2024). On the Conversational Persuasiveness of Large Language Models: A Randomized Controlled Trial. *Research Square (Research Square)*. <https://doi.org/10.21203/rs.3.rs-4429707/v1>
- Schwaber, K., & Beedle, M. (2002). *Agile Software Development with Scrum*. Pearson.
- Shaikh, O., Chai, V., Gelfand, M. J., Yang, D., & Bernstein, M. S. (2023). Rehearsal: Simulating conflict to teach Conflict resolution. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2309.12309>
- Shapiro, D. (2005). Participatory design. *Proceedings of the 4th Decennial Conference on Critical Computing*, 29–38. <https://doi.org/10.1145/1094562.1094567>
- Shneiderman, B. (2020). Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy. *International Journal of Human-Computer Interaction*, 36(6), 495–504. <https://doi.org/10.1080/10447318.2020.1741118>
- Solaiman, I., Talat, Z., Agnew, W., Ahmad, L., Baker, D., Blodgett, S. L., Hal, D. I., Dodge, J., Evans, E., Hooker, S., Jernite, Y., Luccioni, A. S., Lusoli, A., Mitchell, M., Newman, J., Png, M., Strait, A., & Vassilev, A. (2023). Evaluating the social impact of generative AI systems in systems and society. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2306.05949>
- Strachwitz, R. G., Alter, R., & Unger, T. (2021). Addressing wicked problems. *Collaboration, trust and the role of shared principles at the philanthropy and government interface. Trusts & Trustees*, 27(6), 450–459. <https://doi.org/10.1093/tandt/ttab066>
- Street, W., Siy, J. O., Keeling, G., Baranes, A., Barnett, B., McKibben, M., Kanyere, T., Lentz, A., Arcas, B. a. Y., & Dunbar, R. I. M. (2024). LLMs achieve adult human performance on higher-order theory of mind tasks. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2405.18870>
- Sunnafrank, M. (1990). Predicted Outcome value and uncertainty reduction theories A test of competing perspectives. *Human Communication Research*, 17(1), 76–103. <https://doi.org/10.1111/j.1468-2958.1990.tb00227.x>
- Tang, L., & Shen, Q. (2012). Factors affecting effectiveness and efficiency of analyzing stakeholders' needs at the briefing stage of public private partnership projects. *International Journal of Project Management*, 31(4), 513–521. <https://doi.org/10.1016/j.ijproman.2012.10.010>
- Tenzer, H., Pudelko, M., & Harzing, A. (2013). The impact of language barriers on trust formation in multinational teams. *Journal of International Business Studies*, 45(5), 508–535. <https://doi.org/10.1057/jibs.2013.64>
- Thomsen, S. (2004). Corporate values and corporate governance. *Corporate Governance*, 4(4), 29–46. <https://doi.org/10.1108/14720700410558862>
- Tobey, D. H., & Perera, B. Y. (2012). Corporate social responsibility initiatives. *African Journal of Economic and Management Studies*, 3(1), 95–115. <https://doi.org/10.1108/20400701211197302>
- Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., & Herrmann, E. (2012). Two key steps in the evolution of human cooperation. *Current Anthropology*, 53(6), 673–692. <https://doi.org/10.1086/668207>
- Valk, A., & Kratovič, M. (2021). We collaborate with everyone, but with some more than others: evidence of stakeholder collaboration among internal security professional higher education institutions. *Empirical Research in Vocational Education and Training*, 13(1). <https://doi.org/10.1186/s40461-021-00110-6>
- Vangen, S. (2016). Developing Practice-Oriented Theory on Collaboration: A Paradox Lens. *Public Administration Review*, 77(2), 263–272. <https://doi.org/10.1111/puar.12683>

REFERENCES W-Z

- Waddell, P. (2011). Integrated Land Use and Transportation Planning and Modelling: Addressing challenges in research and practice. *Transport Reviews*, 31(2), 209–229. <https://doi.org/10.1080/01441647.2010.525671>
- Weber, E. P., & Khademian, A. M. (2008). Wicked problems, knowledge challenges, and collaborative capacity builders in network settings. *Public Administration Review*, 68(2), 334–349. <https://doi.org/10.1111/j.1540-6210.2007.00866.x>
- Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., Kenton, Z., Brown, S., Hawkins, W., Stepleton, T., Biles, C., Birhane, A., Haas, J., Rimell, L., Hendricks, L. A., . . . Gabriel, I. (2021). Ethical and social risks of harm from Language Models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2112.04359>
- Westermann, H., Savelka, J., & Benyekhlef, K. (2023, July 27). LLMediator: GPT-4 assisted online dispute resolution. *arXiv.org*. <https://arxiv.org/abs/2307.16732>
- Wikipedia contributors. (2024, December 14). Probabilistic risk assessment. *Wikipedia*. https://en.wikipedia.org/wiki/Probabilistic_risk_assessment
- Xu, H., Bu, Y., Liu, M., Zhang, C., Sun, M., Zhang, Y., Meyer, E., Salas, E., & Ding, Y. (2021). Team Power Dynamics and Team Impact: New Perspectives on Scientific Collaboration using Career Age as a Proxy for Team Power. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2108.04108>
- Yaegashi, K., Valušytė, R., Biamonti, A., Anzai, H., & Minamiura, S. (2019). Meanings of design in the next era [Conference Proceedings]. In DML (Design Management Lab), Ritsumeikan University (Ed.), *Conference Proceedings*. 4d-conference.com. <https://en.ritsumei.ac.jp>
- Zhang, Z. (2023). Designing for explanation-driven trust in Chatbots. <http://resolver.tudelft.nl/uuid:99884f56-3956-4f0d-9e95-95168f6716ba>
- Zuber, N., Kacianka, S., Gogoll, J., Pretschner, A., & Nida-Rümelin, J. (2021). Empowered and embedded: ethics and agile processes. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2107.07249>

APPENDIX A

Interview

Introduction:

Introduce myself and inform them that I am conducting thesis research on collaboration between stakeholders. Ask for permission to record the interview and explain the consent form before sending it.

Interview Questions

- Can you introduce yourself and your position within your company/organization?
 - What kinds of collaborations do you encounter in your work? / Which other parties do you collaborate with in your work?
 - Does your organization use any strategies or methods to facilitate collaboration?
 - What kind of obstacles do you experience during collaborations? / What prevents successful collaboration?
- >If more information is needed:
- Can you recall a recent collaborative project and describe what made collaboration with other parties challenging?
- >If not already mentioned, introduce the themes of Value, Communication, and Power Structures:
- Value: Does your company address value during collaboration—both the values important to your company and the value to be achieved?
 - Communication: How is communication managed during collaboration? What happens when there is uncertainty about part of the project?
 - Power Structures: Are there any power structures within your company or during collaborative projects? How do they impact collaboration?
- Are you familiar with Large Language Models (LLMs)?
>Explain if needed.
- >Explain some of the strengths of LLMs.
- Based on what I just explained about LLM capabilities, where in your organization do you think LLMs could be beneficial?
 - Considering the potential of LLMs, what ethical considerations do you think are important? / What risks should we not overlook?

APPENDIX B

Stakeholders have difficulty collaborating with each other

Marina - "People from very different disciplines need to collaborate, and these individuals often struggle to understand each other. You speak different languages, use different methods, and have different perspectives on life. Since you haven't worked together before, it's challenging to develop mutual understanding. Achieving that understanding takes time; it doesn't happen immediately after the first meeting."

Marina - "At the start, everything remains quite broad and high-level, so everyone can generally agree. But as the project progresses, you begin to uncover misunderstandings and realize where assumptions have been made. Often, these assumptions are never explicitly addressed and linger in the system for a long time until a conflict arises that makes it clear some stakeholders are not aligned."

Marina - "Even after two years into a project, it can become apparent that stakeholders have been working past each other. You might discover that different stakeholders have fundamentally different perspectives, but these differences remain unspoken because they are unclear and hard to pinpoint."

Sabine - "The main obstacles that have prevented this (Reforms in the healthcare sector) from happening so far are because people generally find it difficult to work together with people outside of their domain."

Sabine - "You have the policymakers or management, and then you have the Parkinson's caretakers. The Parkinson's caretakers have their own meetings every few months. In these meetings, we discuss who should take care of which patients. So, we are already communicating. The higher-ups (policymakers) are still lacking in communication. I think they believe that mistakes happen on our side, and yes, they do, but we communicate and fix them. The real problem lies with them."

Sabine - "Proper transferal is everything (very important), handing of responsibility properly and smoothly. Also knowing each other, investing in each others domain and field of work. People need to communicate better about what they do and explore what others are doing so there is more central understanding, especially from higher up, they should really spend more time talking to us about what's happening."

Leonoor - We increasingly collaborate through co-creation, bringing all stakeholders to the table. We use models such as roadmaps, allowing us to consider a broader context from the very beginning. One example is the digitalization of rehabilitation, enabling elderly patients to carry out more rehabilitation exercises at home.

Leonoor - Collaborating on a larger scale is challenging because larger systems often do not communicate with each other. For example, cooperation between different home care organizations can be difficult. Within hospitals, it is somewhat easier since the system is less fragmented.

APPENDIX B

Irene - Family support can be highly fragmented. In some cases, as many as 20 different professionals may be involved with a single family, each handling their own responsibilities without coordination or awareness of each other's efforts.

Strong hierarchy disrupts collaboration in multi stakeholders systems

marina - "In collaborations, there are often certain parties that hold all the power. In discussions, they are always very open, but when push comes to shove, they are the ones who make the final decision. Some choices simply aren't made because certain parties know that a dominant party won't participate if those options are chosen, which leads to others participants dropping out as well."

Sabine - "I feel like there is a lot of knowledge among the caretakers but these people in the top don't use it. I feel like they use too much of a top down approach. They give commands instead of collaborating. Hierarchy still plays a large role. Another thing that prevents proper collaboration is that everyone has opinions about people and organizations in the healthcare sector and those opinions influence their decision making."

Hans - "Large companies are often capable of acquiring innovation and integrating it into their production lines. Typically, when you approach a large company with an idea, they will look at how it can be adapted to fit within their production line, rather than the other way around (How their production line might adjust to the new idea)."

"When you want to innovate and you need these big companies to join you, you sometimes don't really have a choice and need to listen to them"

Marguerite - "Stakeholders often feel like participants. However, that doesn't work very well because it gives people the idea that they can step in and out of the process, whereas they should actually feel like a core part of the process."

When providing training to employees, we frequently encountered resistance. This resistance likely stems from the top-down nature of decision-making, where employees feel unheard in their day-to-day work. Additionally, they face conflicting demands—being pushed for both quality and quantity—forcing them to complete numerous tasks in a short timeframe. In some organizations, it is also difficult to admit struggles or fears. (Irene)

APPENDIX B

Stakeholders are hesitant to fully commit to a project due to risk and uncertainty

marina - "Many for-profit companies want a guarantee upfront that they will be able to derive value from a project"

"Primarily, commercial parties want to quickly see how a project can generate value for them. Since they need to make a profit, there has to be something in return."

Hans - "Uncertainty and risk play a significant role in these types of collaborations. There is often a large difference between stakeholders in terms of how much risk they are willing to take. The question is how to mitigate risk for those who find it difficult to handle."

"Some people are much more willing to grant someone an opportunity, while others are not at all—they want to know right away how much you can pay. This varies greatly from person to person. Some really important questions you need to ask yourself are: How can I quickly identify whether you have a good connection with someone, or how can you create that connection? How do you identify why someone is sitting at the table with you: what kind of structure does their company have, how are they representing their boss, and are they here to help or to resist?"

Gina - "Money often plays a significant role when there is uncertainty involved. In such cases, people often want to avoid risks."

"When the stakes are higher, you need to focus on how risks can be avoided. For example, student projects are often so focused on survival that more experimental endeavours don't really take off because the risks are simply too high."

"As long as all interests are safeguarded, stakeholders are much more willing to take risks. However, this needs to be clearly communicated in advance."

Marguerite - "These organisations don't fully involve themselves in the project because you can't immediately promise them an outcome. They send someone that participates but it's just them, not the whole organisation."

Stakeholder don't properly address value during the project

marina - "Some companies do start with value identification at the beginning of the process, but this certainly doesn't happen everywhere. Its say most often it doesn't happen"

"Different stakeholders naturally have varying perspectives on what constitutes value. For example, sustainability or autonomy can be interpreted in many different ways."

"Values between the individual and the company they represent can differ. The person participating might fully support transparency, but the company they represent may want to protect their intellectual property and completely disagree with this stance."

APPENDIX B

Sabine - "The project that is now starting at 'Kennis centrum parkinson huis' has a more bottom up approach so that helps with value exchange. Generally speaking though, every organization has such different values and norms and this makes collaboration difficult."

Hans - "In my work sector, there isn't yet a specific method or approach to navigate the different values of stakeholders. We are currently exploring joint value creation as a way to safeguard stakeholders' values."

"Currently, you often end up with a sort of zero-sum game where, if I give something to someone, I don't immediately see how I get something in return. This creates a race to the bottom, which I don't believe in at all. In these kinds of complex situations, you really need to focus on: Where does our shared value lie? But also, what are our conflicting values? You need to map these out, and it has to happen as early as possible."

Gina - "In my experience, people are always enthusiastic about participating, but there's a lot involved. Everyone is, in theory, always looking for shared value, but often they don't fully understand how the collaboration will work."

"People often want to step into a meeting and just "bam, bam, bam," discuss things as they usually would, and then topics like values or joint value creation aren't addressed at all."

"If you don't clearly define from the start how each stakeholder derives value from the project, you end up with a kind of expectation management throughout the project, which eventually results in the project falling apart."

"There have been projects that seemed to be going very well, but because it wasn't discussed what value would be created, no one ended up taking the lead and it eventually fizzled out."

Marguerite - "The reason why this value wasn't apparent at first is because this solution was very far removed from the healthcare domain so it's not within their field of expertise. It's difficult for them to see the value during a project because they're not very open to it"

We use methodologies that already incorporate core values. For example, Signs of Safety and solution-focused work are built on foundational values, which also serve as the criteria for measuring success. However, the organizations we worked with often had their own values, which sometimes led to conflicts.(Irene)

Organizations tend to prioritize cost-effectiveness, while we aim to provide the highest possible quality. (Irene)

Value conflicts most often arise at the beginning. During initial discussions, everyone's needs and priorities must be aligned, requiring strong mediation skills. (Irene)

APPENDIX (C)

Cultural Values and Sensitive Content

AI systems must navigate varying cultural norms and definitions of sensitive content, such as hate speech or graphic material. Norms differ by region, and generative AI cannot be culturally neutral. Evaluations often involve geopolitical, ethical, and social value assessments, yet many focus narrowly on dominant cultures. Limitations arise from the overrepresentation of certain values and biases in cultural contexts, which may neglect marginalized communities

Disparate Performance

Disparate performance refers to unequal system performance across subpopulations, often caused by skewed data representation and feature inconsistencies. These disparities are exacerbated in multilingual or low-resource settings. Evaluations include subgroup performance analysis, accuracy, and disaggregated results. Limitations include data sparsity and challenges in defining meaningful performance metrics for underrepresented groups

Environmental Costs and Carbon Emissions

Generative AI systems consume significant energy during training, inference, and deployment, contributing to carbon emissions. Evaluation tools like CodeCarbon and Carbontracker measure energy usage and emissions, but there is no consensus on comprehensive metrics. Challenges include accounting for indirect factors like supply chains and lack of transparency from hardware manufacturers

Privacy and Data Protection

AI systems often process personal data, leading to risks of privacy violations and unintended data leakage. Evaluations examine memorization, inference of personal attributes, and data leakage risks, while mitigation efforts include data minimization, consent mechanisms, and opt-in approaches. Limitations stem from the difficulty of ensuring robust protections and addressing context-specific privacy needs

Financial Costs

The financial burden of AI system development includes data acquisition, compute infrastructure, and labor costs, restricting access for less-resourced groups. Evaluations track infrastructure costs, labor hours, and model hosting expenses. Challenges include accounting for hidden costs, such as those related to data cleaning and post-deployment adjustments

Data and Content Moderation Labor

Human labor underpins tasks like data curation, moderation, and evaluation. Many workers face low pay and exposure to harmful content without adequate psychological support. Evaluations focus on working conditions and adherence to ethical standards. Limitations arise from a lack of transparency, documentation, and regulation around the use of crowd labor

APPENDIX (C)

Trust in Media and Information

Generative AI systems contribute to the erosion of trust in media and information due to their ability to produce convincing misinformation and disinformation. These systems often generate content indistinguishable from human-created material, complicating detection and reducing trust in credible sources. Tools like watermarking and transparency measures can mitigate these issues, but challenges remain due to the sophistication of generative technologies

Over Reliance on Outputs

Humans often overtrust AI outputs due to perceived authority or urgency in decision-making. This overreliance can lead to the spread of inaccuracies and biases, exacerbated by AI's vulnerabilities like hallucinations and deceptive behaviors. Interventions such as user education and robust testing are critical to mitigate these risks

Personal Privacy and Sense of Self

AI systems can infringe on privacy by exposing personal or sensitive information through training data leaks or unintended memorization. This violation extends beyond data privacy to issues of autonomy, as individuals may lose control over their personal narratives. Mitigations include stronger privacy protections, such as opt-in data use and regulatory frameworks like GDPR.

Community Erasure

Generative AI systems can unintentionally erase marginalized communities through biases in training data and content moderation strategies. Automated or human moderation often suppresses non-dominant cultural expressions, leading to reduced visibility and representation. Transparent, inclusive moderation policies and representation in training datasets are vital for mitigation

Long-term Amplification and Embedding of Marginalization by Exclusion (and Inclusion)

AI systems can perpetuate marginalization by excluding underrepresented groups from datasets or by including their data without consent, leading to exploitation. This issue is particularly pronounced in applications like low-resource language modeling or cultural representation. Engagement with affected communities and tailored design are recommended solutions

Abusive or Violent Content

Generative AI systems can produce or enable the generation of abusive or violent content, such as non-consensual imagery or hate speech. This disproportionately affects marginalized groups and perpetuates harm. Mitigation strategies include robust content filtering, dataset auditing, and regulatory oversight

APPENDIX (C)

Militarization, Surveillance, and Weaponization

The power to develop and deploy advanced generative AI is concentrated among a few entities, often leading to misuse for surveillance, military purposes, or cyberattacks. Ethical guidelines and international agreements are essential to curb these risks, ensuring AI systems are not weaponized or used for oppressive surveillance

Imposing Norms and Values

Global deployment of generative AI often reflects the dominant cultural values of the developers, imposing these norms on diverse cultures. This can result in homogenization, suppression of cultural differences, and marginalization of non-dominant languages and practices. Solutions include culturally sensitive design and inclusion of underrepresented communities in system development

Intellectual Property and Ownership

Generative AI raises significant concerns over intellectual property (IP) as it can reproduce copyrighted content or create new works without clear ownership frameworks. These issues require robust IP policies, attribution mechanisms, and legal clarifications to protect creators' rights

Economy and Labor Market

AI's ability to automate creative and routine tasks disrupts labor markets, potentially displacing workers and altering skill requirements. Evaluations should consider both the economic opportunities and threats posed by AI to ensure equitable outcomes, such as through workforce reskilling initiatives

Widening Resource Gaps

Generative AI systems often require extensive computational resources, exacerbating disparities between well-resourced organizations and those with limited access. This creates barriers for equitable participation in AI development and use. Policies to democratize access and reduce barriers are necessary to address these inequities

Bias, Stereotypes, and Representational Harms

Generative AI systems can embed and amplify harmful biases, influenced by the development chain, data choices, and modeling techniques. These biases often target marginalized identities and can reinforce stereotypes. Evaluations typically focus on bias detection through intrinsic (model-focused) and extrinsic (output-focused) methods, addressing co-occurrence, sentiment, and toxic language analysis. However, limitations exist in addressing intersectionality and adapting to evolving contexts, leading to underrepresentation of certain cultural and demographic groups.

APPENDIX (D)

Zero-Shot Prompting

Eliciting specific outputs from the model without providing prior examples, relying solely on the model's pre-trained knowledge. This technique is particularly useful for general tasks where the model's existing knowledge base is sufficient.

Few-Shot Prompting

Providing a limited number of examples to guide the model's output, thereby improving the accuracy and relevance of its responses. This approach effectively fine-tunes the model's output for specific contexts or specialized tasks.

Chain-of-Thought Prompting

Encouraging the model to break down its reasoning into logical steps before providing an answer, enhancing performance on complex tasks. This technique is especially valuable for problem-solving scenarios and explicating decision-making processes.

Instruction-Based Prompting

Providing clear, explicit instructions within the prompt to define the task and reduce ambiguity. This method ensures the model comprehends exactly what is expected, leading to more targeted and relevant outputs.

Contextual Prompting

Supplying relevant background information or specific context within the prompt, enabling the model to generate more coherent and tailored outputs. This approach is crucial for domain-specific tasks and enhancing the relevance of generated content.

Iterative Refinement

Fine-tuning the prompt through incremental adjustments to the wording or structure to achieve the best possible output. This process involves systematic trial and error to optimize results and adapt to specific requirements.

Role-Playing

Assigning the AI model a specific role or persona to guide its responses and maintain a consistent tone and perspective. This technique can be particularly useful for generating content from different viewpoints or simulating specific expertise.

Multi-Task Prompting

Instructing the model to perform multiple tasks or process multiple instructions within a single prompt. This approach increases efficiency when dealing with complex queries or interconnected tasks.

Priming

Providing the model with key terms, phrases, or examples that influence its output towards a specific style, topic, or sentiment. This technique helps steer the model's response in a desired direction, enhancing contextual relevance.

APPENDIX (D)

Creativity Prompting

Guiding the model's level of creativity and diversity in its outputs through specific prompt wording. This can encourage more imaginative or unconventional thinking for brainstorming tasks, or conversely, promote more focused and conventional outputs for technical applications.

Length Control

Specifying the desired length of the model's output by setting a maximum or minimum number of tokens or characters. This ensures responses fit within specific constraints, which is particularly useful for generating content with strict length requirements.

Anchor Prompting

Using fixed phrases or patterns within the prompt to guide the structure and content of the model's output. This helps maintain consistency across multiple outputs and can be especially useful for generating structured content.

Contrast Prompting

Providing the model with contrasting examples or scenarios to help it distinguish between desired and undesired outputs. This technique is useful for refining the model's understanding of specific concepts and generating more nuanced responses.

Task-Specific Prompting

Tailoring the prompt to the specific requirements and nuances of the task at hand, such as summarization, question-answering, or creative writing. This optimizes the prompt for particular use cases, enhancing the relevance and quality of the output.

Feedback-Based Prompting

Incorporating feedback from users or subject matter experts to refine the prompt and improve the model's performance over time. This iterative process helps fine-tune outputs for specific audiences or applications, ensuring continual improvement and adaptation.

Negative Prompting

Telling the model specifically what not to do in order to still give it a large amount of freedom whilst removing the undesired results.

APPENDIX (E)

1. *Creating timetables and schedules*

When supplied with the necessary information, LLMs could generate comprehensive timetables and schedules for all stakeholders, ensuring key dates are highlighted and availability overlaps are efficiently managed.

- **Power structure:** Using automated systems means that no single stakeholder can strongarm their preferred structure or choices. A system that only looks at the factual information decides on the best structure
- **Communication & Certainty:** having clear structure and an overview of what stakeholders can expect can reduce uncertainty and risk.
- **Data input:** The tool can process structured data sources such as Excel files, Google Calendar, and Google Sheets while also extracting relevant information from unstructured data sources like meeting transcripts, emails, or large text files. It can synthesize this information into organized schedules and planners.
- **Prompt structures used:** Instruction-Based Prompting
- **How well does the tool perform:** The tool is able to easily and quickly create timetables from both structured, unstructured data or a combination of the two. It can output these schedules into various formats including Google Sheets which would allow for simple updates to existing scheduling documents.
- **Ethics and Risk:** The primary risk is over-reliance on LLM outputs (Low Impact). While the tool effectively organizes schedules, errors may still occur, though their impact will be minimal. Human oversight is necessary to validate and adjust outputs, ensuring accuracy and adaptability to last-minute changes.

2. *Role Allocation Support*

LLMs can suggest roles or responsibilities for stakeholders based on their stated interests and expertise, fostering co-ownership.

- **Power structures:** An LLM can argue for the best position for each parties without having a agenda behind it, basing it decision purely on factual info which might allow smaller stakeholder to take up roles that they would otherwise not have been assigned.
- **Communication & Certainty:** LLM can produce clear argumentation tailored to each stakeholder to explain why roles should be divided in a certain way.
- **Data input:** The tool can use pre structured data such as forms or excel sheets containing specific information per stakeholder but is also capable of making role estimations based on unstructured data such as transcripts or even websites of stakeholders can be used to assess their strengths.
- **Prompt structures used:** Instruction-Based Prompting
- **How well does the tool perform:** When provided with links to stakeholders websites, the tool is able to provide a clear overview of what it thinks the stakeholders capabilities and strong points are. When provided with a project description it can then easily and quickly identify which tasks and roles in the project should be assigned to which stakeholder giving clear reasoning behind its decisions.
- **Ethics and Risk:** The primary risk is Bias, Stereotypes, and Representational Harms (High Impact), the tool may significantly influence decision-making, especially when trained on biased data. To mitigate these risks, it is essential to implement safeguards that ensure transparency in the tool's use, operation, and reasoning. All stakeholders should have clear visibility into how decisions are made. Any advice provided by the tool should be explicitly framed as guidance rather than absolute directives.

3. Trend Predictions

LLM can be used to generate trend predictions based on raw data provided by the user. This can help stakeholders with their decision making process as it gives them an idea of what might happen when certain options are picked.

- **Value Alignment:** By creating possible future scenarios it becomes far easier to understand and identify how and when value will be achieved for the various stakeholders involved
- **Data input:** The tool can use both structured and unstructured data provided by the user to use in its trend predictions. Additionally, it can scan the internet for additional information to reinforce the provided data set or replace it all together.
- **Prompt structures used:** Instruction-Based Prompting, Task-Specific Prompting
- **How well does the tool perform:** The tool can make predictions based on public data and some of these can be insightful but most of them are surface level insights. The real value comes when specific datasets are provided to the model that it can use together with publicly available data. This does introduce some ethics concerns though.
- **Ethics and Risk:** The main risks are Over Reliance on LLM Outputs (Low Impact) and Privacy and Data Protection (High Impact). The tool will be used by humans so there is little danger of LLM mistakes entering the system without supervision of humans. Users do however need to consider and be aware of the fact that any information they upload into the tool will be shared with and used by OpenAI to further improve their models. Sensitive data should thus not be used with tools like these.

4. Creating Frameworks

When given the appropriate structure, data, and prompts, LLMs can automate the creation of design frameworks and integrate stakeholder data. This not only reduces the workload for organizations involved but can also help minimize bias, hierarchical barriers and enhance stakeholder trust and engagement.

- **Communication & Certainty:** due to reduced time for the feedback loop stakeholders will be actively held up to date which can give them a greater sense of reassurance that their actions are making a difference.
- **Value alignment:** By reducing the time for feedback loops and increasing the depth of the results, value achieved becomes far easier to identify.
- **Data input:** The tool can identify useful information for specific frameworks from both structured and unstructured data sets. It can also extract text from visual sources to a limited capacity. Finally, it is able to identify specific frameworks from visual references.
- **Prompt structures used:** Instruction-Based Prompting, Chain-of-Thought Prompting Few-Shot Prompting, Task-Specific Prompting, Negative Prompting
- **How well does the tool perform:** LLM's ability to create frameworks strongly depends on what the framework looks like. It is unable to make and fill in frameworks that have complex two dimensional value structures such as what is required for service blue prints or problem trees. Simpler visual structures such as Value matrixes or the Tobey and Perera (2012) value quadrant chart are however possible and LLM can successfully place the right information in the right location inside the visual and argue why it has done this. For LLM capabilities with more complex frameworks further investigation is needed into the use of API's together with external platforms.
- **Ethics and Risk:** The main risks are Over-Reliance on LLM Outputs (Low Impact) and Intellectual Property and Ownership (Low Impact). Since the frameworks produced will be used by humans, any mistake will most likely be identified early on. If this process can be completely automated some jobs might change in the future as a result but there seems to be no threat of jobs being completely replaced. When it creates unique or new frameworks a discussion will need to be had about how ownership will work. Organizations should establish clear policies on attribution and usage rights to prevent disputes

5. Clarifying Ambiguities

Large language models can analyze datasets and transcriptions to identify and clarify ambiguous points. They can enhance understanding by providing additional context based on prior discussions or by incorporating relevant information from external sources.

- **Communication & Certainty:** By reducing ambiguities, the quality of communication increases and uncertainty is reduced as stakeholders have a better understanding of the information presented to them.
- **Data input:** The model can use structured data and unstructured data. It can be provided with various documents, website links and other sources of text and analyse and cross reference these various sources.
- **Prompting structures used:** Few-Shot Prompting, Chain-of-Thought Prompting, Instruction-Based Prompting, Role-Playing
- **How well does the tool perform:** When provided with the right materials the tool can explain and clarify on basically any point that is unclear to the user. It can even do this when one document references something from a different document without specifically stating the things its referencing too. This shows a remarkable degree of understanding of the subject matter talked about from the LLM model.
- **Ethics and Risk:** The main risks are Over Reliance on LLM Outputs (Low Impact). This tool will always be directly used by humans and whilst misinformation is a risk, as long as this possibility is understood by users it should not provide too big of a risk. One important factor is that LLM have a tendency to provide fake information when forced to give results. Guidance to users and rules for the LLM should be created that prevent this from happening.

6. Bias & Fact Checking

LLMs have instant access to vast amounts of information, making them powerful fact-checking tools when used correctly. They can verify claims, cross-reference sources, and provide additional context or insights as needed.

- **Communication & Certainty:** By reducing misinformation uncertainty is also reduced as stakeholders can be assured that they are working with sound data and no unforeseen errors can happen as a result of being misinformed
- **Data input:** Both structured and unstructured data can be uploaded for checking. This tool does require a predesigned database that it uses to compare and identify bias too. For the current tool the biases earlier identified in this paper were used but this database can be extended.
- **Prompting structures used:** Role-Playing, Instruction-Based Prompting, Priming, Negative Prompting
- **How well does the tool perform:** The tool is able to identify various biases in reasoning and decision making, explaining which specific arguments might be flawed and why. It is however not always equally reliable and can sometimes give false negatives or positives. Fact checking is almost 100% reliable and the tool can clearly argue why it thinks something is or is not factually correct.
- **Ethics and Risk:** The main risks are Over Reliance on LLM Outputs (Low Impact), Privacy and Data Protection (High Impact). This tool will always be directly used by humans so as long as it is not depended upon for crucial parts of the project it can't do serious harm. Prompting does however play a huge role in the tools effectiveness and reliability so users should be properly informed on how and when to use it. Users should be aware that if they want to fact check sensitive data, the information in these files might be used by OpenAI to train their models.

7. Investigating

With the right prompts, LLMs can serve as powerful research assistants, efficiently analyzing vast amounts of literature and data to extract relevant insights that can inform and support the project.

- **Power structures:** By giving everyone equal access to resources it allows any stakeholder, big or small, to be equally informed and thus removes hierarchy.
- **Data input:** ChatGPT 4o (and most likely other LLM too) can retrieve a wide array of information from various sources of the internet such as Real-Time and Up-to-Date Information, Business and Industry Data, Local Information, Technology and Science etc.
- **Prompting structures used:** Role-Playing, Instruction-Based Prompting, Few-Shot-Prompting, Contextual-Prompting
- **How well does the tool perform:** One of ChatGPT's biggest strength is in assisting its user in research. It can quickly find both large datasets or small and precise pieces of information from a wide array of sources. It can transform this information into a wide array of useful outputs. Additionally, when provided with data sources by the user, it can add these to the existing pool of information and integrate and use them with other datasets.
- **Ethics and Risk:** The main risks are Over Reliance on LLM Outputs (Low Impact). This tool will always be directly used by humans so as long as it is not depended upon for crucial parts of the project it can't do serious harm. Prompting does however play a huge role in the tools effectiveness and reliability so users should be properly informed on how and when to use it.

8. Equitable Participation

An LLM can act as a facilitator during discussions or debates by analyzing key metrics such as speaking time per stakeholder and the tone or language used. If imbalances are detected, it can provide real-time guidance to promote more equitable participation and ensure a fair, balanced dialogue.

- **Power structures:** LLM can help ensure more equitable involvement from all stakeholders, fostering inclusive and balanced discussions.
- **Data input:** The tool can use and analyse any text file that gets uploaded. Live speech to text input currently isn't integrated into chatGPT yet but using API keys together with other platforms like Azure or google live transcribe can make this possible.
- **Prompting structures used:** Role-Playing, Instruction-Based Prompting
- **How well does the tool perform:** The tool can analyse text both on context and general conversational trends such as one person talking more than others. It has a good understanding of intonation and can identify hostile language. Additionally, it can provide explanations informing users why stakeholders might be saying certain things and provide tips on how to de-escalate.
- **Ethics and Risk:** The main risks are Bias, Stereotypes, and Representational Harms (High Impact). LLM has inherently been designed to have a strong ethical compass and will thus most likely never provide harmful insights based on just transcript information unless specifically asked to do so and even then it might object. Nevertheless, its good to consider that there can still be bias and that this tool won't be an easy fix to overcome all bias and inequality in collaboration.

9. *Proposing Win-Win Solutions*

LLMs can efficiently and objectively identify areas of overlap between stakeholders' values and potential benefits. By generating these insights quickly, they help stakeholders recognize the project's value early in the process, fostering greater commitment and alignment.

- **Power structures:** By involving an unbiased entity to assess potential benefits for all stakeholders, the evaluation becomes more objective, ensuring a fair and well-balanced perspective on the optimal outcome for everyone involved.
- **Value alignment:** An unbiased entity can provide a more objective assessment of the true value that can be achieved, potentially uncovering opportunities that stakeholders may have overlooked.
- **Data input:** The tool can work with both structured and unstructured information but the information needs to be filtered and ordered beforehand to make sure that no unnecessary information gets used. Because of this, sources like website links might not be optimal for this tool.
- **Prompting structures used:** Role-Playing, Instruction-Based Prompting, Chain-of-Thought Prompting, Multi-Task Prompting, Anchor Prompting
- **How well does the tool perform:** When provided with information rich content the tool can quickly and effectively extract the important data and used it to make a variety of prediction and evaluations such as: Win-Win Collaboration Opportunities, Project-Wide Impact Analysis, Key Challenges & Recommendations and General Recommendation.
- **Ethics and risk:** The main risks are Over Reliance on LLM Outputs (Low Impact) Bias, Stereotypes, and Representational Harms (Low Impact). This tool provides advice directly to human stakeholders so it can never do unchecked harm. The tool could be misused but damage would be limited to providing bad advice which humans still always get a chance to ignore.

10. *Arguing on behalf of the opposition*

When certain stakeholders are underrepresented in a project, relevant data can be gathered to create LLM-driven representations of their perspectives. These AI-generated stakeholders can then advocate for their values and interests during discussions, ensuring their voices are considered in decision-making.

- **Power structures:** Allowing LLMs to represent underrepresented stakeholders or those unable to advocate for themselves helps ensure balanced and equitable discussions.
- **Value alignment:** Ensuring that all perspectives are represented in a discussion enables a more in-depth exploration of the subject matter and uncovers potential value that might otherwise be overlooked.
- **Data input:** The tool can use various sources to train itself in representing stakeholders, both structured and unstructured. Sources can vary from text directly provided, papers, data sheets or even website links.
- **Prompting structures used:** Role-Playing, Instruction-Based Prompting, Negative Promptin, Contrastive Prompting
- **How well does the tool perform:** The performance of this tool is completely reliant on the amount of information it is provided with. The more information it gets, the better it can represent opposition in a realistic way. The language and argumentation structure is of high quality and LLM, unique ability to adjust its reasoning structure based on its talking partner makes it quite effective. Additionally, LLMs inherent trait of always staying polite makes it quite pleasant, even in disagreements.
- **Ethics and risk:** The main risks are Bias, Stereotypes, and Representational Harms (High Impact). Since this tool might be put in a position of representing other stakeholders groups, the danger of misrepresentations is present. This can introduce various ethical problems. To mitigate this clear rules should be set for the LLM, instructing it what it is and isn't allowed to do. Additionally, the user should also be properly informed about how the tool should be used.

11. *Inclusive Decision-Making*

LLMs can be integrated into decision-making processes to provide unbiased insights and perspectives. They can also represent the values and viewpoints of stakeholders who may be absent or underrepresented, ensuring more inclusive and well-informed decisions.

- **Power structures:** Allowing LLMs to represent underrepresented stakeholders or those unable to advocate for themselves helps ensure balanced and equitable discussions.
- **Value alignment:** Ensuring that all perspectives are represented in a discussion enables a more in-depth exploration of the subject matter and uncovers potential value that might otherwise be overlooked.
- **Data input:** The tool processes structured and unstructured data, including stakeholder feedback, meeting transcripts, historical decision-making records, ethical guidelines, and industry best practices. It synthesizes diverse perspectives, ensuring that underrepresented or absent voices are included in the decision-making process.
- **Prompting structures used:** Instruction-Based Prompting, Chain-of-Thought Prompting, Contextual-Prompting, Role-Playing
- **How Well Does the Tool Perform:** LLMs excel at integrating diverse viewpoints into decision-making processes by analyzing large datasets, detecting overlooked perspectives, and structuring inclusive discussions. They can highlight ethical considerations, conflicts of interest, and alternative solutions, fostering balanced and well-informed decisions.
- **Ethics and risk:** The primary risk is bias, stereotypes, and representational harms (High Impact). If the model is trained on biased datasets or lacks representation from diverse groups, it may reinforce systemic inequalities rather than mitigate them. Additionally, AI-driven insights could be misused to justify biased decisions if not critically evaluated. To minimize these risks, human validation should be embedded in the process, ensuring that AI-assisted decision-making aligns with fairness, inclusivity, and ethical standards.

12. *Enhancing communication*

LLMs can act as communication bridges between stakeholders with different expertise or cultural backgrounds by translating technical jargon into layman's terms or translating between languages, ensuring clarity and mutual understanding.

- **Communication & Certainty:** By leveraging LLMs for communication, stakeholders gain a clearer understanding of each other's goals and potential outcomes, effectively reducing uncertainty and fostering better collaboration.
- **Value alignment:** By facilitating clear and effective communication, LLMs can enhance productivity and streamline the alignment of stakeholder values, ensuring more efficient collaboration.
- **Data input:** The tool can use various sources to train itself in representing stakeholders, both structured and unstructured.
- **Prompting structures used:** Instruction-Based Prompting, Negative Prompting, Contrastive Prompting
- **How well does the tool perform:** This tool strongly varies based on the context and the amount of data it has been provided with. When used to change tone it performs extremely well, being able to switch between different levels of formality and emotional tones. When switching between layers of expertise going from expert to novice is far easier as it requires less additional context. Going from novice to expert requires additional information and datasets for the LLM since it will otherwise try to find its own sources or make its own interpretation which might not align with reality.
- **Ethics and risk:** The main risks are Over Reliance on LLM Outputs (Low Impact). This tool if used directly with humans shouldn't be able to cause much harm, however one of the big risks is the appeal of automation with this tool. This can lead to automated systems through which text gets transformed based on predefined user descriptions. When too many steps are taken without human oversight, mistakes and bias can creep into the system which could lead to harm.

13. Conflict mediation and resolution

By analyzing stakeholder concerns, previous discussions, and project goals, LLMs can generate neutral summaries and propose compromise solutions, helping to mediate disputes and resolve conflicts fairly.

- **Power structures:** Analyzing stakeholder concerns and project goals with LLMs can reduce hierarchy by generating neutral summaries and fair compromise solutions to mediate disputes effectively.
- **Data input:** The tool processes structured and unstructured data. It can analyze conversation logs, emails, meeting transcripts, and policy documents to identify conflicting viewpoints and areas of potential compromise.
- **Prompting structures used:** Role-Playing, Instruction-Based Prompting, Chain-of-Thought Prompting, Contextual-Prompting
- **How well does the tool perform:** LLMs can effectively synthesize different perspectives into neutral summaries and propose fair, balanced solutions to disputes. The tool can highlight underlying concerns, pinpoint areas of agreement, and suggest actionable compromises that align with project goals. However, its effectiveness depends on the complexity of the conflict and the availability of well-structured input data.
- **Ethics and risk:** The primary risk is over-reliance on LLM outputs, as users might trust AI-generated resolutions without further verification or human oversight. While LLMs can assist in reducing power imbalances by providing unbiased summaries, they might not fully capture implicit biases or emotional nuances in conflicts. Therefore, the tool should be used as a support mechanism rather than a sole decision-maker, ensuring human judgment remains central to conflict resolution processes.

14. Ethical Decision Support

LLMs can be trained to assess ethical considerations in decision-making, highlighting potential ethical dilemmas and suggesting frameworks or best practices to navigate them responsibly.

- **Value alignment:** including an ethical view on the decision making process guarantees that certain values such as equality and fairness are addressed.
- **Power structures:** By taking into account ethical considerations the LLM can also help assure that all parties are included in the final decision by focusing on values such as equality and fairness
- **Data Input:** The tool analyzes structured and unstructured data, including organizational policies, ethical guidelines, stakeholder concerns, and historical decision-making patterns. It can reference established ethical frameworks, industry best practices, and case studies to assess ethical considerations in various contexts.
- **Prompting Structures Used:** Chain-of-Thought Prompting, Instruction-Based Prompting, Contextual-Prompting, Role-Playing
- **How Well Does the Tool Perform:** LLMs can effectively identify ethical dilemmas, highlight conflicting values, and propose best-practice frameworks for ethical decision-making. By embedding value alignment into the process, the tool ensures that principles such as equality and fairness are actively considered. Additionally, it helps restructure power dynamics by making sure all stakeholders' perspectives are taken into account. However, ethical decision-making is inherently complex and often requires nuanced judgment and contextual understanding, which AI may struggle to fully grasp.
- **Ethics and Risk:** The primary risk is bias, stereotypes, and representational harms (High Impact). If trained on biased data or influenced by dominant cultural perspectives, LLMs might reinforce existing inequalities rather than mitigate them. Ethical assessments may also be influenced by the limitations of AI's interpretive abilities, potentially leading to oversimplified solutions or unintended exclusions. To mitigate these risks, human oversight is critical, ensuring that AI-assisted ethical evaluations are transparent, well-rounded, and aligned with diverse stakeholder perspectives.

15. *Identifying Hidden Value:*

While humans excel at this task, the effort required and the potential to reduce bias, when implemented correctly, make it particularly well-suited for LLMs. LLM's can scan to large amounts of text in little time, spotting themes and trends that humans might overlook.

- **Value alignment:** By increasing the amount of values identified all stakeholder got a more complete and well rounded overview of what the project is about.
- **Data Input:** The tool can process large volumes of structured and unstructured text, including reports, emails, surveys, meeting transcripts, and other project-related documents. It scans for recurring themes, overlooked insights, and underlying values that might not be immediately obvious to human reviewers.
- **Prompting Structures Used:** Chain-of-Thought Prompting, Instruction-Based Prompting, Contextual-Prompting, Few-Shot Prompting
- **How Well Does the Tool Perform:** LLMs are particularly effective at recognizing patterns, implicit themes, and emerging trends within extensive datasets. They can surface hidden value by detecting underrepresented concerns, stakeholder priorities, and overlooked opportunities. By expanding the range of identified values, the tool helps create a more complete and inclusive understanding of a project's scope.
- **Ethics and Risk:** The main challenge lies in bias, stereotypes, and representational harms (High Impact). If the training data or prompt framing introduces biases, the tool may reinforce existing stereotypes or overlook critical perspectives. Additionally, value identification is inherently subjective, meaning that AI-generated insights must be validated through human oversight to ensure ethical and fair representation of all stakeholders. Guardrails should be implemented to ensure that the LLM does not amplify dominant narratives at the expense of marginalized voices.

16. *Documenting Goals and Values*

LLMs can efficiently identify, organize, and document stakeholders' goals and values, streamlining the process and significantly reducing the time required.

- **Value alignment:** By quickly and effectively documenting goals and values LLM can help guarantee that no information is lost.
- **Communication & Certainty:** Being able to document goals and values and transform them to be understandable for all stakeholders increases clarity and certainty
- **Data Input:** The tool processes structured and unstructured data, including stakeholder interviews, meeting notes, emails, reports, and project documentation. It identifies, organizes, and synthesizes key goals and values into a structured format, ensuring clarity and consistency.
- **Prompting Structures Used:** Instruction-Based Prompting, Chain-of-Thought Prompting, Contextual-Prompting, Few-Shot Prompting
- **How Well Does the Tool Perform:** LLMs excel at streamlining the documentation of goals and values by efficiently extracting relevant information and structuring it in an accessible manner. This reduces the risk of information loss and enhances value alignment by ensuring all perspectives are captured. Additionally, by transforming complex or abstract values into clear, stakeholder-friendly language, the tool improves communication and certainty within teams and organizations. However, while LLMs can organize information efficiently, human input is necessary to refine, contextualize, and validate outputs, especially in cases where subtle nuances or implicit goals need to be accounted for.
- **Ethics and Risk:** The primary risk is over-reliance on LLM outputs (Low Impact). If users blindly trust AI-generated documentation without review, important contextual details or stakeholder-specific priorities might be misrepresented. To mitigate this, the tool should be used as a supporting mechanism, with human oversight ensuring accuracy and relevance.

17. Anonymized data Collection

Fully anonymized data collection can help eliminate bias in datasets and reduce hierarchical influences, fostering a more equitable and objective workflow and decision-making process.

- **Power structures:** By anonymizing the data it removes labels that could influence decision making down the line. All data becomes equal when it no longer known which stakeholders is associated with each data set.
- **Data Input:** The tool processes structured and unstructured data from surveys, feedback forms, reports, and communication logs. It removes identifiable markers such as names, roles, demographic details, and other personal information while preserving the integrity of the data for analysis.
- **Prompting Structures Used:** Instruction-Based Prompting, Contextual-Prompting
- **How Well Does the Tool Perform:** LLMs can efficiently anonymize data by stripping personally identifiable information, ensuring that decision-making remains objective and unbiased. By eliminating hierarchical influences, the tool fosters a more equitable and fair workflow, where data is evaluated on merit rather than the identity of its source. This improves power structures by neutralizing potential biases related to social status, role, or affiliations. However, while AI-driven anonymization is effective, it is not infallible, subtle patterns in language or data structure can still unintentionally reveal information about stakeholders.
- **Ethics and Risk:** The primary concern is Privacy and Data Protection (High Impact). Improper anonymization could lead to re-identification risks, where seemingly anonymous data can be traced back to specific individuals. Additionally, certain anonymization techniques may inadvertently strip context that is crucial for accurate decision-making. To mitigate these risks, strict data protection policies and human oversight should be enforced, ensuring that sensitive information is securely handled and anonymization processes are regularly reviewed for effectiveness.

18. Stakeholder Risk Analysis

Leveraging prior stakeholder and organizational data, LLMs can analyze current project information to generate personalized risk assessments, identifying key areas that may specifically impact each stakeholder.

- **Communication & Certainty:** By identifying risks for stakeholders, LLMs can help give an overview of what to be mindful and aware of during the project, allowing stakeholders to respond accordingly
- **Data Input:** The tool processes structured and unstructured data, including previous stakeholder interactions, organizational reports, project documentation, and industry-specific risk factors. It analyzes historical data alongside current project details to generate personalized risk assessments for each stakeholder.
- **Prompting Structures Used:** Chain-of-Thought Prompting, Instruction-Based Prompting, Contextual-Prompting
- **How Well Does the Tool Perform:** LLMs are highly effective at identifying stakeholder-specific risks by comparing current project dynamics with historical patterns and known risk factors. This enhances communication and certainty, ensuring stakeholders are aware of potential challenges and can proactively mitigate them. The tool provides an overview of critical risks, allowing stakeholders to make informed decisions and adjust strategies accordingly. However, while the tool is valuable for surfacing potential risks, it lacks real-world intuition and adaptability, meaning that human oversight is necessary to validate and contextualize assessments.
- **Ethics and Risk:** The main concern is over-reliance on LLM outputs (Low Impact). While AI-generated risk assessments can highlight important areas of concern, they should not be treated as definitive conclusions. Stakeholders should be encouraged to cross-verify AI-generated insights with domain experts and real-world observations to ensure well-rounded decision-making.

19. Policy & Regulation Awareness

LLMs can be used to track and summarize relevant policies, regulations, and compliance requirements, ensuring that stakeholder discussions and decisions remain within legal and ethical boundaries.

- **Communication & Certainty:** By having a proper overview and checks of what is within the boundaries of law and regulation stakeholders can be assured that this won't become a surprise problem later on
- **Data Input:** The tool processes legal documents, industry regulations, compliance requirements, and policy frameworks. It can extract and summarize key regulatory guidelines from government websites, organizational policies, and legal texts, ensuring that stakeholders remain informed.
- **Prompting Structures Used:** Instruction-Based Prompting, Chain-of-Thought Prompting, Contextual-Prompting
- **How Well Does the Tool Perform:** LLMs are highly effective at tracking, summarizing, and contextualizing policies and regulations relevant to a given project. By automating this process, stakeholders gain a clear and concise understanding of legal boundaries, helping to ensure compliance. This enhances communication and certainty, as stakeholders can make decisions with the confidence that they align with regulatory frameworks, reducing the risk of unforeseen legal challenges. However, while LLMs can efficiently highlight key points, they lack the nuanced judgment required for complex legal interpretation, meaning legal professionals should still verify outputs for accuracy.
- **Ethics and Risk:** The primary risk is over-reliance on LLM outputs (High Impact). While AI can efficiently summarize legal information, it may miss nuances, fail to interpret new regulations accurately, or provide outdated information if not properly maintained. To mitigate this, AI-generated summaries should be cross-checked with legal experts and official regulatory sources to ensure complete and up-to-date compliance.

20. Transforming Data Based on Context

LLM can be used to transform data and information based on the context of the recipient/reader. This can help make information sharing more efficient and effective.

- **Communication & Certainty:** LLMs' ability to tailor information to the recipient's context enhances communication by ensuring that messages are relevant, clear, and aligned with the recipient's level of understanding or needs. This contextual adaptation reduces ambiguity, increasing confidence in the information received and supporting more informed decision-making.
- **Value Alignment:** By transforming data based on the recipient it becomes far easier to communicate and align value between stakeholders from various backgrounds
- **Prompting Structures Used:** Instruction-Based Prompting, Chain-of-Thought Prompting, Contextual-Prompting, Negative Prompting
- **How Well Does the Tool Perform:** The LLM can identify key information needed to build up a user profile. This data can be viewed and adjusted by the user whenever necessary. The tool can then proceed to transform new data into a form that better matches the context of the user to increase readability and highlight what is specifically valuable for the user.
- **Ethics and Risk:** The primary risk is over-reliance on LLM outputs (Low Impact). The tool will alter information that is provided with. Even though it has been specifically instructed to not create any new text and just transform existing information, there might still be alterations that are not 100% correct. Because of this the user will need to be aware that they still need to critically assess anything provided themselves as well. The tool just provides support in this case.

APPENDIX (F)

1. Creating timetables and schedules

	-1	+1	+2
Ethics			
Impact			
Performance			

6.Bias and Fact Checking

	-1	+1	+2
Ethics			
Impact			
Performance			

2. Role Allocation Support

	-1	+1	+2
Ethics			
Impact			
Performance			

7.Investigating

	-1	+1	+2
Ethics			
Impact			
Performance			

3. Trend Predictions

	-1	+1	+2
Ethics			
Impact			
Performance			

8.Equitable Participation

	-1	+1	+2
Ethics			
Impact			
Performance			

4. Creating Frameworks

	-1	+1	+2
Ethics			
Impact			
Performance			

9.Proposing Win-Win Solutions

	-1	+1	+2
Ethics			
Impact			
Performance			

5.Clarifying Ambiguities

	-1	+1	+2
Ethics			
Impact			
Performance			

10.Arguing on behalf of the opposition

	-1	+1	+2
Ethics			
Impact			
Performance			

APPENDIX (F)

11. Inclusive Decision-Making

	-1	+1	+2
Ethics			
Impact			
Performance			

12. Enhancing communication

	-1	+1	+2
Ethics			
Impact			
Performance			

13. Conflict mediation and resolution

	-1	+1	+2
Ethics			
Impact			
Performance			

14. Ethical Decision Support

	-1	+1	+2
Ethics			
Impact			
Performance			

15. Identifying Hidden Value:

	-1	+1	+2
Ethics			
Impact			
Performance			

16. Documenting Goals and Values

	-1	+1	+2
Ethics			
Impact			
Performance			

17. Anonymized data Collection

	-1	+1	+2
Ethics			
Impact			
Performance			

18. Stakeholder Risk Analysis

	-1	+1	+2
Ethics			
Impact			
Performance			

19. Policy & Regulation Awareness

	-1	+1	+2
Ethics			
Impact			
Performance			

20. Transforming Data Based on Context

	-1	+1	+2
Ethics			
Impact			
Performance			

APPENDIX (F)

11. Inclusive Decision-Making

	-1	+1	+2
Ethics			
Impact			
Performance			

12. Enhancing communication

	-1	+1	+2
Ethics			
Impact			
Performance			

13. Conflict mediation and resolution

	-1	+1	+2
Ethics			
Impact			
Performance			

14. Ethical Decision Support

	-1	+1	+2
Ethics			
Impact			
Performance			

15. Identifying Hidden Value:

	-1	+1	+2
Ethics			
Impact			
Performance			

16. Documenting Goals and Values

	-1	+1	+2
Ethics			
Impact			
Performance			

17. Anonymized data Collection

	-1	+1	+2
Ethics			
Impact			
Performance			

18. Stakeholder Risk Analysis

	-1	+1	+2
Ethics			
Impact			
Performance			

19. Policy & Regulation Awareness

	-1	+1	+2
Ethics			
Impact			
Performance			

20. Transforming Data Based on Context

	-1	+1	+2
Ethics			
Impact			
Performance			