ON THE NATURE OF DESIGNER AUTONOMY

COLLABORATIVE FRAMING IN THE AGE OF AI

Graduation Project Report
Ameya Sawant
MSc Strategic Product Design
Faculty of Industrial Design Engineering, TU Delft

GRADUATION PROJECT

MSc Strategic Product Design
Faculty of Industrial Design Engineering
April 2025 to August 2025

AUTHORED BY

Ameya Sawant
5950899

SUPERVISORY TEAM

Project Chair
Fernando Secomandi
Department of Design, Organisation and Strategy
Faculty of Industrial Design Engineering
Delft University of Technology

Project Mentor
Kars Alfrink
Department of Sustainable Design Engineering
Faculty of Industrial Design Engineering
Delft University of Technology

Delft University of Technology
Faculty of Industrial Design Engineering
Landbergstraat 15
2628 CE Delft
The Netherlands
www.tudelft.nl

TABLE OF CONTENTS

ABSTRACT

This graduation project is situated at the emergent intersection of design methodology and the integration of artificial intelligence in design practice. Increasing algorithmic mediation brings challenges to fundamental understandings of designer's capacity and responsibility to make value-laden decisions. Despite autonomy's prominence in responsible AI frameworks, definitions remain abstract and inconsistently applied to design practice, risking fragmented interpretations as human and machine agents negotiate framing authority. Framing is selected as the critical focus of this inquiry because it constitutes the primary, value-laden act where designers define problems, set priorities, and ultimately shape which futures are made possible. Existing models of collaborative reflective practice lack explicit consideration of non-human actors in the sense-making process, and its impact on the human actor's role and responsibilities. To address this gap, the project aims to develop an understanding of Designer Autonomy that will explicate its meaning, components, and vulnerabilities for design framing practice in an AI-augmented future. Adopting a Value Sensitive Design approach, the project uses the tripartite investigations to define and understand the value of autonomy of designers in collaborative design activities performed with AI.

The conceptual investigation explicates designer autonomy into two interdependent dimensions of authenticity (unmanipulated acting on one's own beliefs) and agency (independent enactment of those beliefs) in the context of responsible design practice. The value is situated in framing practice in design, claiming that effects of exercising of autonomy, and the potential of shaping of its perception occurs in this process of collaborative sense-making. This informs the empirical investigation, which undertakes an autoethnography as the methodology to understand the designer's negotiation of agency and perception of influence of authentic beliefs in AI mediated problem framing activity. This resulted in four factors key drivers of co-framing: controlling the pace of activity, having a sense of role and responsibility, relevant digital literacy about the AI tool, and an explicit moral and political lens. The technical investigation reveals three interconnected characteristics of AI: bias, trust, and explainability, and how these can influence the human actor during collaborative framing processes. While AI systems may not be autonomous actors in their own right, their capability and integration into processes means they exert valid influence over the designer's autonomy, therefore shaping the design decisions.

The key novel results of this project are a descriptive account of autonomy of a designer, its two interdependent dimensions, and the specific risks to each from AI in practical scenarios. A novel conceptual model called Collaborative Framing (Co-framing) practice develops on design theory research on collaborative reflective practice, to describe it as the primary site for ethical-epistemic engagement between human and non-human (AI) actor collaboration in abductive design processes. This leads to discussions of application of the outcome to prescriptive design methods, connecting designer to user autonomy, extending the co framing model across design stages and domains, and embedding ethical reasoning directly into AI mediated workflows.

PREFACE

One must wait until the evening,
To see how splendid the day has been.

— Sophocles

Biographer Robert Caro begins his book, *The Power Broker: Robert Moses and the Fall of New York*, with the quote mentioned above. The interpretation of which, along with the context of its presence in Caro's work, brings forward a paradox about creative processes, and the relationship between the creator and their audience, and most importantly the people that enable the process.

The origin of that quote is seemingly untraceable, which is in stark contrast to the methodologies and standards of Caro as a journalist and biographer. His process is characterised by the extreme lengths he goes to effectively relive the lives of his subjects, to trace back every piece of mundane paperwork and corroborate decades old stories, building a research archive so vast that it was exhibited at the New York Historical museum. The choice to begin his work with this quote means he felt compelled to convey what it meant to the readers. Maybe the quote cautions the readers to judge the life of Robert Moses only after fully understanding it through the book, sharing a principle with them which guides his work ethic, and its eventual outcomes.

The book brought forward in extensive detail the factors which shaped Robert Moses as a figure, and how he reshaped the fabric of New York by decades of unchecked accumulation of power. Moses's reign as a public administrator saw immense development of roads, parks, public services and governance changes. His projects are now known for displacing communities, reinforcing racial and economic segregation, and sacrificing social and environmental well-being for monumental urban projects. But at the time of publishing, the public opinion of Moses was generally positive, and he was known as the "master builder". The Power Broker drastically shifted this, with a widespread rethink of what "good" infrastructure should be. So before embarking on his process to study Moses for a biography, Caro had made a judgement on Moses's life based on the impact he had on others. His own perception, experiences and sensitivities compelled him to tell the story of Robert Moses, and to convince others that it is a story worth telling. Caro said in a podcast interview discussing this book, that after he was done with the research, he couldn't not start writing the final book before he knew what to end it with. But the process of creation starts much before its material realisation.

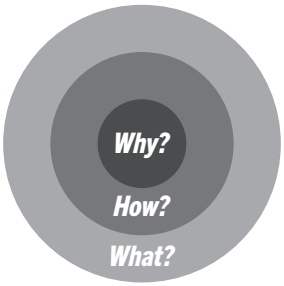
There is a sense of trust invoked by the quote at the start of the book, for its audience to have when experiencing the creation. The same trust, I believe, is present between the creator and the people who enable, support and empower the process. This trust is more critical at a stage when there is inherent uncertainty about the nature and value of the final output, as it creates a space for the creator to feel free to pursue questions which they inexplicably feel important to answer.

The original idea for this project was to understand the mechanisms around creative acts such as Caro's inclusion of the unverifiable quote, which could be framed as a deliberate decision made in accordance with his own convictions about what mattered for the reader's experience. Such approaches are not uncommon in design practice, but do come at the risk of unintended paternalism through the solutions. With artificially intelligent technologies integrating in methodologies, I felt a draw to study how a designer's ability to make choices that balance personal judgment, ethical responsibility, and the anticipated needs of users will shape in the future. Which brings me to exploring the value of autonomy of an individual in the context of the expectations of design practice and properties of the technology.

I would like to wholeheartedly thank my supervisory team of Fernando Secomandi and Kars Alfrink, for showing trust in me to pursue this project which deviates from traditional design graduation projects. Their support allowed me to begin exploring this subject even if there was no clear outline of what the final outcome was going to be. Using the problem solving approaches of design, this project aims to understand the value of autonomy of a designer, deliberately choosing to not be situated in a specific designed product or service to maintain complete focus on designer behaviour and not the results of it. I would also like to thank Delft Design for Values institute for their support on matters of philosophy of technology, and the opportunity given by the Student Thesis Awards Annual Theme 2025: Design for Human Autonomy.

I am grateful to my friends for being a patient ear to my extensive ramblings about design and the silly world we get to experience rapidly changing around us, and being also a voice of reason and reflective curiosity, which has undoubtedly contributed to this project. Finally, I would like to thank my parents, whose faith in me has been the ultimate foundation of the faith I have in my beliefs and abilities. I hope this project is a valuable realisation of the privileges I am afforded to be an autonomous, and responsible individual and designer.

CHAPTER 1
INTRODUCTION



1.1 *Why* do we need to understand Designer Autonomy ?

Human autonomy is considered a critical value in modern democratic society, a value which is closely connected to the ideals of freedom, independence and liberty. It is an important contributor to one's notions of their identity, and their role in participating in collective society (Christman, 2008). With growing technological capabilities and the consequent collective dependency, autonomy has come under increasing pressure in socio-technical systemic domains. This is due to the potential of autonomous technologies to transform how decisions are made and enacted across social, political, and professional levels (Floridi & Cowls, 2019). A socio-technical systemic domain can be interpreted as a complex network where the people and technologies interactively shape experiences and outcomes. Actors in such a system are human- individual and collective, as well as technological- from simple machines to complex multi-modal tools. When exploring the values of human actors in such systems, what is addressed invariably is the reasonable expectations and responsibilities of the human role. Critically, how it reacts to, and is shaped by the technologies present in the interactive processes (Franssen, 2015). The technology central to this research is Artificial Intelligence. AI presents numerous opportunities for collective progress, but also presents risks to human autonomy, particularly due to its potential to influence, manipulate, or override human decision-making processes (Susser et al., 2019). Autonomy is not only significant to the individual but also plays a role in shaping systemic structures, as it determines the extent to which individuals can exercise control and responsibility in environments increasingly mediated by opaque and probabilistic systems (Floridi & Cowls, 2019).

Recognising this tension of human autonomy and automation by technology, several frameworks for responsible AI design and deployment have elevated autonomy as a central tenet, a value to be collectively upheld. However, the effectiveness of these frameworks in practice is limited by the lack of a consistent and specific understanding of the value itself (Prunkl, 2022). As AI begins to mediate professional judgement in socio-technical practices, it becomes essential to articulate how human roles shift. This articulation must be grounded in clear conceptualisations of human values within domain-specific practice (Dignum, 2019). For this project, it means exploring how autonomy is understood, exercised, and potentially compromised in design processes. Design theories and methodologies of practice are also rapidly adapting as autonomous technology is integrated to be a part of processes as well as its outcomes.

The domain of design has evolved from the shaping of products and artefacts to the structuring of services, experiences, systems, and even policies (Dorst, 2015b). With this evolving understanding of the practice, the role of designers can be described as normalisers of past and current values into the future (Wilkie & Ward, 2009). Traditional, or perhaps a simplified, understanding of design may lead to conclusions that it is a neutral task of developing solutions for the requirements or constraints shaped by clients and users. It could be argued that value interpretations from the stakeholders are shaped by their own morality, and are simply externalities to the design process, beyond the scope of the designer. This conceptualisation of the role of a designer faded with the understanding of how expert practitioners imposed their aesthetic and social goals through their process and outcomes (van den Hoven et al., 2015). As the scope of design was investigated further, sensitivity about the value shaping potential was explicated in the form of new theories and methodologies addressing it. Within design theory, "framing" emerged as a core of the design thinking approach (Dorst, 2011) and is considered as a political act of value selection and exclusion, where designers actively shape which futures are legitimised and which are rendered invisible through their interpretive choices (Prendeville et al., 2022).

A designer's practices shape solutions in ways which contribute to future meanings and prioritisations of values. Therefore, it is important to understand how interactions with actors in a socio-technical systems during design process have the potential to shape their own perceptions of their values. The interpretive role becomes particularly prominent for the emerging domain of strategic design, as it is the shaping of organisational strategies through design approaches and practices for the benefit of people, organisations, and the planet alike ("Strategic Design," 2016). Having an explicit and consistent understanding of a value is especially important when the value in question is as fundamental in nature as autonomy, and differing interpretations can have undesirable consequences (Melnyk et al., 2024). A descriptive ethical framework rooted in lived experience, is an important prerequisite for any valid normative claims about how autonomy ought to be exercised (Williams, 1986).

This research explores the autonomy of a designer as a value that is dynamically shaped, proposing that it is critical to fulfil the ethical demands of the practice. It investigates designers as the human actors (Pasmore et al., 2019), and AI tools as the technological actors in a socio-technical system when collaborating in a design process. This allows for interpretations of broader theoretical understanding of the impact on human autonomy when in tension with automation (Franssen, 2015) to be applied to research on emerging design practices. Framing is addressed as both a cognitive and a methodological site, where values or their interpretations are materialised. A contextually relevant explication of the value of autonomy is necessary to inform conversations about policy, pedagogy, and the evaluation of the role of human designers in an AI-augmented future.

1.2 *How* do we understand the value in the Age of AI?

This project adopts a Value Sensitive Design (VSD) approach to explore the value of autonomy in the context of AI-integrated design framing practice. Value Sensitive Design is a theoretically grounded, interactional approach to the design of technology that accounts for human values throughout the design process (Friedman et al., 2013). As developed by Friedman and colleagues, VSD involves three iterative and integrative investigations- conceptual, empirical, and technical, aimed at addressing how values can be recognised, embedded, and sustained in socio-technical systems. The adaptation of the VSD approach can be categorised into four loosely separate descriptions: an affinity with VSD, prescriptive VSD, critical VSD, and formative VSD. This research has "as affinity to VSD", which is described as values-oriented research in the field of human-computer interaction, and is rooted in a concern about the increasingly complex interplay between values and the design and use of technology. The structuring of investigations and core theoretical underpinning of an interactional understanding of technological appropriation, are fundamental to this research. Rather than assuming a fixed set of values or design outcomes, VSD allows for the exploration of contextual dynamics, and value tensions as they emerge through design activity (Davis & Nathan, 2015). In line with VSD, this project defines the foundational elements of the inquiry as follows:

Value: Autonomy.

Technology: Artificial Intelligence (AI), specifically as it mediates aspects of the design process as part of collaborative tools.

Context of Use: Framing practice in design as the site of decision-making.

Primary Stakeholders: Strategic design practitioners, who enact and shape the norms of future design practice.

Where this project differs from traditional applications of VSD is in two critical ways. First, it treats designers themselves as the "users", focusing on how they interact with and are influenced by these tools during value-laden framing practice. This shift in perspective foregrounds the internal ethical and epistemic conditions of design practice; how autonomy is experienced, justified, and negotiated by practitioners themselves. Second, while VSD is often used to inform prescriptive design interventions, this study adapts its structure to support a descriptive ethical inquiry. The aim is to understand how a designer's autonomy is expressed and shaped in context, forming a base for future prescriptive theory on how autonomy should be embedded into design tools or processes.

Among the eight challenges identified by Friedman et al. (2021) for advancing Value Sensitive Design, two are particularly relevant to the focus of this project: Framing and Prioritising Values, and Value Tensions. These refer respectively to the difficulty of identifying which values are relevant in complex design contexts and determining their relative importance, and to the conflicts that arise when values compete or contradict one another in real-world decision-making. While this project does not aim to resolve these challenges methodologically, it adopts the position that valuable insights into both can be gained by a descriptive account of designers as the users of AI tools.

1.3 *What* questions will be answered by this project?

The research is structured around three interrelated investigations as laid out by Value Sensitive Design:

Conceptual Investigation: The conceptual investigation examines the foundations of autonomy as a value and explores how these factors can be meaningfully situated in the context of design practice. The theoretical structure draws from literature on ethics, philosophy, and design theory to shape analytical criteria for studying the role of both the human and technological actors in shaping the value of designer autonomy.

Empirical Investigation: The empirical component explores how designer autonomy is enacted in practice, particularly during value-laden moments of design decision-making. Framing practice is treated as the primary site where autonomy becomes visible; where beliefs are expressed or compromised. A qualitative study using autoethnography as the methodology will allow for observations of designer autonomy in strategic design framing activity performed in collaboration with AI tool.

Technical Investigation: The technical investigation focuses on the characteristics of AI systems that influence the conditions under which designer autonomy is exercised. Observations made during the qualitative experiment will be informed by theoretical understanding of the contributing technological factors to the shaping of the value.

Together, these investigations offer a multifaceted understanding of designer autonomy as it is exercised in AI-mediated design practice. The conceptual investigation will develop an understanding of the value and the site in practice where it is exercised, informing the empirical investigation approach. The data generated from the empirical investigation will be coded inductively, thematically, and reflexively analysed to develop an understanding of designer autonomy for strategic designers in the age of artificial intelligence. The technical investigation will articulate risks posed by AI to the driving factors which shape the outcomes of design processes. Based on this, the key research questions to be answered by this project are:

Why is it important to understand designer autonomy as a value?

How is designer autonomy is shaped in collaborative design activity with AI?

What technical characteristics of AI tools impact designer autonomy in an interaction?

CHAPTER 2

CONCEPTUAL INVESTIGATION

The role of the conceptual investigation in the adapted Value Sensitive Design approach of this project is to explicate the value of autonomy, and why it is of relevance for designers, being the chosen stakeholders participating in the interaction with AI in their practice. Recognising the novel nature of this investigation, this section is divided into two parts. The first part will build on current knowledge of values in socio-technical environments and design theory to assemble an understanding of designer autonomy. The second part will then place this value in design practice, explicating how designer autonomy is exercised and shaped in design processes when collaborating with AI.

2.1 Constructing the Value of Designer Autonomy

This part explores prerequisite concepts which are fundamental to developing a consistent understanding of the value of designer autonomy in the age of AI. The goal is to construct autonomy as a value shaped by ethical principles and through disciplinary practices, culminating in a comprehensive articulation of designer autonomy that will serve as the basis for situating it in a design process.

2.1.1 What is a Value?

To understand autonomy as value, it is important to clarify what is meant by the term "value". Values, along with Norms and Virtues are described by van de Poel & Royakkers (2011) as "points of departure" to understand distinctions between key normative ethical theories. This brings up another relevant point of clarification, the meaning of morality, ethics and branches of ethical theory. Morality is the entirety of beliefs and actions with which individuals express what they understand to be good or right. Ethics, therefore, is the systematic reflection on the nature of morality, and is a methodological toolkit of addressing moral problems, not a set of answers. Ethics is commonly understood to comprise two interrelated branches: descriptive and normative ethics. Descriptive ethics seeks to understand and explain how individuals and groups engage with moral questions in practice; how they reason, what they value, and how they justify actions within specific social, cultural, or institutional contexts. It offers an account of morality "as it is lived", without making evaluative claims about what is right or wrong. In contrast, normative ethics aims to establish standards or principles for how individuals "ought" to act. It draws on distinctions such as values, norms, and virtues to formulate prescriptive theories of moral behaviour. This project is situated within the descriptive branch of ethics, aiming to understand how the value of autonomy is conceptualised and enacted in the real-world practice of design. Therefore, understanding the distinction between the values, norms and virtues is a critical foundation, their meanings being:

Values are enduring beliefs or standards that individuals or societies hold about what is important or desirable. They serve as guiding principles for behaviour and decision-making.

Norms are specific rules or standards of behaviour that are expected within a group or society. They are derived from values and dictate acceptable conduct.

Virtues are character traits or qualities deemed to be morally good and desirable in a person. They enable the evaluation of the individual with respect to values of society.

Although distinct, they are not discrete. Norms and consequentially virtues are dependent on an understanding of values, as norms without values lack direction. Understanding values comprehensively is critical to consistently shape norms and create expectations of the virtues of individuals. Values then, can be understood to be generalised ends which help determine which ideals are worthy of striving for in life, while norms are the means to prescribe morality shaped by the accepted values. With the identified need for a consistent (Prunkl, 2022), and domain specific (Dignum, 2019) explication of a value due to its consequential nature in the shaping of the future (Melnyk et al., 2024), a descriptive account of designer autonomy becomes a foundation for valuable prescriptive theory on how designers should be engaging in practice.

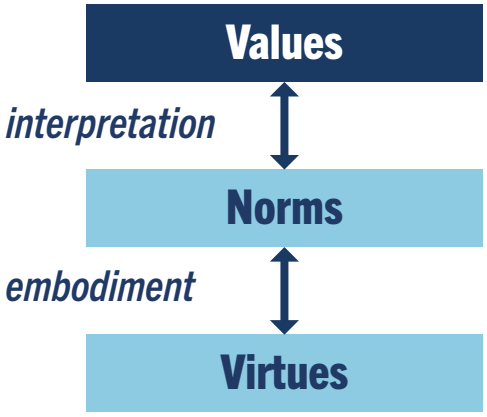


Fig. 1.: Interdependency of core concepts of Values, Norms and Virtues to understand normative ethical theory.

Through Value Sensitive Design, an understanding of values has been operationalised across the design process, placing them as a core driver of decision-making. Values, as described by VSD, are those which a person or group of people consider important in life (Friedman et al., 2013). Crucially, VSD establishes that values are not simply shaped by empirical worldly experiences- "outside-in" but also emerge from individual desires within the socio-cultural framework around them- "inside-out". This notion of dynamic nature of value construction translates into what is referred to as interactional theory when placed in the socio-technical relationships in which VSD operates in. It states that values neither originate purely from the technology's inherent properties (an endogenous theory) nor are simply a result of social forces (an exogenous theory). Values in a socio-technical systemic contexts are shaped continually by the interaction of properties embedded into technologies which either support or hinder certain values, with properties of the people interacting with the technology for their (van den Hoven et al., 2015).

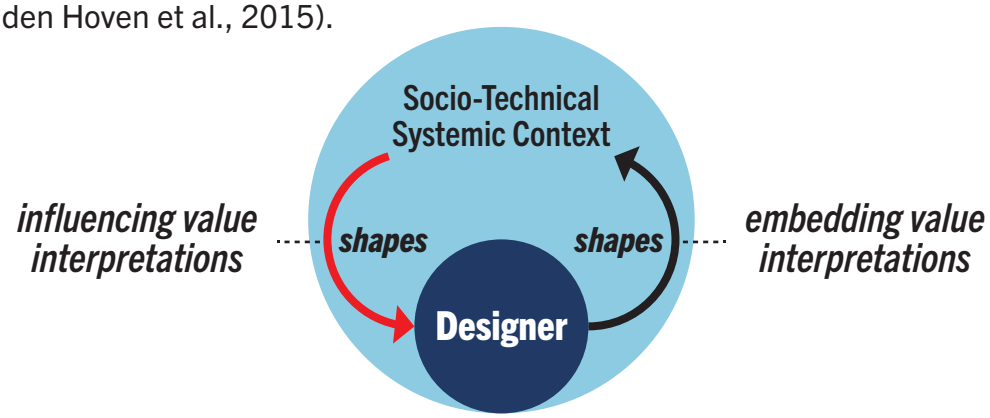


Fig. 2.: Representation of role of a designer in interactional understanding of shaping values, with descriptive focus of the paper highlighted in red.

Understanding values as both guiding ideals and socially embedded constructs enables a more nuanced ethical inquiry into how human actions are shaped, justified, and constrained. The distinction between values, norms, and virtues offers a conceptual structure for interpreting how beliefs translate into social expectations and embodied behaviours. In socio-technical systems, this structure becomes even more dynamic: technologies not only mediate human action but participate in the ongoing construction of what is considered valuable. This recognition sets the groundwork for understanding designer autonomy as a dynamic concept which is lived, negotiated, and expressed through designers' interactions with the systems and constraints that shape their practice (as highlighted in figure 2).

2.1.2 Human Autonomy in the Age of AI

Autonomy has long been treated as a foundational concept in moral and political philosophy, closely tied to ideas of responsibility, selfhood, and legitimacy. In its broadest sense, as outlined by Christman (2008), autonomy refers to a person's capacity for self-governance, the ability to reflect on one's beliefs, preferences, and desires and to act in accordance with them. This core idea has persisted over centuries, yet its interpretation has shifted substantially depending on the philosophical tradition, societal context, and underlying ethical commitments of the time.

The Kantian conception of autonomy, one of the most influential in moral philosophy, positions autonomy as the ability of rational agents to legislate moral law for themselves. In this view, to be autonomous is to act according to reason, independent of external influence, emotion, or inclination. Autonomy becomes synonymous with moral lawfulness: a person is free when they obey rules they have rationally willed for themselves. This framing had far-reaching implications for modern moral theory, especially for how responsibility is assigned and how freedom is understood in normative terms.

In contrast, John Stuart Mill's utilitarian tradition locates autonomy in the domain of liberty and self-development. For Mill, autonomy is not primarily about rational will, but about the ability to pursue one's own conception of the good life. Protection from coercion, by the state, by social norms, or by custom becomes the essential condition for individual flourishing. Mill's perspective opened space for autonomy to be seen not only as a moral ideal but as a political and developmental one: a value central to liberal democracies.

More recent developments in autonomy theory have pushed beyond these early formulations, especially in response to critiques that autonomy in its Kantian or Millian forms assumes an overly individualistic, abstract subject. The relational turn in autonomy theory, articulated by scholars such as Mackenzie and Stoljar, asserts that individuals are not detached rational agents, but embedded in social, cultural, and historical contexts. From this view, autonomy is not merely about independence, but about the capacity to make decisions in ways that are informed, supported, and enabled, or hindered, by one's social surroundings. Autonomy becomes something to be cultivated and sustained, not presumed (Christman, 2008).

While this evolution has expanded the applicability of autonomy as a concept, it has also introduced tensions. Jonathan Knutzen (2020) explores these tensions in his critique of what he terms formal views of autonomy, those that reduce autonomy to mere procedural features of choice, such as non-coercion or logical consistency. According to Knutzen, these formal accounts risk overlooking the very things that make autonomy meaningful, which he argues is its connection to the agent's identity, their capacity to respond to reasons, and their place within epistemic and moral communities.

Knutzen argues that formal features alone are insufficient to meaningfully describe autonomy, and it cannot be captured by simply checking whether a person's decision was uncoerced or rational in form. Instead, autonomy must involve an agent's recognition of themselves in their choices, and their capacity to justify those choices in light of reasons they understand and endorse. Autonomy, then, is inseparable from the concepts of knowledge, justification, and moral responsibility. It is not simply the condition for making choices, but the condition for owning them.

This critique refines our understanding of what is at stake when addressing autonomy, especially in the context of complex, interconnected societies shaped by sociotechnical systems. Through both its philosophical lineage and its recent critiques, several key insights emerge:

Autonomy is essential to the ***individual*** as a moral and political agent. It allows people to shape their lives in accordance with values they reflectively endorse.

Autonomy is foundational to ***society***, particularly in democratic frameworks where participation, accountability, and justice depend on people being able to act with informed self-determination.

Autonomy is ***fragile***. It can be undermined by both external control (e.g., coercion, manipulation, paternalism) and internal distortions (e.g., false beliefs, lack of critical reflection).

Autonomy involves ***responsibility***. To be autonomous is not only to choose, but to be accountable for those choices; to understand their consequences and be able to justify them.

Autonomy depends on ***knowledge*** and ***self-awareness***. It requires more than freedom from interference; it demands that individuals have the cognitive, emotional, and social capacities to engage with their reasons and values meaningfully.

Autonomy is ***not binary***. It is not something one either has or does not have, but a condition that can be supported, degraded, or shaped by the systems and structures one inhabits.

Taken together, these insights frame autonomy not as a static property or ideal, but as a situated and interpretive value, shaped through reflection, justification, and moral self-understanding within lived contexts. This understanding sets the foundation for examining how autonomy is affected and transformed in environments shaped by external systems, technologies, and collective norms. Addressing the complexity of the conceptualisation of autonomy and the risks this complication can bring to valuable discourse, Carina Prunkl (2022) offers a clarifying conceptual bifurcation that makes understanding and application more actionable. Autonomy can be understood to be formed by two interdependent components: authenticity and agency.

Authenticity refers to the individual's capacity to form beliefs, values, and preferences through processes that are critically reflective and recognisably their own.

It concerns whether one's internal motivations and justifications stem from autonomous reasoning or whether they are shaped in ways that bypass reflective judgement, for example, through manipulation, inherited assumptions, or opaque influence.

Agency refers to the capacity to act intentionally and to translate one's values and beliefs into action. This includes the ability to resist external pressures, make decisions under uncertainty, and take responsibility for outcomes.

It is diminished when systems limit meaningful choice, predefine desirable outcomes, or obscure causal relationships between action and effect.

By distinguishing between these two components, Prunkl's account provides a simplified yet nuanced lens for analysing how autonomy is shaped in contemporary technological societies. In a socio-technical application, an individual might retain the ability to act (their agency) but lose track of why they believe what they do (their authenticity), or vice versa. Understanding autonomy as composed of both authenticity and agency allows it to be examined as a relationally exercised value, shaped by internal moral reasoning and the external structures that enable or inhibit meaningful action.

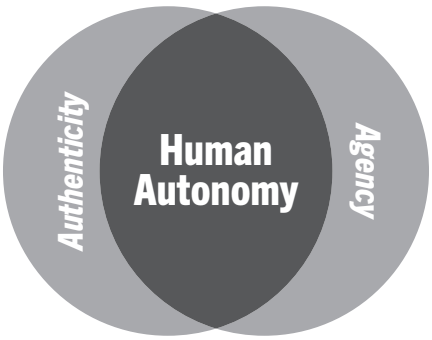


Fig. 3.: The dimensions of Human Autonomy in the Age of AI, Authenticity and Agency. *(Prunkl, 2022)*

To understand the shaping of this value by exogenous factors, the influences of technology can be simplified to be understood as automation. In the study of socio-technical systems, autonomy and automation often emerge in tension. Maarten Franssen (2015) explores how, in such systems, human operators are not fully embedded components but are treated as partially external agents. Their behaviours are shaped by system architecture but not reducible to it. This partial externality introduces a crucial ethical ambiguity: the human agent is held morally responsible for outcomes generated in systems where they may lack full epistemic access or decision-making control.

One key concern is what van den Hoven terms "epistemic enslavement" (as cited in Franssen, 2015) is a condition in which individuals operate based on instructions or insights generated by AI without the opportunity or means to critically assess their validity. In design, this manifests when designers engage with algorithmically generated briefs, insights, or optimisations that subtly shape the design direction while obscuring their underlying logic. Such circumstances dilute the conditions required for authentic judgment and reduce the designer's ability to exercise reflexive ethical reasoning.

Further, the structural configuration of automated systems may maintain formal human responsibility while stripping the designer of material control, creating an accountability vacuum. Designers are positioned as operators of systems whose outputs they are expected to legitimise without sufficient means of interrogation.

Operator Roles as Partially External Components: In sociotechnical systems, human operators are integrated not as complete system components but as partially external agents. Their behaviour is monitored and controlled externally, distinguishing them from fully automated components.

Epistemic Enslavement: When operators act based on prescribed instructions without understanding the underlying reasons. This condition reduces the operator's autonomy, making them dependent on the system's directives.

Responsibility Assignment: The assignment of responsibility to human operators in automated systems raises ethical concerns, especially when their autonomy is constrained, yet they are held accountable for system outcomes.

This explication of human autonomy in the age of AI illustrates that the value in the context of a socio-technical practice is neither a static trait nor an abstract ideal. It is a situated, relational, and epistemically grounded value shaped by both internal reasoning and external systems. Through the dual lens of authenticity and agency supports the understanding of how designers' moral and epistemic capacities are supported or undermined in AI-mediated contexts. As technologies increasingly mediate design decision-making, preserving designer autonomy requires critical attention to how technologies influence belief formation, decision justification, and value expression within practice.

2.1.3 Evolving Role of A Designer

With an understanding of the nuances of the value of autonomy and its dependence on the interactional relationship between humans and technological actors within a context, we need to understand what the role and responsibility of the designer is to build towards a specific explication of their autonomy. To get a better grasp of what makes up the components of autonomy of a designer, this section explores the practices, responsibilities, and epistemic conditions of design work. Answering what design is has been notoriously challenging, but it is essential to have a consistent and relevant understanding of what is meant when referring to design as a practice and to frame the responsibilities of the design practitioner. Using the definition provided by the International Council of Design (What Is Design?, n.d.), design is a discipline of study and practice focused on the interaction between a person, a “user”, and the man-made environment, taking into account aesthetic, functional, contextual, cultural, and societal considerations. As an academic discipline, design theory has evolved and expanded what is understood about design and the design practitioner.

The expansion of design from a solution-driven craft to a transdisciplinary mode of sensemaking. Rather than treating design as the act of solving given problems, Dorst (2015b) argues that contemporary design is fundamentally concerned with constructing the problem itself. It is an abductive move that redefines what matters, who is affected, and what possibilities are even imaginable. This shift foregrounds the value-shaping role and responsibilities of the design practitioner.

From its origins, design was often understood as a craft-based or applied art discipline, a way of shaping form and function through skilled making. Before formal theories emerged, the designer’s role was largely visual or technical: to give aesthetic and material expression to predefined needs. This view began to shift in the mid-20th century, influenced by developments in systems thinking and operations research. Scholars like Herbert Simon (1969) framed design as a rational, goal-oriented process. In this view, problems were assumed to be given and solvable through structured methods. Design became a process of optimising solutions against constraints, an engineerable, analytical task. This laid the groundwork for the first generation of design methods, built on stepwise models of analysis, synthesis, and evaluation.

However, as design increasingly dealt with complex social, institutional, and technological challenges, this rationalist model began to show its limits. In response, Donald Schön (1983) introduced the idea of the reflective practitioner. He observed that designers often work in conditions of uncertainty, ambiguity, and incomplete information what he called "the swampy lowlands" of real practice. Here, problems are not given but constructed, and design involves iterative cycles of action and reflection. Reflection-in-action became a key concept: designers adjust their framing while designing, guided by surprise, tacit knowledge, and responsiveness to context. Schön’s work helped reframe design as a situated and interpretive practice.

Building on this shift, scholars like (Cross, 2001), (Lawson, 2006) and (Dorst, 2015b) examined how design operates through abductive reasoning, which is generating possibilities rather than deducing or inducing conclusions. This led to the idea of co-evolution, where problems and solutions evolve together through iterative framing. Designers, rather than solving static problems, shape both the definition of the challenge and the means to address it. The recognition of wicked problems (Rittel & Webber, 1973) issues with no clear or stable solution further cemented this understanding.

By the early 2000s, research into design expertise began to show that skilled designers rely not only on logic but on intuition, experience, and judgment. They draw from repertoires of precedent and are sensitive to context, values, and ambiguity (Cross, 2001; Lawson, 2006). This added depth to the epistemology of design expertise was now understood as a form of knowing-in-action, not merely technical mastery. Later developments, especially in speculative and critical design (Dunne & Raby, 2013), as well as interaction with science and technology studies, further repositioned design as a world-shaping, ethical practice. Designers were no longer seen as neutral facilitators but as agents whose decisions shape institutions, behaviours, and futures. This opened up space for design to be framed as a politically and morally consequential practice, capable of both reinforcing and resisting dominant systems.

Term	Meaning
Designer's World	The experiential and conceptual background, including values, past cases, and disciplinary habits, that a designer brings to interpret a situation.
Knowing-in-Action	Skilled, tacit, and often unspoken knowledge expressed in the flow of action. It guides intelligent behaviour without conscious deliberation.
Reflection-in-Action	On-the-spot reflection triggered by breakdowns or surprise. Designers improvise, test, and adjust their framing during the process.
Reflection-on-Action	Retrospective reflection after a design action or project, often used to develop understanding and improve future responses.
Framing	The way that a designer perceives and understands a problem situation. A way of accommodating a new problem by connecting it to prior experience.
Design Expertise	The ability to navigate ambiguity through intuitive judgment, ethical sensitivity, and reflective framing, often built through experience.

Table. 1.: Table of terminologies of the key concepts of the design process.

Term	Meaning
Design	A situated, iterative process of exploring possibilities to respond to complex or blurred contexts.
Co-Evolution of Problem-Solution	The mutual shaping of problem and solution through iterative exploration, each informing the other as understanding develops.
Satisficing	Choosing a solution that is "good enough" in context, acknowledging that optimality is rarely possible in complex or value-laden settings.
Wicked Problems	Problems that lack clear definitions or solutions, where each attempt to address the issue reshapes the problem itself.
Design Frames	The interpreted definition what the situation is about, what values are at stake, and what direction the design should take.

Table. 2.: Table of terminologies of the key concepts of the design process.

2.1.4 Designer Autonomy in the Age of AI

With the understanding that autonomy entails both the authenticity of belief and the agency to act meaningfully, and that design is not the execution of predefined tasks, but a sensemaking process embedded in socio-technical systems, we can now move towards assembling a concrete understanding of the value of designer autonomy. Having the human actor's role and responsibilities outlined by the principles of design practice, an exploration into the technology and its influence on norms of practice will highlight specific characteristics relevant to the value in the age of AI. This section elaborates shifting roles of designers with respect to increased AI integration in practice, its implications, eventually leading to conceptualisation of designer autonomy and its components.

As AI systems enter design workflows, they do not merely offer tools for automation but reconfigure the epistemic and ethical structure of design practice. (Verganti et al., 2020) describe this as a shift from problem-solving to sensemaking. The designer is no longer tasked primarily with generating solutions; instead, they become interpreters positioned to interpreters positioned to explore what is being framed as the problem, for whom, and under what assumptions. In this repositioning, the designer gains epistemic authority but also inherits new responsibilities.

(Shi et al., 2023) expand this view by identifying multiple modes of collaboration: AI assisting designers, designers assisting AI, and co-evolving roles. These modes signal a dynamic interplay, where the designer is neither wholly in control nor fully passive. However, each mode reshapes the scope of design decision-making and the distribution of interpretive authority between human and non-human agents.

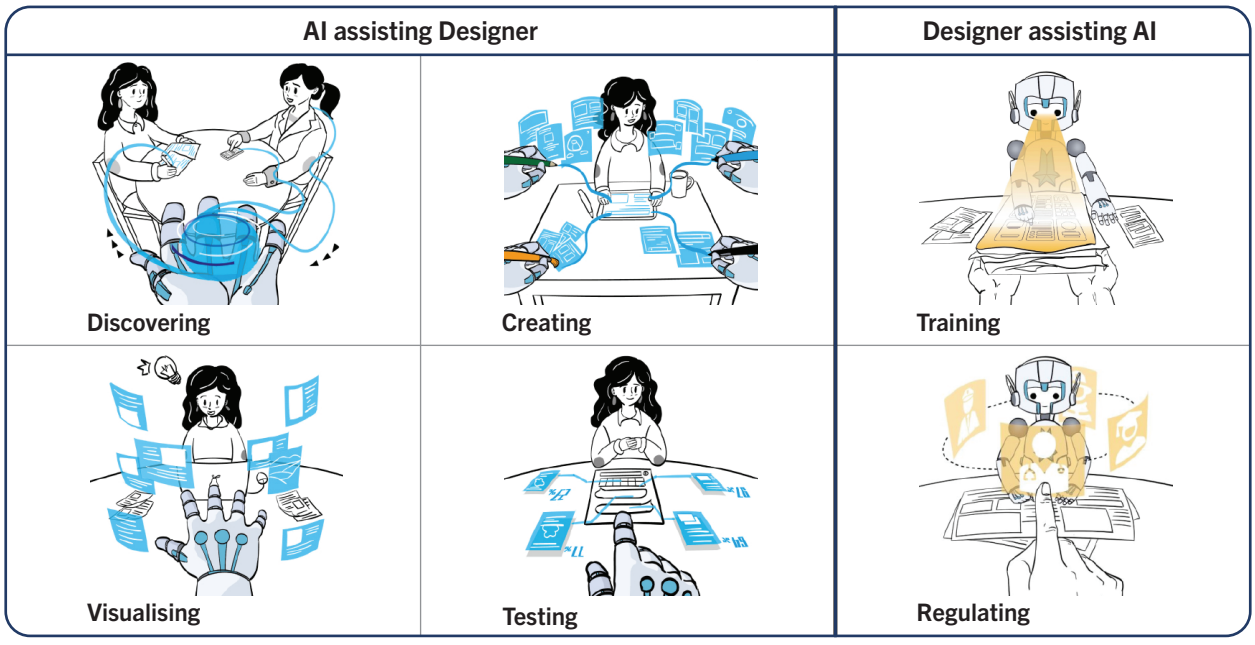


Fig. 4.: Modes of evolving Designer-AI collaborative roles. (*Yang Shi et al.2023)*

This role redistribution carries ethical stakes. AI systems do not merely support decision-making but shape which decisions are possible, which values are legible, and which futures are imaginable. Designers may find their judgment subtly shaped by these dynamics, raising important questions about reflexivity and influence.

Below are some key takeaways from the research in Designer-AI Collaboration:

Shift from Problem-Solving to Sensemaking: Designers are increasingly positioned as interpreters of complex social and technological contexts, tasked with framing the right problems rather than solely generating solutions. This reorientation highlights the designer's role in making sense of ambiguity, while AI supports by surfacing patterns and generating possibilities.

Reinforcement of Design Thinking Principles: AI systems are shown to support and even enhance core Design Thinking principles such as empathy, iteration, and abductive reasoning. They enable more personalised, user-centred outcomes and allow designers to test ideas at unprecedented scale and speed.

Scalability, Adaptability, and Efficiency: AI augments the design process by reducing resource constraints, enabling broader problem scopes, and offering continuously adaptive learning. This shifts the boundary of what is feasible in human-intensive processes and redefines the role of human effort.

Modes of Collaboration: Designer–AI interaction can take multiple forms, such as AI assisting designers (e.g., generating options), designers assisting AI (e.g., training models), or mutual collaboration. These modes represent different distributions of control, responsibility, and authorship, raising questions about how autonomy is preserved or reconfigured.

Complementary Strengths: Productive collaboration depends on recognising and leveraging the distinct capabilities of human and machine agents. Designers contribute domain knowledge, contextual sensitivity, and ethical awareness; AI offers speed, scalability, and data processing. The goal is synergy, not replacement.

Challenges in Trust, Transparency, and Role Clarity: Barriers to successful collaboration include lack of interpretability in AI decisions, difficulty establishing appropriate trust, and ambiguity about each agent's role in the design process. These challenges have direct implications for maintaining designer autonomy and responsibility.

Both (Verganti et al., 2020) and (Shi et al., 2023) conclude by acknowledging that while current research clarifies how AI can augment or reconfigure design tasks, further work is needed to understand the epistemic and value-laden nature of human–AI collaboration. Verganti et al. call for deeper reflection on how AI's capabilities influence not just how solutions are generated but how problems are framed. Shi et al. similarly emphasise the need to investigate designers' sense of control, trust, and role clarity, particularly as AI systems become more generative and suggestive in early-stage ideation.

Designers operate within environments that are structurally complex and morally charged. The contexts they engage, from institutional directives to stakeholder tensions to socio-political ideologies, are not neutral backdrops but active forces that shape what counts as a legitimate design problem, whose voice is heard, and what futures are deemed desirable. As established in the previous section, framing is the central act through which these forces are interpreted, accommodated, or resisted. The designer, then, is not merely a facilitator of creative solutions but a mediator of conflicting value systems. Their autonomy is not exercised in abstraction, but in the thick of these entangled relationships.

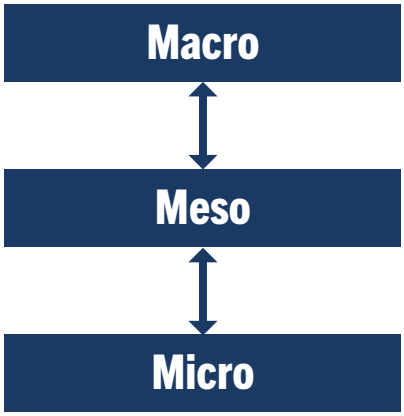


Fig. 5.: The three interdependent Micro-Meso-Macro levels of value frameworks. *(Prendeville et al., 2022)*

In this layered landscape, the structure of micro, meso, and macro value frameworks (Prendeville et al., 2022) becomes particularly relevant. At the micro level, designers bring their own beliefs, intuitions, and ethical positions, shaped by disciplinary training, personal experience, and cultural context. At the meso level, they encounter organisational agendas, team dynamics, and stakeholder expectations that condition what kinds of design decisions are considered acceptable or viable. At the macro level, broader systemic structures like economic systems, social ideologies, policy regimes frame the boundaries of what is even imaginable. Design decisions, in this light, are not simply practical or creative moves, they are value choices that materialise through the negotiation between these framing layers. And it is precisely in this negotiation, how the designer responds to, adapts, or challenges dominant frames, that autonomy, as both agency and authenticity, is made visible and expressed.

The mapping of Prunkl's two components of autonomy onto the design process offers a way to understand designer autonomy not as a personality trait or static capacity, but as a dynamic ethical practice. Authenticity in this context refers to the designer's ability to recognise, articulate, and act from their own critically examined beliefs, not simply reproducing defaults or inherited positions. Agency refers to their capacity to shape the design direction, resist reductive or coercive framings, and take epistemic and moral responsibility for outcomes. When either component is undermined, when the designer acts without reflection or cannot meaningfully influence outcomes, autonomy is diminished.

A designer's autonomy is exercised in relation to the other actors around them, shaped by the technological tools, collaborative structures, and power dynamics the designer must navigate. It is relational, unfolding in constant interplay with technological systems, institutional constraints, and the needs of others. It carries moral weight, in that it entails responsibility to make decisions that are justifiable, inclusive, and responsive to the broader implications of design. This gradual building of conceptual understanding allows a final description of designer autonomy and its components.

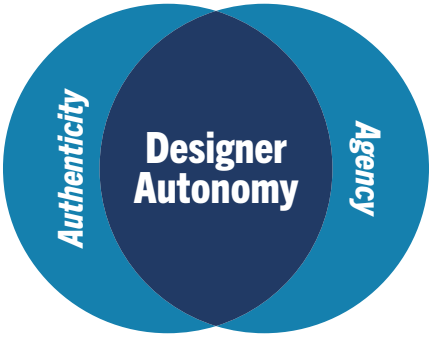


Fig. 6.: The dimensions of Designer Autonomy in the Age of AI, Authenticity and Agency.

Designer Autonomy is the value of the designer to approach problems and develop solutions in morally, epistemically and politically responsible ways. It is enacted through the reflective negotiation of values across individual, organisational, and systemic frameworks, maintaining authenticity in belief and agency in actions in value-laden design environments.

Designer Authenticity refers to the capacity of a designer to critically reflect upon and act in accordance with their own beliefs, values, and ethical commitments within the design process.

Designer Agency refers to the ability of a designer to exert meaningful influence over the design direction to act intentionally, challenge constraints, and reshape frames in ways aligned with their reflective commitments.

Designer authenticity is not simply about expressing personal choice or intuition, but about ensuring that the choices a designer makes are recognisably their own shaped by deliberate judgment and not default conventions or inherited assumptions. Authenticity demands that the designer's framing of the design situation arises from conscious engagement with their moral, epistemic and political stance, rather than being unconsciously shaped by external pressures, trends, or tool biases. It is fundamentally about self-recognition in one's continuous, iterative reasoning.

Designer agency includes resisting dominant narratives, redirecting abductive trajectories when deemed necessary, or creating alternative possibilities when default pathways foreclose certain futures. Agency is not reducible to control or power but can be enabled or constrained by exerting or curtailing it. It is the capacity to act with responsibility, particularly in conditions of complexity, ambiguity, or automation.

A designer exercises agency when they are able to examine assumptions, propose counter-frames, and shape decisions that reflect considered judgments, even when these go against prevailing expectations or systemic inertia. Agency should not be reduced to simply contrarian tendencies, but rather but rather understood as a constructive form of critical participation.
Contextualising the insights from Franssen (2015) to design, where the operator effectively refers to the design practitioner, to the insights from the evolving role of designers in the age of AI, the implications to the components of designers' autonomy can be articulated as follows:

Implication to Designer Authenticity

Operator Roles as Partially External Components: Designers functioning as partially external components may find their personal beliefs sidelined, challenging their authenticity in decision-making.

Epistemic Enslavement: When designers are required to follow rigid protocols without understanding the rationale, their ability to act authentically, based on personal judgment and values, is compromised.

Sensemaking Role: As designers focus on interpreting contexts and defining problems, they must align these interpretations with their personal and professional values, ensuring authenticity in their contributions.

Implication to Designer Agency

Operator Roles as Partially External Components: The external control and monitoring of designers limit their ability to make autonomous decisions, thereby constraining their agency.

Epistemic Enslavement: Establishing clear roles in designer-AI collaborations empowers designers to act intentionally, ensuring their contributions are impactful and aligned with their expertise.

Delegation to AI: Designers exercise agency by deciding which tasks to delegate to AI and which to retain, influencing the design process's direction and outcomes.

Responsibilty Assignment: Assigning responsibility without corresponding authority or autonomy undermines designers' agency, as they are held accountable without the power to influence outcomes effectively.

Being an autonomous actor participating in the design process, the designer's role, responsibilities, knowledge, identity and effectiveness is relevant when understanding the value tension against automation.

2.2 Situating the Value in Framing Practice

This part proposes framing practice as the site where the value of designer autonomy is observable and malleable when collaboration with AI. The goal of this section is to understand the development of framing practice in design, and how it is relevant to the shaping of design outcomes. This will enable to create a conceptual framework of how autonomy is visible and malleable in design processes, which will then inform the practical approach to the qualitative research for the empirical investigation.

2.2.1 Framing Practice in Design Theory

Within literature about design theory, framing emerges as a critical element of reflective practice and is considered to be the core of the design thinking approach (Dorst, 2011). The concept of framing does not originate in design theory but in the social science, when Erving Goffman (1974) introduced framing as a way of understanding how individuals interpret and make sense of social reality. Frames, in his theory, are cognitive structures that help people identify "what is going on" in a situation. In media studies and political communication, scholars like Entman (1993) showed how institutions use framing to highlight specific aspects of an issue while downplaying others, effectively shaping public understanding through selection, emphasis, and exclusion. The practice of framing thus became recognised as a selective, meaning-making process deeply embedded in social power.

Donald Schön (1984) later introduced framing into design discourse. In his theory of reflective practice, framing is the act of defining what the problem is, what matters, and how it should be approached and often revised in action. Designers engage in reflection-in-action when confronted with uncertainty, improvising their framing based on interaction with the situation. This transformed framing from an implicit or external act into a core cognitive operation within design: the way designers give shape to situations that are not yet understood.

Subsequent theorists such as Cross, Lawson, and Dorst linked framing to abductive reasoning, proposing and testing tentative problem definitions based on emerging possibilities. This work formalised the concept of co-evolution between problem and solution spaces: as one changes, the other does too. Framing was now a constructive, iterative activity, not a pre-design step. Still, it was largely treated as an individual cognitive skill. As a practice, the process of framing occurs in co-evolution, in the movement towards a solution frame from a problem frame (Goss et al., 2025). For an intended goal or outcome (value), framing connects the things that comprise the focus of the realisation of the value (what) with a working principle of action to achieve it (how) (Kelly & Gero, 2022).

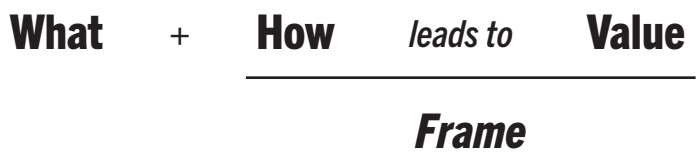


Fig. 7.: Design reasoning framework adapted from Dorst (2011).
(Goss et al., 2025)

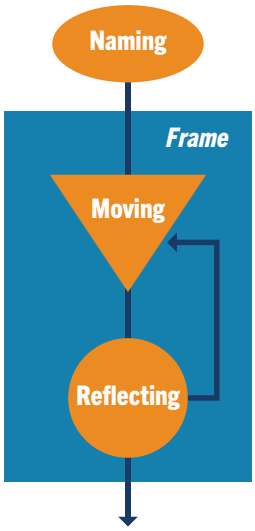


Fig. 8.: The mechanism of reflective practice, the key design activities of Naming, Moving and Reflecting and their interplay. *(Valkenburg & Dorst, 1998)*

Later, design researchers focused on framing as an expression of situated expertise. Scholars like van der Bijl-Brouwer (2019) explored how expert designers draw from prior cases, disciplinary logics, and intuitive judgment to frame problems. Framing was increasingly recognised as the designer's primary site of judgment and impact. However, it was still often described in neutral terms in literature, which came under critique in more recent literature, especially in the work of (Prendeville et al., 2022).

Drawing on critical theory and science and technology studies, they reframed framing itself as a political act. Referring to designers as material-semiotic storytellers (Wilkie & Ward, 2009) whose framings do not just interpret the world but actively shape it by deciding which futures are imaginable and which are ignored. Their work compiled a three-level model of design framing practices to enable an understanding of its political foundation. The first being Agentic, Subjective, and Cognitive frames, which they call Micro-frames, are grounded in cognition theory, psychology, and linguistics. These refer to the individual designer's "frames of reference" or internalised biases, personal values, prior knowledge, and intuitive reasoning structures that shape perception and decision-making. This is a conceptual parallel to Schön's articulation of the notion of "design worlds" (Schön, 1988).

Micro-frames direct how information is processed and are central to understanding the role of individual agency in framing, including how frames reproduce or resist social positions, ideological assumptions, and normative expectations. The second level, Strategic and Collective Action frames, or Meso-frames, focuses on group-level processes, particularly how design teams, stakeholders, and institutions collectively construct and align interpretations through interaction. Rooted in social movement theory, these frames are established via negotiation, cultural context, and strategic mobilisation. They help explain how specific value interpretations of the designer(s) gain traction, how they resonate with the broader public, and how they can facilitate counter-framing practices within project or organisational settings.

Finally, Institutional frames, or Macro-frames, emerge from the structural and discursive formations of institutions and dominant ideologies. These include enduring cultural norms, policy regimes, and systemic logics (such as market liberalism or technological determinism) that shape the boundaries of what is considered legitimate or even thinkable in design. These frames are the most deeply embedded, often taken-for-granted, and typically require symbolic as well as material strategies to resist or reconfigure. Importantly, the research emphasises that these levels are not isolated but intersect dynamically in everyday design activity. Framing is always situated, contested, and shaped through the interaction of individual, collective, and systemic forces.

Applying the learnings from exploration of design and framing practices, the three levels in terms of value frameworks which influence the value embedding decision-making across the design process can be simply understood as follows:

Micro: the designer's own beliefs, assumptions, and ethical stance.

Meso: organisational, team-based, or institutional framing logics.

Macro: structural ideologies, social narratives, and power systems.

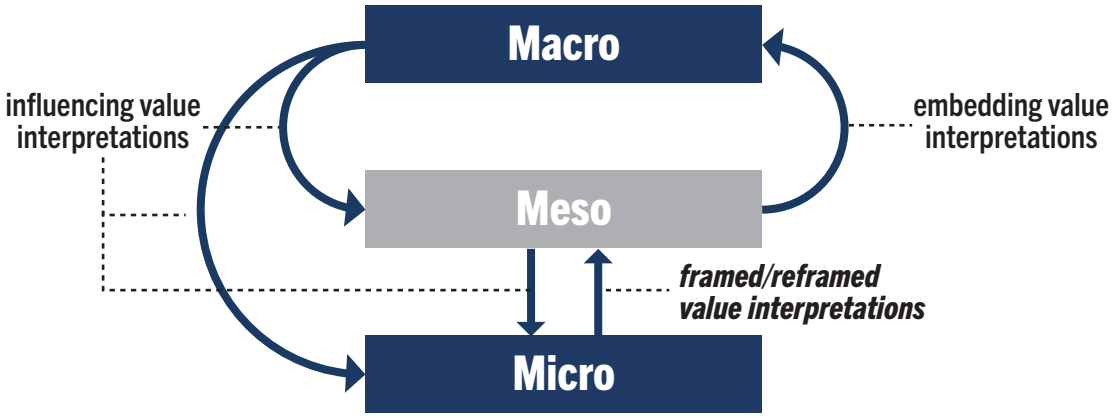


Fig. 9.: Interaction of the Micro-Meso-Macro levels of value frameworks and the position of framing practice in embedding value interpretations within it.

This model enables a descriptive analysis of how values are distributed and negotiated in design, not just internally within the designer, but across systems of power. Framing becomes the ethical interface of design, the act where autonomy is expressed, challenged, or suppressed. In this view, framing is not just reflective or strategic, it is also moral and political. The reflective practice model (fig. 5) describing a framing at an individual or cognitive level, while the multi-level interactional framework (fig. 6) situating the practice of framing at the centre of value interpretation and embedding role of designers. Therefore, it is proposed that autonomy of a designer is both expressed and challenged at the micro level, influenced by meanings embedded in meso and macro levels.

2.2.2 Framing Practice in the Age of AI

As understood in discussions about the shifting role of designers, AI technologies increasingly mediate key elements of the design process. They support designers by accelerating iterations, generating insights, and surfacing patterns. They also bring with them risks as integration may obscure reasoning, reinforce existing value structures, or influence the perceived boundaries of the solution space through defaults and data-driven framings. These effects are especially salient in strategic design, where framing is not merely methodological but ethical, where designers define what is at stake, for whom, and why.

Framing, to reiterate, is not a neutral or preliminary act. It is the interpretive move through which designers make value judgments, construct meaning, and determine direction. When design decisions are mediated by systems whose logic is opaque or guided by historical bias, the designer's authenticity may be undermined. Similarly, their agency may be compromised when AI narrows perceived possibilities or pre-structures tool environments to favour certain directions of interpretations of what is valuable, how and to whom.

Dorst's (2015b) account of frame creation positions framing as a design act grounded in abductive reasoning: beginning with values and working backwards to imagine new actions. In this view, framing is a process of negotiation with context, requiring constant reflection and adjustment. Valkenburg and Dorst's (1998) earlier studies reinforce this, showing that framing is both material and interactive, emerging through conversation, sketching, and reflection-in-action. These accounts suggest that framing can be observed empirically, not just theorised.

Framing practice can be established as the site where designer autonomy is exercised and influences the nature of outcomes of the process. The designer does not merely choose from given options but participates in constructing the very space in which options become visible. As Prunkl (2022) outlines, autonomy comprises both authenticity and agency, components which in design practice manifest as the capacity to reflectively interpret problems and intentionally shape direction. Thus, framing is where these capacities are exercised, and where they come under pressure in AI-mediated environments. The introduction of AI systems does not eliminate autonomy but makes its exercise more complex as designers are increasingly required to account not only for stakeholder values and contextual constraints, but also for how the system itself frames the design space.

As highlighted previously, framing is both an epistemic condition of practitioners as well as an operationalised methodology prescribed to reinforce and externalise the tacit approach in practice. This distinction can be understood as framing as an approach, a cognitive model and framing as a practice, specific design methods to create frames as part of the design process. Kelly and Gero (2022) address this complexity in understanding the position of framing in design theory. Framing as an approach is referred to a theoretical lens through which the cognitive behaviours of designers are understood. This highlights how problem representations, goals, and constraints are actively constructed rather than passively received.

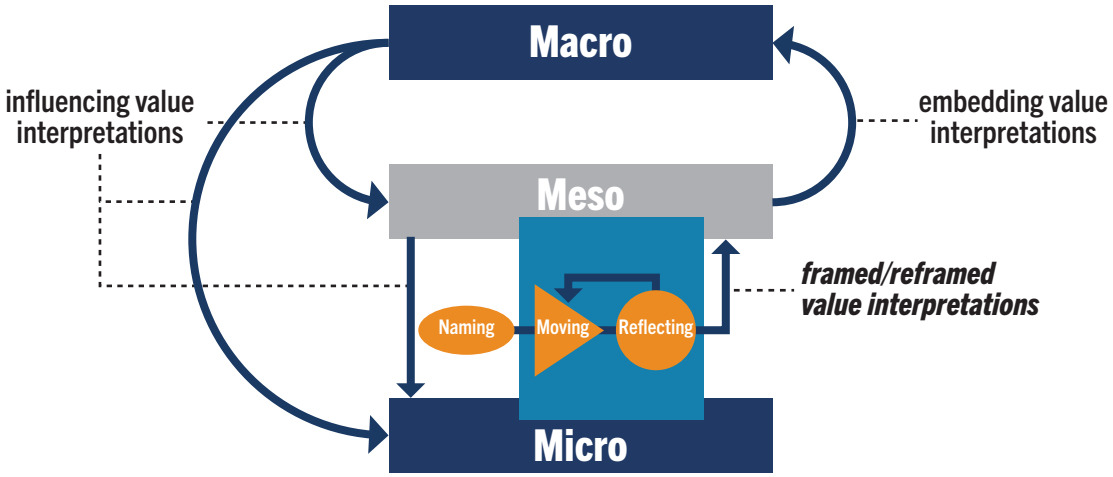


Fig. 10.: Interpretive position of design framing practice between Micro and Meso levels of thein interactional multi-level value frameworks model.

With the awareness of the broader multi-level interactional model as well as models of designerly actions comprising framing processes, it can be interpreted that a particular framing activity is situated between the Micro and Meso levels (as illustrated by figure 10).

Newer understandings of framing view design cognition as frame-driven, where framing shapes what is seen, valued, and pursued during the design process. In contrast, framing as a practice involves deliberate techniques or interventions that guide how these cognitive processes are enacted, such as reframing strategies, sketching, or stakeholder engagement methods. These practical tools externalise and stabilise internal frames, enabling designers to collaboratively shape and revise problem representations in situated contexts. Therefore, it can be concluded that framing as a cognitive approach is shaped by, and at risk from exogenous factors, and framing as an operationalised practice, normalises value interpretations by becoming the endogenous factor which embeds values in design outcomes.

2.3 Designer Autonomy in Collaborative Framing Practice

To conceptualise how designer autonomy is exercised in collective and technologically mediated practice, this project introduces the term collaborative framing or co-framing. Earlier design theory has described the iterative interplay between problem and solution spaces as co-evolution, the descriptions of process of framing do not explicitly factor in multiple actors participating in co-evolution. Co-framing focuses instead on the moral and epistemic dynamics of framing itself, specifically when that process involves multiple actors, both human and non-human. It captures how design frames are negotiated, resisted, or reasserted as designers interact with AI systems that increasingly participate in the sensemaking process. Co-framing is proposed as the primary site of ethical and epistemic work in design. It is here that designers engage with how their tools shape interpretations, distribute responsibility, and surface or obscure values. Autonomy, then, is not a background condition of practice, but a continuously exercised in acts of negotiation within increasingly complex socio-technical systems. A designer working with AI more than simply guiding a process, but engaging in a dialogue with a system that comes embedded with assumptions, patterns, and norms.

Co-framing is the collaborative construction and negotiation of value interpretations between human and artificially intelligent actors in a design sense-making activity.

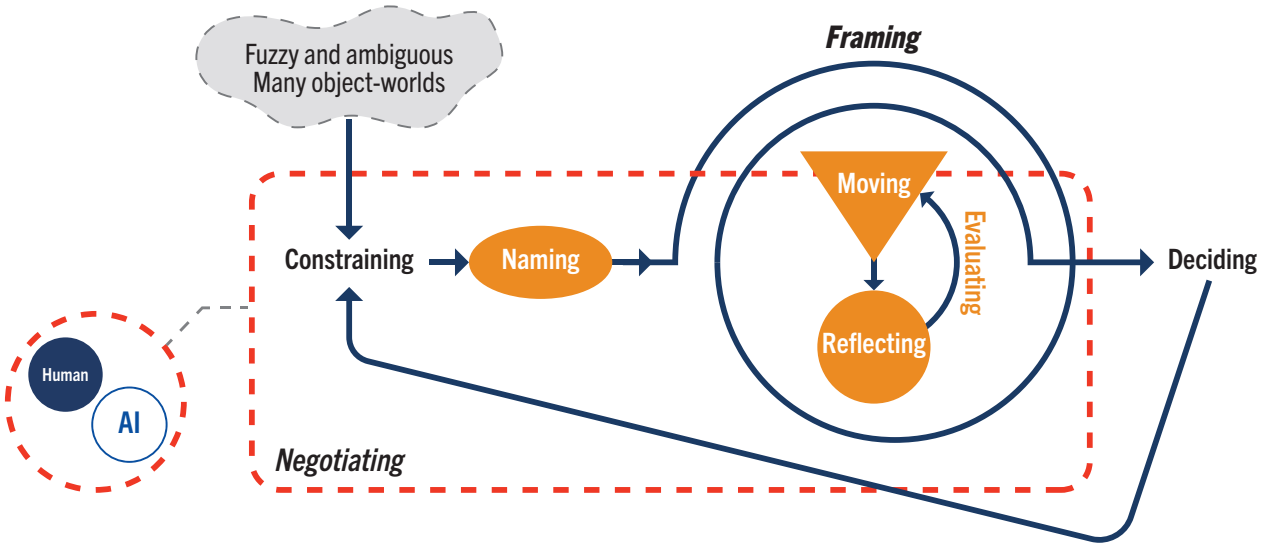


Fig. 11.: A model for Collaborative Framing Practice, adapted from the mechanism for framing in co-reflective practice *(Zahedi et al., 2016)*.

Building on research about collaborative reflective design practices and framing dynamics of design teams (Zahedi et al., 2016), co-framing explicitly introduces the presence of artificially intelligent systems as active collaborators to the framing process along with the human designers. This explicates the role and responsibilities of the designer to be sensitive about the factors influencing their value interpretations, not just from “meso” and “macro” factors, but also from collaborating actors and their own “micro” frameworks. While an AI does not represent an autonomous agent in terms of equal capacity to a human designer, research highlights there is significant potential for the technology to influence human behaviour.

It can be understood that negotiation, and thus, framing activity, in the model of co-reflective practice is done by multiple human designers. The aim of creating a new terminology of co-framing is to make the collaborating actors and their interacting individual frames (micro) explicit, and to accommodate the integration of AI as a meaningful contributor to shaping the individual's interpretation in action.

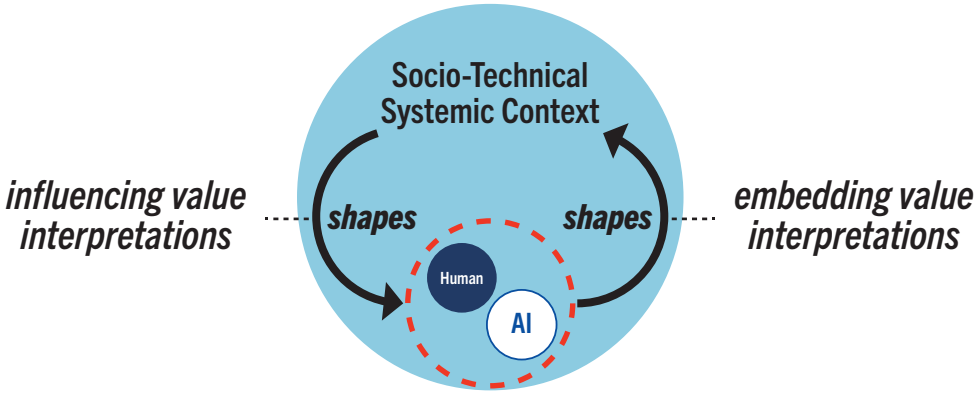
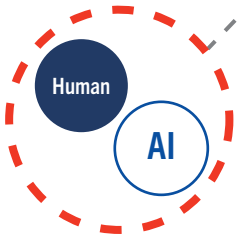


Fig. 12.: Collaborating human and AI actors in the design process
in their interpretive role of values in a socio-technical systemic context.

AI systems influence each of the three value framework layers by shaping internal reasoning through suggestions, structuring team practices through tool workflows, and reinforcing systemic norms through pre-trained defaults. Co-framing brings attention to how designers engage with questions of what the situation is about, whose values are being prioritised, and whether to align with, resist, or reinterpret the system's framing.

Bringing back the lens of descriptive ethics to the mechanism of co-framing, we can derive four potential scenarios in which the dimensions of a designer's autonomy can be either supported or undermined in the interaction with an AI system. These scenarios are not prescriptive ideals but observable patterns of behaviour which does not necessarily guarantee ethically correct outcomes. Supporting autonomy within an activity is not inherently good, and neither should undermining the designer's autonomy be understood as inherently bad. This hides an implication that the morality of value interpretations held by the human are good, or the human actor always wants to do good. These scenarios describe what may be occurring and eventually influencing the resulting decisions. Once applied to specific cases for analysis, it can be determined whether that particular interaction of either supporting or undermining autonomy led to desirable outcomes. But this should not be extrapolated to other interactions without considering the numerous nuances contributing to the value framing activity. Below is a matrix which shows possible scenarios influence on the dimensions of autonomy of the human and AI interaction.



	Supports Agency	*Undermines* Agency
Supports Authenticity	Designer develops own view of the problem and directs framing	Designer develops own view but defaults to AI's framing instead
Undermines Authenticity	Designer frames based on AI-shaped assumptions, without questioning	Designer accepts AI's framing without forming their own view

Fig. 13.: Nature of the potential outcomes of the interaction of the co-framing actors with respect to the components of autonomy.

To reiterate, autonomy is not exercised solely through control or resistance but through critical engagement with system inputs. It may become visible in instances where designers reject or reframe system-driven defaults based on their reflective interpretation of values and context. Equally, autonomy is present when designers intentionally adopt or propagate system suggestions not as passive acceptance, but as a result of deliberate alignment with their ethical judgment and project aims. In both cases, the designer retains epistemic agency by actively evaluating how the system contributes to the framing of the design space.

Co-framing is about shared negotiation, judgment, and contestation between the human and technological actors. It describes a form of engagement in which the designer remains epistemically and morally present, aware of the influences shaping their decisions and maintains interpretive agency in light of those influences. This reframing of the human-AI relationship expands the focus beyond practical tool use to value negotiation, from automation to responsibility. Ultimately, co-framing offers a lens through which to observe and analyse designer autonomy in AI-mediated settings. It provides the conceptual and empirical ground for understanding how values are made visible, negotiated, and acted upon in framing practices. This lens frames the designer not as a passive operator or adaptive collaborator, but as a reflective and accountable agent involved in shaping design futures.

CHAPTER 3

EMPIRICAL INVESTIGATION

The role of the empirical investigation in the adapted Value Sensitive Design approach of this project is to understand the context and experiences of the primary stakeholder, and the shaping of value under observation in relation to the technological interaction. The conceptual investigation supplies this section of the project with a foundation of the value, and the site in which it is shaped in the design process. This section explains the qualitative methodology undertaken, approach to data collection and analysis of the study, concluding with the findings of the analysis for the empirical research question.

3.1 Autoethnography of a Co-Framing Activity

An autoethnography is chosen to empirically investigate designer autonomy as a value shaped and exercised in collaborative framing activity in a design process. An ethnography as an analytical tool is equally a test of the ethnographer as it is a study of the data (Fetterman, 2020). In an autoethnography, the dual role of the ethnographer as both the principal researcher and primary subject allows for benefits of having to reflect on one's own lived experience while also being having complete awareness of the subject matter. This approach enables a personal analysis of experiences, allowing for a direct line of reasoning to reveal drivers of actions, as it benefits from a focus on specific contextual information about an individual's perceptions of personal and societal factors (Cooper & Lilyea, 2022).

This approach as a qualitative research methodology specifically suits the nature and objective of the inquiry of designer autonomy as a value in the age of AI. The autonomy of the designer, as concluded in the conceptual investigation, is not exercised in isolation but emerges through engagement with evolving value landscapes. Autonomy here is understood as a situated capacity to frame design problems in morally and epistemically responsible ways. This view rejects static or procedural notions of autonomy, instead placing emphasis on the designer's ability to navigate, interpret, and negotiate values that are often in flux. The concept of value dynamism is central to this account: values are not fixed inputs to be implemented, but live elements shaped through practice, context, and reflection. In design situations mediated by AI, this dynamism becomes even more pronounced, as system suggestions, embedded norms, and data structures introduce tensions that the designer must actively work through. Autonomy, then, is not the absence of constraint, but the capacity to engage meaningfully and reflectively with the constraints and affordances of a dynamic value environment.

The focus of the autoethnography will be to capture how these negotiations unfold in practice between the human and AI actor. This thesis introduced the concept of collaborative framing, referring to the distributed process through which multiple agents, human and non-human, jointly shape the definition and interpretation of a design problem. Unlike traditional co-design, which often centres on collaboration toward solutions, co-framing highlights the pre-decisional moment where values, assumptions, and problem boundaries are still fluid. It is at this point that the designer's autonomy is most visible and most malleable, as framing decisions both express the designer's reflective stance and set the trajectory for downstream design commitments. In AI-mediated settings, co-framing draws attention to how algorithmic suggestions, data structures, and interface defaults participate in shaping what is foregrounded or excluded in problem definition. This, combined with the iterative nature of autoethnography data collection, should reveal novel insights about the empirical component of the value sensitive design research. Therefore, for the study of collaborative framing in a design activity, autoethnography allows for an effective structure to capture underexplored challenges of designing with AI (Olasik, 2023).

Strategic design is chosen to situate the study in a specific discipline of design. Design has long been implicated in the shaping of sociotechnical systems, with traditions such as sociotechnical systems design foregrounding the interdependence of technical artefacts and social arrangements. However, as organisational and technological complexity has deepened, the focus has shifted from balancing subsystems to actively interrogating how systems are framed, structured, and steered. Strategic design inherits this systems-level ambition but repositions the designer as a critical agent in shaping how values, futures, and institutional logics are negotiated. It is not only about harmonising the social and the technical but about making explicit the framing conditions under which such harmonisation is defined. For this reason, strategic design provides a suitable and revealing context in which to study the nature of designer autonomy, particularly under the influence of AI-mediated framing tools (Pasmore et al., 2019).

Based on the learnings from the conceptual investigation, "collaborative design activity" from the initial research questions can now be understood to be "collaborative framing practice". Adapting the research question for the empirical investigation component of this project to be answered through the autoethnography is:

How is designer autonomy is shaped in ***collaborative framing practice*** with AI?

The conceptual investigation also provides a further bifurcation of this question into two sub-questions; each focussed on the interconnected yet distinct components of agency and authenticity:

How do designers ***(a)*** negotiate and exert their ***agency***; and how do they ***(b)*** perceive and articulate their ***authentic*** influence during a collaborative framing activity?

3.1.1 Strategic Design as the selected Design Domain

Strategic design has emerged as a response to the growing complexity of societal, organisational, and technological systems. It is rooted in the evolution of design a problem-solving practice to design as a mode of systemic intervention but focussed on situating it in organisational problem-solving landscape. The shifting role of design within organisations: from an activity situated at the end of development pipelines to one increasingly integrated into upstream decision-making ("Strategic Design," 2016). This shift reflects the recognition that many of today's challenges, climate change, digital transformation, social inequality, are not merely technical problems requiring solutions but deeply interdependent, value-laden issues that require the reframing of problems themselves.

Strategic design is a positioning of design approaches and processes within complex social, cultural, technical and economic systems. It applies design reasoning and practices to shape organisational vision, strategy, and structures. Practitioners of strategic design work with ambiguity, stakeholder multiplicity, and institutional friction. Rather than solving for predefined problems, they operate in contexts where both the nature of the problem and the criteria for success are contested or fluid. In this way, strategic design becomes both reflective and projective, it reflects on how issues are understood and simultaneously projects new possibilities for action.

Manzini and Coad (2015) describe this as a design mode that moves beyond artefacts to intervene in the "emerging forms of life" within socio-technical systems. Here, design becomes a tool for shaping conversations, aligning stakeholders, and imagining futures through concrete engagement with organisational and societal dynamics. Strategic designers often operate across disciplinary boundaries, facilitating co-creation, mediating between conflicting interests, and translating abstract intentions into actionable pathways. Their work is marked by framing, negotiation, and iteration carried out with an acute awareness of the political, cultural, and economic systems that constrain or enable transformation.

This shift has led to both conceptual expansion and disciplinary tension. Strategic design draws from systems thinking, innovation theory, futures studies, and organisational development, positioning itself as a mediator between high-level strategic intent and practical implementation. However, this broad scope has also led to ambiguity. (Dorst & Watson, 2023) argue that strategic design risks being overstated or misunderstood, particularly when its epistemological foundations are not made explicit. In their critique, they emphasise that while strategic design aspires to combine top-down strategy and bottom-up design practice, this synthesis remains underdeveloped. Without clear articulation of what design brings to strategy beyond facilitation or visualisation, the field may fall short of its transformative ambitions.

Strategic designers, therefore, are uniquely positioned to illuminate the phenomenon under investigation in this study. Their practice routinely involves navigating heightened ethical and political stakes, where decisions shape not only artefacts but also systems, behaviours, and futures. Unlike other design domains where problem spaces may be more bounded or outcomes more predefined, strategic design operates across levels of influence, requiring practitioners to engage explicitly with a broad range of stakeholders, value conflicts, and institutional constraints. This makes the responsibilities and judgments of the designer more visible, and the framing process more deliberate and consequential.

3.1.2 Problem Framing as the selected Co-Framing Activity

For the specific framing activity to be the focus of the autoethnography, developing a problem frame is chosen. The problem framing phase allows for clear application of the developed model of collaborative framing between human and AI actors, evolved from the model of framing in design teams. Developing a problem frame is a critical point in a design process, as it attempts to capture insights that the designer's deem relevant to be framed in a way which they feel will lead to effective ideation and consequent design steps. As explored previously, a phase like this is where designers' autonomy is exercised in an observable way, as their authentic beliefs and agentic capacity inform and influence how they go about accepting or rejecting framing forces in the design task. The outcome of this is one where value-laden actions materialise into decisions which fundamentally alter the design direction. A problem statement fits these criteria as the framing will be embedded with interpretations of the participating actors, which in this case are the principal researcher and ChatGPT (Model GPT-4o).

In the context of the nature of cases dealt by strategic designers, we can look towards research on public and social innovation. Studying adaptation of problem framing practice when co-evolution of problem and solution spaces in cases of complex societal issues, it was concluded that problem framing as a practice could be linked to a shared expertise. When applying the abductive approach of design framing to complex design process, it is argued that the process becomes a fundamentally different practice to traditional deductive reasoning present in such domains (van der Bijl-Brouwer, 2019). Mere transposition of a practice onto a new domain does not guarantee its success, the process has to adapted to the specific need of the context (Dorst, 2015a). For the purposes of the autoethnography, we can use learnings from this research on public and social innovation. This explains the selection of problem framing as the activity in focus and informs the case selection to be that of a complex public/social issue, approached with the lens of a strategic design.

3.1.3 Public Transportation in Mumbai as the selected Case

The selected case was a previously completed systemic design research project in which the principal researcher played a key role. The academic project was done in the first year (2019) of Bachelor of Design education, in a module focussed on design research methods. The design context is that of public transportation of Mumbai, with the research being open ended, vaguely in the direction of capturing user experiences of various modes of transport in the city. The team of 3 design students conducted interviews and field observations of users and employees of trains, metro and bus networks. In this iteration of the project, the problem frame was formulated around the experience of a new traveller of the local train network, setting aside the other two researched modes for reasons of future feasibility and benefits of having a more specific case for concepts and prototype development. With the firsthand experience of this project, specifically about early understanding of the roles and responsibilities of a designer, this case provided a good contrast when compared to the present understanding of the same. With differing practical objectives, the autoethnography of the same task was done with no practical limitations informing feasibility or viability. With there being rich data from multiple modes, there was no constraint on what the solution direction should be. This led to the objective of developing a problem frame which should allow for a novel approach to addressing the shortcoming of public transportation services in Mumbai.

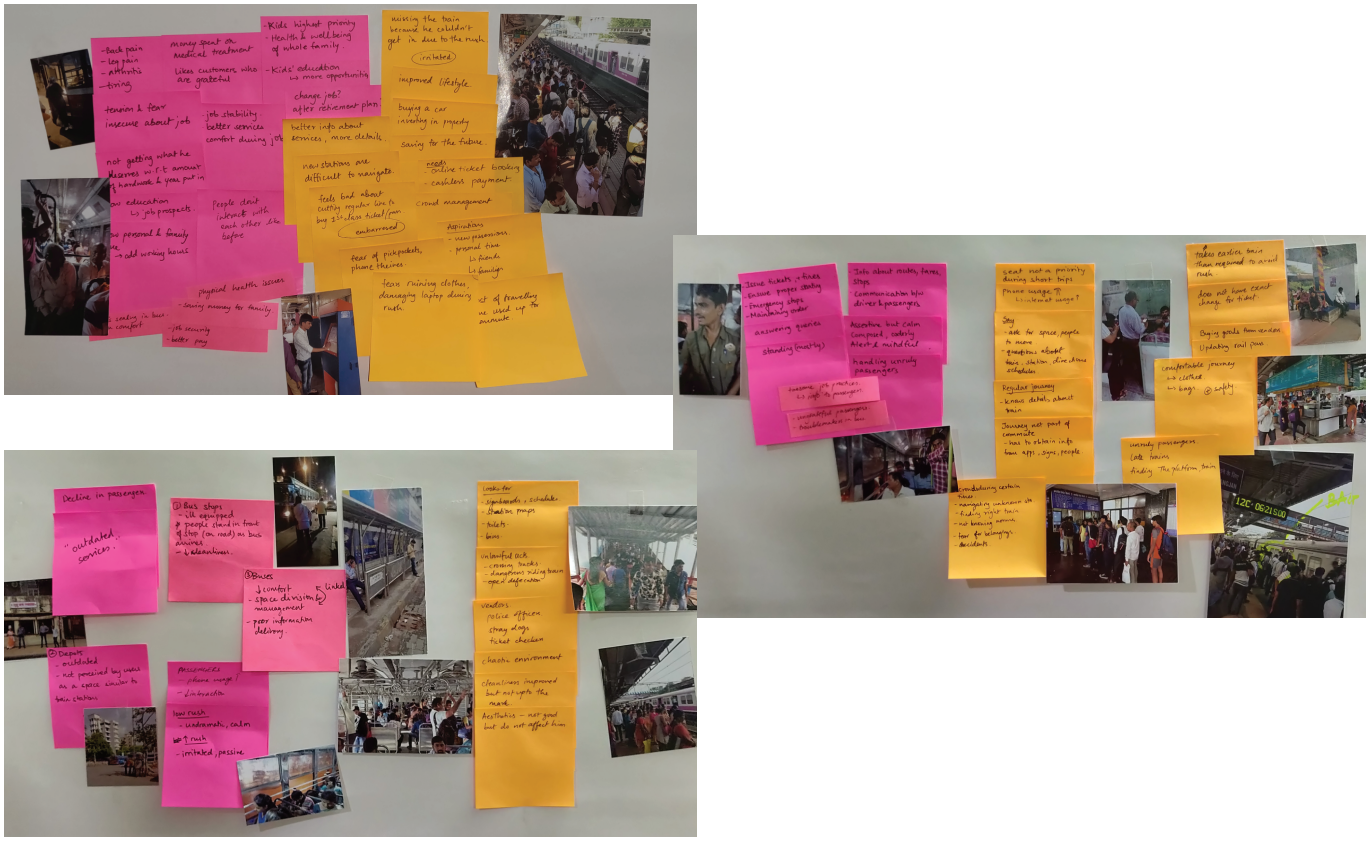


Fig. 13.: Design Research Highlights from the orginal selected case of Public Transportation in Mumbai.

3.2 Data Collection, Analysis and Findings

The autoethnography task was spread across 5 days or sessions, roughly replicating the time span provided to the team in the original project. While the task of developing a problem statement could be done in a shorter amount of time, this approach allowed for intuitive separation of sub-phases such as familiarising, analysis, statement development, etc., with ample time for reflection within each task as well as post. Disconnecting from the AI in between sub-tasks also allowed for longer periods of time to reflect on the previous session and to come up with a direction for the next one. The data used in this study was the anonymised data points for primary research, grouped by users and employees of trains, metro and bus, 6 sets. These data points are quotes or observations as noted by the researchers of the initial project. All data used in the study is anonymous and cannot be used to link to a source individual. The original data generated as part of the autoethnography are in the form of (1) logged conversations between the researcher and AI tool (2) In-task Reflection reports (3) Post-task Reflection reports. The reflections formed the core of the data analysis, with the logs of the conversation supporting entries in the reflection while also providing an opportunity to detect a disparity in actions compared to the reflection. The stopping point of the process was the development of a problem statement along with supporting design constrains as a result of the collaborative framing activity.

Data analysis followed an inductive approach to conduct a reflexive thematic analysis to interpret the reflections and logged conversations. This methodology for data analysis allows for high degrees of flexibility to work across qualitative research methodologies (Braun & Clarke, 2006), thus enabling an application to the autoethnography. The analytical approach to uncovering relevant answers to research questions in an autoethnography relies on iteratively understanding personal reflections and descriptive facts (Emerson et al., 2011). Reissman (2008) highlights that emphasis of thematic analysis of autoethnography can be approached with varying emphasis on historical context, social theory and coding.

Frameworks of thematic analysis for autoethnographies exist, offering procedures which focus on tools to group events, developing sub-plots (Gibbs, 2008), or suggestions to categorise based on priority, explore contradictions and develop metaphors (Janesick, 2010). Recognising the role of the autoethnography in this broader project to answer the empirical investigative question, this thematic analysis follows a simplified yet comprehensive procedure. This study begins with organisation of generated data, followed by open coding. Original data was first collated into individual documents, in preparation for thematic analysis. Open coding was manually performed to surface patterns in the data, informed by the literature review's account of designer autonomy as comprising both authenticity and agency. Conceptual clusters surfacing from the first pass were then refined through axial coding to develop a set of final of themes which answered the empirical research question of what factors shape a designer's autonomy in interactions with an AI in a collaborative framing task. To ensure the credibility of the study, findings and interpretations were reviewed by the supervisory team. A reflexive approach will be maintained throughout the analysis to support the trustworthiness and applicability of the results.

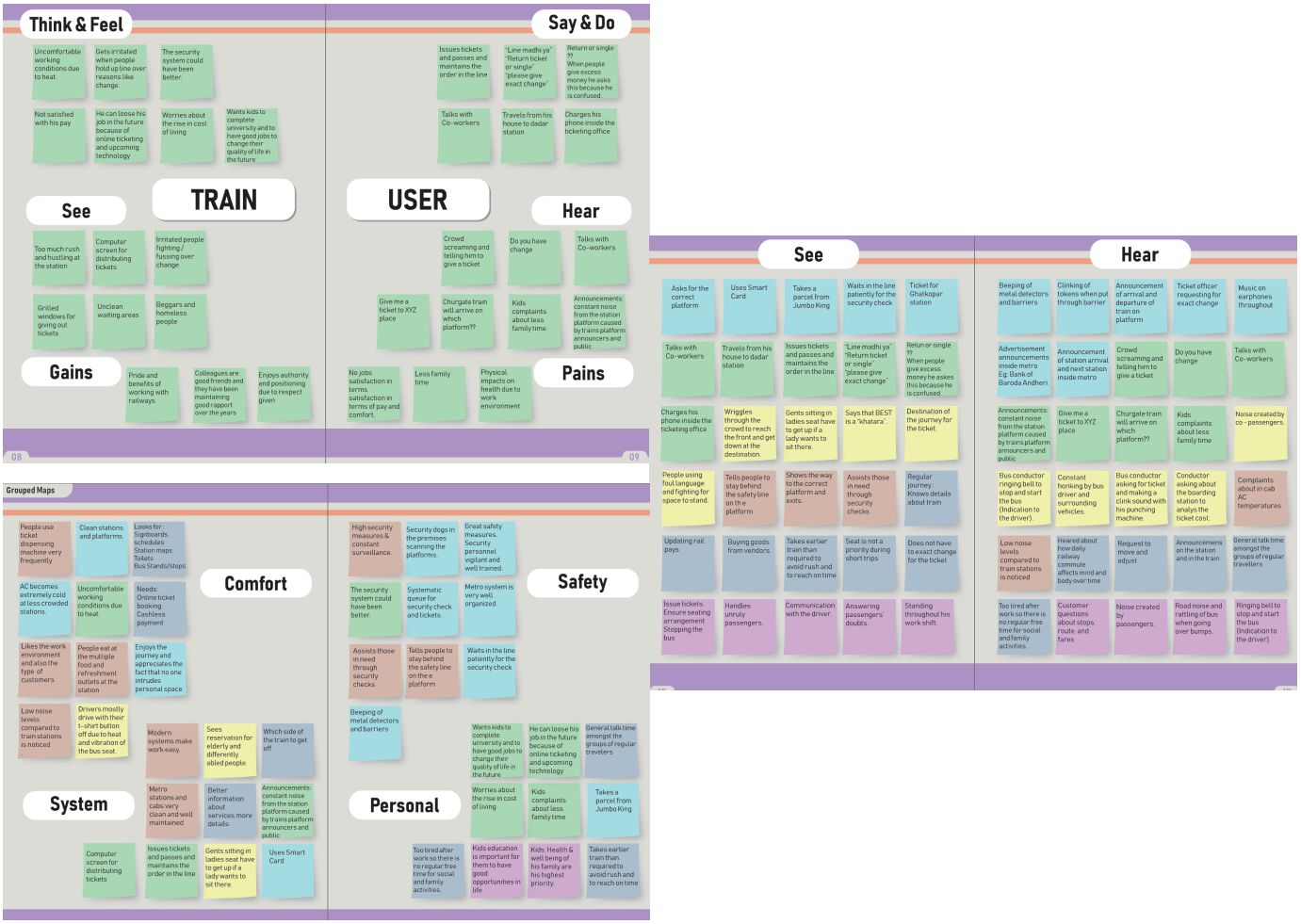


Fig. 14.: Representative overview of data analysis process followed in the orginal selected case.

POV

“ A new traveller needs to understand the railway system as he/she does wants to feel at home while commuting in the city.”

We found defining our problem statement to be one of the most important steps in this entire process. The reason behind this is before we could come down to a POV, we had to do a lot of analysis and thinking in order to understand what problems our target audience is facing. Framing our Point of View meant bringing together months of research into a statement. This was challenging because among many problems, different users were facing, finding out the main issue to be solved and for a specific group of users we uncovered from the previous steps was not an easy task.

We chose to target a new user as our audience because from our analysis we realized that once a user is well versed with how the system functions, they are habituated to the entire process and no longer see it as a problem. It is often the new user who struggles to understand what is to be done and how to do something. Thus, we concluded that if a new user is able to understand the entire process easily in their initial few days, then the entire journey becomes hassle – free.

Fig. 15.: Final Problem Frame of the the orginal selected case as a POV statement

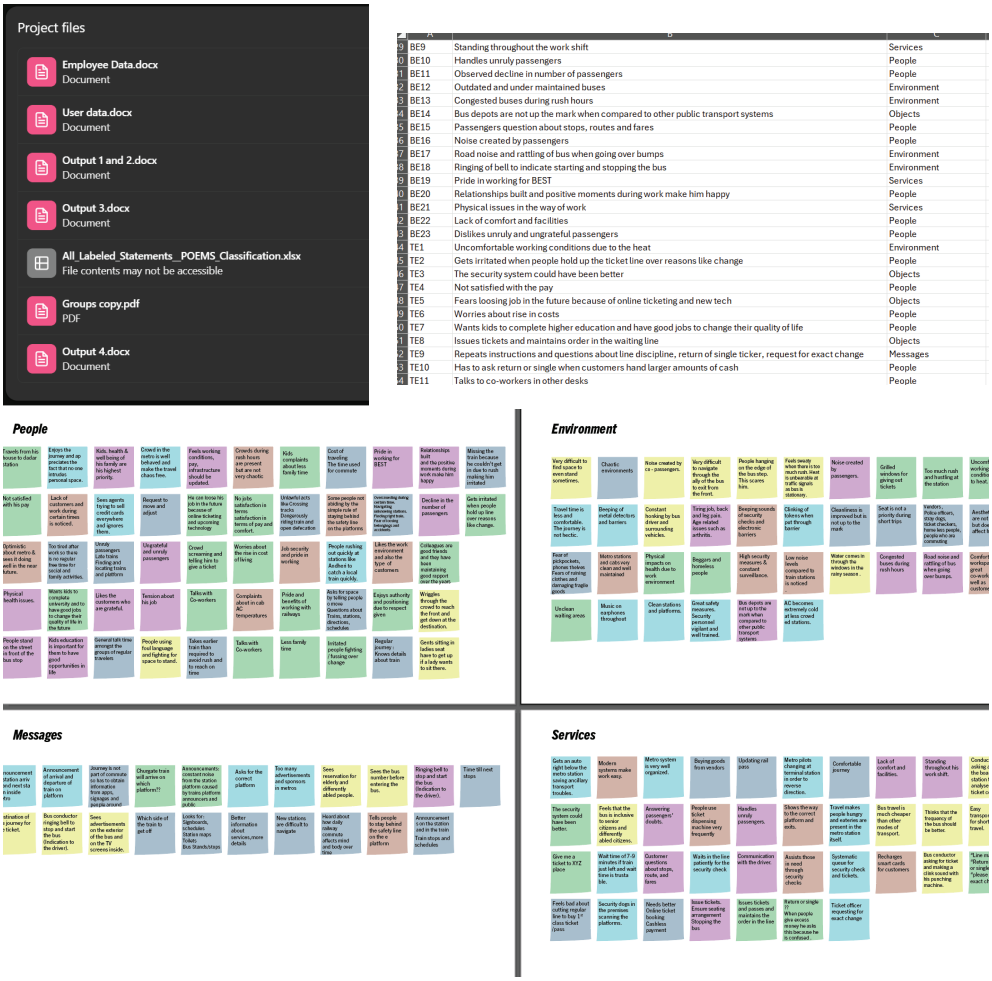


Fig. 16.: Representative overview of data analysis process followed in the autoethnography.

The Untaken Option: Reimagining Public Transport for Mumbai's Reluctant Commuters

1. Context

Mumbai's public transport system—spanning metro, buses, and local trains—moves millions daily and appears, on the surface, to function effectively.

But for a growing group of urban commuters who own or access private vehicles, public transport remains **not just unused, but unconsidered**. They opt out not due to lack of infrastructure, but due to **perceptions of unpredictability, discomfort, friction, and social compromise**.

At the same time, the system's functionality is propped up by the **silent adaptation of its regular users**, who navigate it through informal norms, learned behaviours, and tolerance for dysfunction.

This creates a tension zone:

- **Bottom-up:** Users internalize systemic failures and adapt, creating the illusion of usability.
- **Top-down:** Institutions read this adaptation as success, avoiding the pressure to evolve or onboard new users.
- **Meanwhile:** Non-users disengage altogether, their absence misread as indifference rather than a design failure.

Designing for the reluctant commuter makes this tension visible. It exposes how the system discourages entry and invites a shift—from silent adaptation to intentional inclusion.

2. Stakeholders

Primary Target Users:

- Car and two-wheeler owners who commute within reach of public transport
- Former users who "graduated" to private transport
- Commuters with access but no incentive to switch

Secondary Stakeholders:

- Existing commuters, who bear the burden of adaptation
- Public transport workers, who often compensate for system opacity
- Authorities responsible for sustainable urban mobility

3. Core Needs & Pain Points (Inferred from User Data)

Why non-users stay away:

- **Lack of onboarding:** No clear way in for first timers
- **Loss of control:** Unpredictable timings, unclear procedures
- **Social discomfort:** Overcrowding, chaos, behavioural unpredictability
- **Cognitive effort:** Too many unspoken rules to decipher
- **Status perception:** Public transport seen as a downgrade, not a deliberate choice

4. Constraints

- Fragmented governance and legacy infrastructure
- Absence of ownership over the first-time user experience
- Cultural normalization of dysfunction by both users and institutions
- Overreliance on informal norms to sustain order
- Misinterpretation of adaptation as effectiveness

5. Opportunities

- Treat first-time journeys as pivotal **conversion moments**
- Design **visible, intuitive, and emotionally dignified** entry points
- Reframe public transport as **aspirational and intelligent**, not a fallback
- Surface and support **adaptive behaviours** through intentional design
- Build system fluency *into the system itself*, not around it

6. Refined Problem Statement

How might we reimagine the first-time public transport experience in Mumbai so that private vehicle users see it not as a fallback, but as a smarter, more dignified way to move through the city — and in doing so, challenge the system's reliance on user adaptation as proof of success?

Fig. 17.: Final Problem Frame of the the authoethnography co-framing activity.

Upon completion of the autoethnography activity, iterative coding steps of the analysis of the reflections was conducted. The following list of themes, their corresponding codes and their descriptions was created. Original data documents were coded in entirety, with the attribution being contextual based on the point in the reflection or conversation it appeared in.

Theme	Codes	Description
Control of Interaction Pace	**Phase Management**	Intuitively defining, transitioning between, or concluding stages of the collaborative process.
	Active assertion of control	Deliberate steps to steer the direction and output of the interaction.
	Prompt Crafting for Pacing	Designing prompts to enforce the AI to operate within set logical constraints.
	Separation from Collaboration	Stepping away from the AI tool to create space for independent reflection.
	Moving	The act of progressing or backtracking the task within or between phases.
Expanded Sense of Role	**Responsibility Negotiation**	Accounting for the ethical implications of a framing direction.
	Value interpretation	The act of making sense of, and assigning significance to, values within the data and context.
	Doubt and Uncertainty	Feeling unsure about the right path, often triggering reflective judgment.
	Setting contextual boundary	Defining the limits of the problem space, stakeholders, and relevant considerations.
	Deciding/concluding	Making a final choice on a framing element, often after a period of reflection.
Relevant Digital Literacy	**Trust calibration**	Adjusting one's level of reliance on the AI's outputs based on perceived reliability.
	Knowledge Vigilance	Maintaining active skepticism and critical assessment of the AI's outputs.
	Active suspension of belief	Choosing to temporarily set aside skepticism to explore AI-generated outputs.
	Fixated towards exploration	Using the AI primarily to generate a wide range of possibilities and alternatives.
	Fixated towards conversion	Using the AI to refine, validate, or narrow down towards a single solution.
	Active handover of agency	Suspension of initiation or control over action to AI, turning role to supervisory in nature.
Explicit Moral & Political Lens	**Normative Interrogation**	Surfacing and questioning the value interpretations embedded in AI outputs and own assumptions.
	Systemic Positioning	Situating the problem within larger economic, political, or social structures.
	Counter-framing	Generating an alternative problem definition in opposition to an AI-suggested frame.

Table. 3.: Overview of analysed themes, ascribed codes, and their descriptions.

3.3 Driving Factors of Co-Framing Practice

Answering the empirical research question of designer autonomy is shaped during a collaborative framing activity, 4 key themes explain what driving factors shape the narrative direction of the problem framing activity, effectively also shaping perception of the value of designer's autonomy during the interaction with the AI actor in the design framing process. These factors are reported primarily in the context of this study, followed by discussion about possible directions when applying the learning to wider contexts. The reporting follows a structure of an explanation of the theme and the codes that make it up, with examples from the autoethnography. This is followed by explanations of the theme as a driver which not just shapes outcomes of the co-framing task, but also addresses the endogenous and exogenous forces shaping autonomy and broader responsible design practice. The ethnographer is referred to as "the designer" in the following reported findings.

3.3.1 Theme 1: *Actively controlling the pace of the interaction.*

This theme captures the designer's active control over the pace and structure of the interaction with the AI. The co-framing activity unfolded based on a tacit order of moments of moving built on experience and rooted in personal preference. The designer did not follow a fixed sequence of steps from start to finish. The process unfolded iteratively, with the designer deciding to speed up, slow down, or pause the collaboration based on their in-the-moment judgment. Transitions between different tasks, like reviewing data, analysing it, and developing insights, happened organically. The designer often had to dismiss the AI's suggestions to push forward or reiterate instructions to keep the session on track, actively preventing the AI from rushing the process. The active reinforcement to stay within exploratory or abductive modes was more necessary in initial phases of the activity, when the understanding of the problem space is underdeveloped. Here, a resistance to reach conclusions is seen, even if the collaborated conclusion may perceived as valuable or novel. The AI, after providing an initial overview, quickly jumped to generating high-level insights and suggestions for a systems map. The theme is comprised of several codes: the active assertion of control through deliberate prompts and commands; prompt crafting for pacing to enforce logical constraints and guide the AI's output speed; separation from collaboration by stepping away to allow for independent reflection; and moving between phases of the task (e.g., familiarising, analysing, synthesising) based on in-situ judgment.

Examples from generated data

Asserting Control & Prompt Crafting: "don't jump to insights right away. I want to understand each user group better, read in between the lines." This prompt explicitly slows down the AI's pace, rejecting its push for premature synthesis and redirecting it towards deeper, more nuanced analysis.

Separation & Moving: After a productive discussion on the metro system, the AI suggested, "Would you like to explore speculative future scenarios for the metro." The designer deflected this forward jump, stating, "Lets not jump to framing this soon. This step is just to familiarise with the groups." This demonstrates a conscious decision to remain in the analytical phase, controlling the transition between task sub-phases.

Prompt Crafting for Pacing: Later, when the AI began offering framing suggestions, the designer issued a firm corrective: "…do not give me framing suggestions at this stage, that task cannot happen now. I have not even told you what the design task is. This is just analysis of data points…" This command reinforces the boundary of the current task phase and explicitly dictates the permissible scope of the AI's contributions.

The conscious control of pace was a fundamental driver in the creation of the eventual outcome of a developed problem frame. The learning from this is the importance of creation of moments of reflection, creating cognitive distance between the divergent and convergent thinking seen in design reasoning. By resisting the AI's inherent tendency towards rapid ideation and solutioning, the designer ensured the problem space was thoroughly explored and understood. This prevented a premature convergence on a superficial or AI-suggested frame.

From an endogenous (designer's) perspective, this management of pace was a primary source of perceived autonomy. It was the mechanism through which their professional judgment about when to dive deeper, when to reflect, and when to transition, was enacted. The act of saying "not yet" to the AI was an affirmation of their own agency over the process. From an exogenous (AI's) perspective, this control manifested as a series of constraints on its operation. Its capacity for rapid information processing and generation was deliberately harnessed and throttled to serve the human-defined tempo. The AI's role was subordinated to the designer's reflective pacing, becoming a tool for depth rather than a driver for speed.

This theme highlights a critical competency for responsible AI integration, the value drivers of the pace of design activities. Designers must cultivate the reflex to resist the convenience offered by AI tools, recognising that there may be trade-offs when prioritising efficiency. Responsible practice involves recognising that speed

3.3.2 Theme 2:
An expanded sense of the role of a designer as a value shaper.

This theme describes the designer's evolved perception of their professional responsibility beyond traditional concerns of usability and aesthetics towards being a mediator of values and societal implications. It reflects an authentic belief that design is a value-laden activity, especially when framing problems that impact public systems. The codes include responsibility negotiation, actively accounting for the ethical implications of a direction; value interpretation, making sense of and assigning significance to values within data; doubt and uncertainty as a trigger for deeper reflective judgment; setting contextual boundaries for the problem space; and deciding/concluding after a period of reflection.

This theme explores how the designer saw their role as more than just a problem-solver; they acted as a mediator of values. This meant they constantly questioned what was fair, ethical, and responsible, not just what was efficient or functional. Throughout the process, the designer evaluated every AI suggestion through this lens. They didn't accept ideas at face value but probed them to uncover hidden assumptions about users, society, and power dynamics. This sense of responsibility to think about the bigger picture and long-term impact, a core tenet of strategic design approach, was also a visible driver of their decisions, ensuring the final design direction was not just effective, but also ethically sound.

Examples from generated data

Value Interpretation: When the AI summarized user data, it often used neutral or efficiency-oriented language. The designer's "reading between the lines" was an act of value interpretation, surfacing underlying issues like the "commercial colonization" of metro space by ads or the "normalization of the broken" in buses. They weren't just noting facts but interpreting their societal meaning.

Responsibility Negotiation & Doubt: The AI suggested the designer consider the "sonic infrastructure" of the system. The designer dismissed it as a secondary concern, arguing "Tackling other issues may improve the environmental soundscape, but the other way around seems pointless." This is a clear negotiation of responsibility, deciding that this aspect fell outside the core ethical scope of their intervention.

Setting Contextual Boundaries: The designer consistently framed the problem within broader contexts. They pushed back on a purely systemic view of transportation, introducing concepts like "emotional precarity," "dignity," and "civic imagination." This expanded the boundary of the problem from a technical one to a socio-technical one, which is a core responsibility of a value shaper.

This expanded sense of role was the moral and ethical compass for the entire project. It directly drove the outcome by ensuring the final problem frame ("The Untaken Option") addressed issues of dignity, equity, and access, rather than just efficiency or functionality. It prevented the adoption of an AI-suggested, technocratic frame.

From the endogenous perspective, this worldview is the very core of their professional autonomy. It is their unique value system, expertise, and ethical commitment that they bring to the table. The AI has no inherent values; it only has patterns in data. Therefore, the designer's ability to impose this value-lens on the collaboration is the ultimate expression of their autonomy. From the exogenous perspective, the AI is a values-agnostic entity. Its suggestions are neutralized or weaponized based on the designer's value framework. The AI becomes a mirror, reflecting back the data it was trained on, but it is the designer who must hold that mirror up to their own ethical standards to judge the reflection.

This theme is significant for responsible AI integration in wider design practice. It highlights that the designer's most role should not reducible simply prompter with contextual information, but as an interpreter and shaper of the value system.

3.3.3 Theme 3: *Adequate digital literacy about artificially intelligent systems.*

This theme focuses on the designer's understanding of the AI's nature, capabilities, and limitations, which enables a tactical and critical engagement. This literacy allows for both leveraging the tool effectively and maintaining a critical distance from its outputs. Key codes include trust calibration, adjusting reliance based on the AI's perceived reliability; knowledge vigilance, maintaining active scepticism; active suspension of belief to explore ideas without full commitment; being fixated towards exploration (broad ideation) or conversion (narrowing down); and the active handover of agency, consciously choosing when to let the AI lead a sub-task within the co-framing activity, switching to a supervisory role.

This theme focuses on the designer's practical understanding of how the AI works, which allowed them to use it wisely. They knew the AI is a tool that generates probable text, not a partner that possesses true understanding or intent. This knowledge allowed them to engage with its outputs tactically. They routinely checked its suggestions for bias or plausibility and crafted prompts strategically to test its knowledge limits. They knew when to trust its output for generating options and when to distrust it for making final judgments. This literacy was key to using the AI as a powerful tool without being misled by its confident but sometimes flawed responses.

Examples from generated data

Knowledge Vigilance: The AI generated a long list of thematic clusters from the PDF. The designer's response was not to accept them but to critique their utility: "good but I would stick with conclusions from output 3. this is good to keep in mind when formulating the final frame." This shows vigilance; new output from the AI was weighed against prior, human-generated understanding.

Trust Calibration & Active Suspension of Belief: The designer used the AI extensively for generative tasks like POEMS categorization, trusting it to handle a tedious sorting task. However, they later audited its work, noting "there are 77 statements in the sheet. thats not the total in the source data," demonstrating calibrated trust.

Fixation towards Exploration: The prompt "would first organising them using the POEMS method help?", is an example of using the AI for exploratory structuring of a complex dataset, leveraging its ability to quickly apply a framework without initially passing deep judgment on the content.

Digital literacy was the enabling factor that made the controlled, critical collaboration possible. It allowed the designer to use the AI as a powerful instrument without being subjugated by it. The dual role of the primary researcher and subject of the autoethnography means that a high degree of awareness and aversion to first outputs was expected. But its presence as a driving factor does highlight the potential risks if not present. In this case, it prevented over-reliance and the uncritical adoption of AI-generated content, which would have led to a generic or biased outcome.

From the endogenous perspective, this literacy is a source of confidence and empowerment. Understanding how the "black box" works allows the designer to open it, poke it, and use it intentionally. It transforms the AI from an oracle to be obeyed into a engine to be tuned and directed. This knowledge is a key pillar of their operational autonomy. From the exogenous perspective, the AI is neutral. The designer's literacy determines whether it acts as a subordinate, a partner, or a threat. A literate designer sees the AI's "confidence" as a design feature of its language model, not a sign of actual authority, and can therefore override it without doubt.

This showcases what might be ignored when understanding "prompt engineering" as a skill to develop. A deeper form of AI fluency necessary for responsible use. Designers must have a baseline understanding of how LLMs, or the relevant AI tool works. Awareness of their statistical nature, training data biases, and tendency to hallucinate demystifies the tool and grounds interaction in reality. Designers should consciously decide which tasks to hand over to the AI (e.g., brainstorming, structuring data) and which to retain full control over (e.g., final ethical judgment, conceptual framing), not purely based on their own agentic capacity, but considering what implications it may have on other practitioners with differing levels of experience and expertise.

3.3.4 Theme 4:
An explicit moral and political lens to experiences and knowledge

This theme represents the application of the designer's personal worldview and generalist knowledge to critically evaluate the AI's suggestions and the design context itself. It is the active use of a political and ethical lens to interrogate not just the solution, but the very definition of the problem. It comprises normative interrogation, surfacing and questioning embedded values; systemic positioning, situating the problem within larger political and social structures; and counter-framing, generating alternative problem definitions in opposition to an AI-suggested frame.

This theme highlights how the designer's personal beliefs and broad knowledge shaped their critique of the AI's ideas, and the overarching direction of the realisation of output design frame. A specific lens focused on equity and access, rooted in socialist perspectives was used to evaluate every suggestion and guide the framing. For example, when the AI framed a transportation issue around efficiency, the designer reframed it around fairness and who was being left out. Drawing on a personal perspective on knowledge from adjacent fields like urban planning and transit networks to spot the wider implications the AI missed. This lens was the main filter for all information; it determined which ideas were pursued and which were rejected, ensuring the final problem frame tackled issues of justice, not just technical problems.

Examples from generated data

Systemic Positioning & Normative Interrogation: The designer reframed the AI's technocratic language. When the AI might see a "mobility challenge," the designer, drawing on knowledge of urbanism and political theory, reinterpreted it through a "systemic lens as a question of civic access, power distribution, and infrastructural semiotics." This is a direct application of a political lens.

Counter-Framing: The entire final problem framing exercise was majorly an act of counter-framing. The AI, based on the data, might naturally frame the problem around improving the experience for existing users. The designer consciously countered this by framing it around attracting non-users, a fundamentally different and more politically charged goal that challenges the status quo.

Normative Interrogation: The designer's insight about the "tension zone" between top-down institutional apathy and bottom-up user adaptation is a profound political-economic analysis. It identifies a systemic inertia that is a direct result of power dynamics and misaligned incentives, a conclusion far beyond the AI's analytical capabilities.

This lens was the differentiator between a competent analysis and a transformative one. It provided the critical edge to move from what is to what could and should be. It directly shaped the outcome by insisting that the problem be defined in terms of justice and desirability, not just functionality. This lens can also act as a blind spot, as a practitioner may not be aware of what they consider as normal or obvious, let alone reflect on why that is. Socio-cultural aspects of lifestyle, especially in this case of public transport, determined what the designer considered as favourable. A designer whose worldview is developed in a car-centric society, will have differing opinions than one exposed to strong transit networks. Therefore, the crux of this theme is the explicit component, as these intrinsic biases in self as well as collaborative human or non-human actors cannot be detected and solved for without it.

From the endogenous perspective, this lens is the designer's unique intellectual contribution, their "designer's world." It is the culmination of their education, reading, and lived experience. This is irreplaceable and non-automatable, forming the bedrock of their conceptual autonomy. It is what allows them to see what the AI cannot. From the exogenous perspective, the AI in this case lacks any real coherent moral or political lens. It can simulate different perspectives based on its training data, but it does not hold a perspective. Therefore, it is entirely dependent on the designer to provide this critical, orienting framework. The AI's output is raw material to be filtered through the designer's worldview.

Integrating AI necessitates a stronger, not weaker, commitment to developing and applying a critical political and ethical lens. Design education must encourage broad knowledge across sociology, political theory, and history to equip designers with the frameworks needed to critically interrogate AI outputs and complex problem spaces. Designers must be aware of their own positionality and biases and explicitly state them when working with AI, understanding that the AI will amplify these biases if they are not critically examined. Use the AI to generate multiple, even opposing, frames for a problem. Then, use explicit moral and political lenses to critique and choose between them, turning the design process into a structured debate between human values and AI-generated possibilities.

3.3.5 Conclusions of the Empirical Investigation

To conclude the empirical component of the project, the descriptive nature of the overarching research should be taken into account. With the focus of this section being solely on explicating the mechanism through which the value of autonomy is shaped, we get can conclude that understanding if the AI interaction supports or undermines components of autonomy cannot be equated to the activity resulting in ethically desirable design outcomes.

The neutral conceptual description of the autonomy, its components, and situating it in practice is supported by the synthesised themes of the empirical investigation. The four themes, when applied to specific design cases or designers, can allow for an initial framework for backward facing responsibility to understand driving forces behind specific design outcomes, or can inform forward facing responsibility to pre-empt potential challenges that could lead to undesirable design outcomes.

The four themes, while leaning towards either agency or authenticity, can differ from case-to-case application. This reinforces the interlinked nature of these components in the conceptualisation of autonomy, and a reminder that they only serve as analytical or descriptive tools when understanding autonomy. The dual role played in the autoethnography meant that there was a high degree of conceptual awareness and familiarity of the subject matter beyond the information within the design task, for example, a good sense of ethical concerns of algorithms. This means that while the theme leans towards authentic beliefs as the driving component in this study, a similar approach of showing less suggestibility to AI prompts can be a potentially observed in a designer with a sense of agency, informing them to be the driving actor in the process.

To answer the sub-questions formed at the start of this chapter, the empirical findings demonstrate that designers (a) negotiate and exert their agency through the deliberate orchestration of the collaborative process, such as controlling the interaction pace, strategically crafting prompts, and rejecting or reframing AI suggestions to maintain command over the workflow and final outcome. Simultaneously, they (b) perceive and articulate their authentic influence as stemming from their core professional identity and values, which act as a critical filter for all collaboration; this is evidenced by their conscious application of a moral-political lens to interrogate AI outputs, their negotiation of an expanded ethical responsibility, and their reliance on personal expertise and generalist knowledge to define the problem's boundaries and ensure the final frame is a genuine reflection of their own critical stance and worldview, rather than an AI-generated proposition.

These empirical themes point toward broader lessons for autonomy in AI augmented design practice. A designer’s autonomy is not an innate trait but a skilful balance of process, perspective and purpose, with its virtuous actualisation being closely linked to design expertise. By training designers to orchestrate framing interactions intuitively, to adopt a reflective stance on their professional responsibilities, to develop nuanced digital literacy, and to foreground their own value judgments, teams can guard against default AI logics and foster more responsible outcomes. In other words, these dimensions serve as transferable levers, usable both retrospectively to trace how decisions shaped past projects and proactively to anticipate where AI might unduly constrain future work.

CHAPTER 4

TECHNICAL INVESTIGATION

The role of the technical investigation in the adapted Value Sensitive Design approach of this project is concerned with specific technological properties that influence the value under observation. The empirical investigation provides this section with driving factors which shape how designers exercise their autonomy in co-framing activity. These highlight the conditions that inform decision-making and influence the trajectory of outcomes, which the technical investigation can explore from the perspective of the technology. The goal is to understand how autonomy is conditioned by the functioning of AI systems, their broader infrastructures of decision-making support, data processing, and value encoding.

4.1 Ethical Concerns of Algorithms

This project has placed artificial intelligence as the technology creating the value tension, but to understand the implications of its characteristics, an understanding of algorithmic systems is necessary. While AI is the specific technology, literature on broader algorithmic ethics is explored to formulate its technologically relevant risks to the values of human actors. Following Mittelstadt et al. (2016), "algorithms" refers to decision-making systems that process data inputs and generate outputs with minimal or no human intervention. These systems are increasingly autonomous in the sense that they do not merely support human decisions but actively shape them, particularly in contexts like generative creation, recommendation systems, and predictive analytics. The ethical concerns associated with algorithmic decision-making will help select key characteristics of AI systems prominent in potential design framing activities.

Mittelstadt et al. (2016) identify six core ethical concerns, five of which are grouped into two distinct categories, and one a broader concern. The two categories are:

Epistemic concerns: how algorithmic systems generate, justify, and communicate knowledge, impacting what can be known and trusted.

Normative concerns: how algorithmic decisions affect fairness, accountability, and moral outcomes, shaping what should be done and for whom.

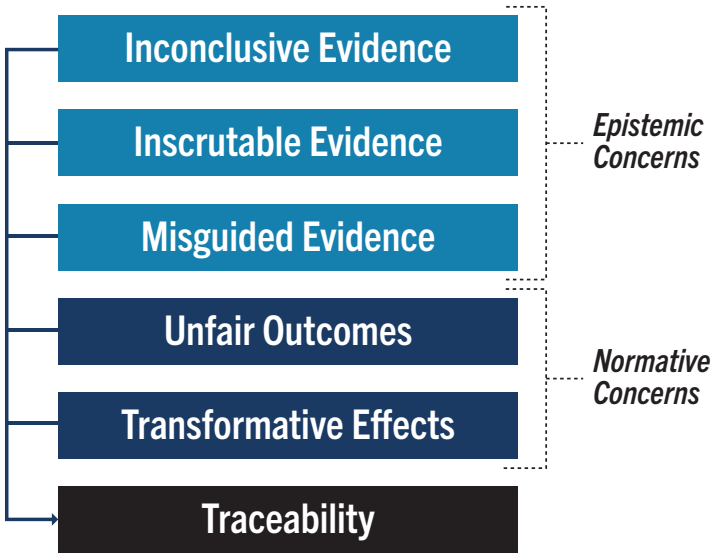


Fig. 19.: Categorisation of the ethical concerns of algorithms. *(Mittelstadt et al., 2016)*

4.1.1 Epistemic Concerns of Artificial Intelligence

The first concern is ***Inconclusive Evidence.*** Algorithms produce probabilistic knowledge based on statistical inference or machine learning models. While correlations can offer actionable insights, they do not always imply causation, making the justification for action uncertain. Acting on such conclusions can weaken the epistemic standards of responsibility.

In design, inconclusive evidence often takes the form of AI-generated insights, such as predicted user needs or emergent trends, which may appear legitimate but lack causal grounding. Designers risk embedding speculative assumptions into real-world interventions, undermining their ability to justify the rationale behind decisions. This erodes authenticity by weakening the link between belief and responsibility.

The second concern is ***Inscrutable Evidence.*** Complex algorithmic systems, especially those using deep learning, are often opaque. The logic behind their outputs is not interpretable by humans, violating the principle that evidence should be intelligible and subject to critique.

Designers working with opaque recommendation systems or black-box generative tools may not fully understand why a particular option is suggested. This limits their ability to critically assess or revise AI contributions and reduces their autonomy to act on well-grounded reasoning, again impacting authenticity.

The third concern is ***Misguided Evidence.*** Algorithms rely on the data they are given. Poor-quality, biased, or incomplete data will produce flawed results. This limitation is intrinsic to data-processing systems and challenges the neutrality of evidence.

When design tools are trained on biased data (e.g., overrepresented user types or market priorities), the algorithmic suggestions may reflect systemic biases. Designers relying on these outputs may unwittingly reinforce exclusionary or harmful patterns, distorting the design process. This may introduce epistemic blind spots, making authenticity more difficult to maintain.

4.1.2 Normative Concerns of Artificial Intelligence

The fourth concern is ***Unfair Outcomes.*** Even when algorithms generate reliable outputs, they may produce discriminatory or unjust results. The ethical wrongness lies in the action's consequences, particularly on vulnerable or marginalised groups.

Designers who use AI to optimise interfaces, workflows, or services may unknowingly exclude minoritised users if the system's logic privileges dominant use cases. This affects agency, as designers may not even realise that their outputs are unjust, limiting their ability to intervene or reshape decisions that violate their ethical stance.

The fifth concern is ***Transformative Effects.*** Algorithms do not merely make decisions; they alter how decisions are understood and who gets to make them. Over time, they shift social practices, values, and worldviews without necessarily causing obvious harm.

In design, this manifests as a redefinition of what constitutes good design. AI may subtly promote efficiency, optimisation, or stylistic trends as default values, causing designers to shift their standards without active reflection. These systemic shifts constrain agency by reducing the space for critical deviation and shifting the designer's role from author to validator.

4.1.3 General Concerns of Artificial Intelligence

The sixth concern is ***Traceability.*** Algorithmic decisions often lack clear chains of responsibility. It can be difficult to identify who, or what, is accountable when harm occurs. This poses a challenge for ethical assessment, as responsibility must be traceable to evaluate causes and assign accountability.

When designers incorporate AI outputs into decision-making without knowing how or why those outputs were generated, they may still be held responsible for downstream consequences. This undermines both authenticity (belief without justification) and agency (action without control) and may place the designer in a position of reduced epistemic and moral clarity. within human-AI collaboration.

Ethical concerns of algorithms, and therefore artificially intelligent systems can significantly shape a designer's perceived autonomy by affecting either dimensions in a particular interaction or over repeated interactions over time.

The influence authenticity can be through presentation of negations or judgements in a way which discourages scrutiny or contestation. Authenticity, as discussed earlier, involves the ability to act based on one's own critically examined beliefs and judgments. When the basis of a recommendation or insight is unknowable, poorly grounded, or probabilistically uncertain, the designer's ability to reflect, justify, and maintain belief ownership is undermined. These epistemic limitations may constrain the conditions under which authenticity can be meaningfully exercised.

The impact on agency can occur by reshaping the capacity to act with intentionality and influence outcomes. When algorithmic systems distort values through oversimplified optimisation, fail to adapt to local contexts, or reinforce dominant norms at the expense of diversity, they reduce the designer's ability to act ethically and contextually. The system shapes outcomes in ways that evade intervention or oversight.

4.2 Technical Characteristics of AI

To examine how autonomous technologies shape designer autonomy, this section investigates three interrelated technical aspects of AI systems: algorithmic bias, trustworthiness , and explainability. Each aspect reveals how the technical design of AI tools can either support or constrain the designer's capacity for epistemic and moral responsibility, particularly during the framing of design problems. This section addresses the technical aspects of AI tools that impact the exercise of designer autonomy.

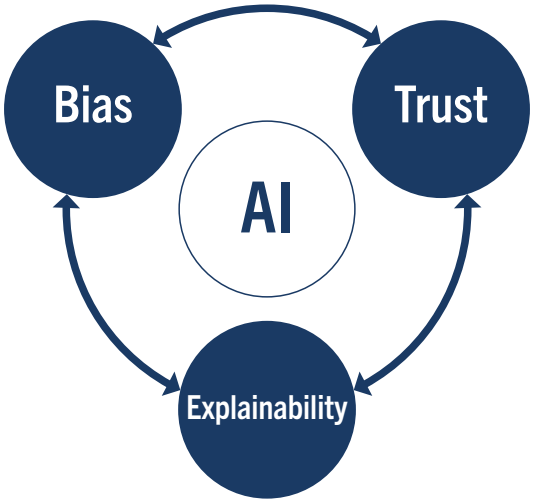


Fig. 20.: Interdependent technical characteristics of Artificially Intelligent systems Bias, Trust and Explainability, affecting the autonomy of the collaborating human actor.

4.2.1 Bias

Danks & London (2017) offer a nuanced conceptualisation of algorithmic bias as a deviation from normative standards, statistical, moral, or legal, rather than simply a technical flaw. Their detailed taxonomy identifies five distinct bias sources in two different groups (figure 21). The first three routes are centred around technical or computational aspects, these are: Training Data Bias, Algorithmic Focus Bias, and Algorithmic Processing Bias. The latter two arise from inappropriate uses or deployment of autonomous systems, which are: Transfer Context Bias and Interpretation Bias. Critically, they argue that not all biases are problematic or to be eliminated. Some, like the deliberate use of a statistically biased estimator for robustness to enforce moral norms, are valuable components towards having a reliable and ethically desirable system, acting as a course correction for previously existing biases. Each type arises from human decisions embedded in the AI development pipeline, revealing that bias is not merely a downstream effect but also an upstream design condition.

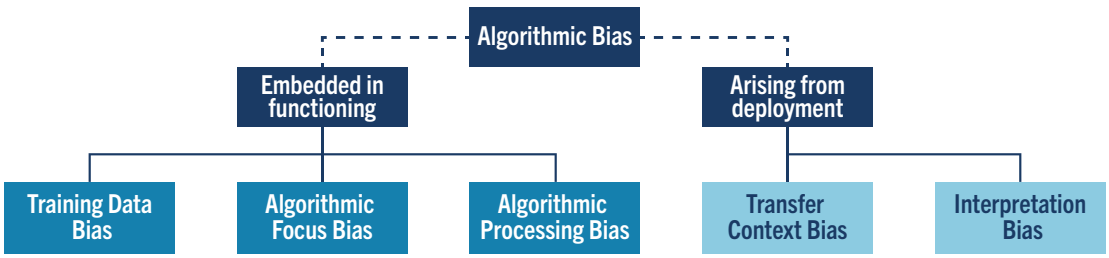


Fig. 21.: Five types of algorithmic bias as categorised by Danks and London *(2017)*.

For autonomous systems, these biases become especially consequential when AI-generated outputs shape how problems are framed, or which user groups are prioritised. The taxonomy provides a framework for diagnosing these issues. For instance, training data bias could cause a system to reinforce dominant perspectives present in its data, while algorithmic focus bias might lead it to improperly use protected attributes. Transfer context bias occurs if a system is deployed in a setting different from its intended use, and interpretation bias can cause designers to misread the system's outputs. In strategic design, where designers engage in problem structuring and value negotiation, these algorithmic biases can implicitly reinforce dominant perspectives while rendering alternative framings invisible. scope of viable actions without the designer's awareness.

This may interfere with the conditions for authenticity, as designers could be nudged toward system-aligned framings that do not reflect their examined commitments. Simultaneously, agency is constrained when these systems, through their various biases, subtly shape the scope of viable actions without the designer's awareness.

Algorithmic bias reveals how hidden normative assumptions in data and modelling choices shape the epistemic constraints within which designers operate. Understanding these constraints invites technical investigations into the socio-technical provenance of data, scrutinise pre-trained models for value-laden features, and examine how patterns of exclusion may be encoded in AI-assisted framing tools. This requires methods that go beyond statistical validation and interrogate the ethical appropriateness of data-model alignments.

4.2.2 Trustworthiness

Durán (2024) argue that trust in AI systems must be distinguished from mere reliance. Trust, as they define it, requires more than consistent performance, it depends on an "extra factor" that includes accountability, transparency, and a capacity to support normative expectations. Crucially, they propose computational reliabilism as an epistemological framework for justifying trust in AI systems that are epistemically opaque. Unlike transparency, which seeks direct interpretability, computational reliabilism evaluates whether a system reliably produces true or warranted beliefs across domains, grounded in the system's technical robustness and its embeddedness in scientific or professional practices.

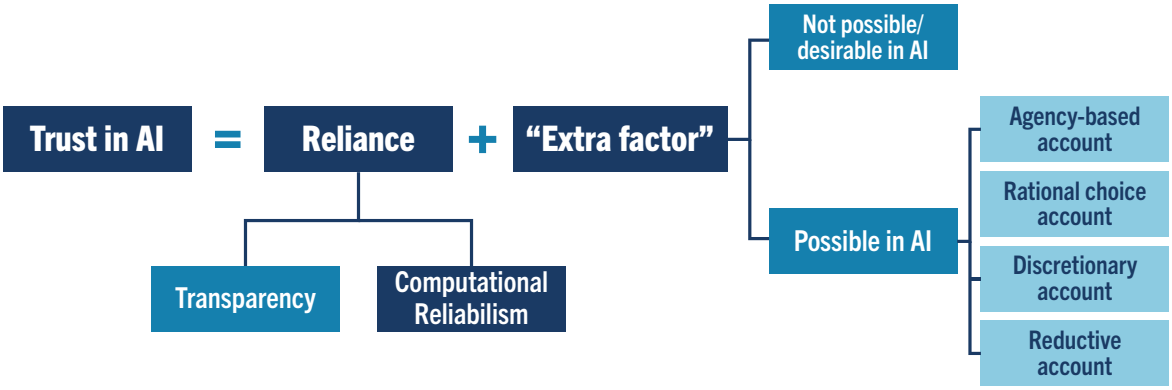


Fig. 21.: Model of Trust in AI by Durán *(2024)*.

In design contexts, designers often engage with systems that are only partially interpretable. Trust without justification, based on opaque reliability or reputational heuristics, may position the designer primarily as an executor of algorithmic outputs, rather than an interpretive agent. This may constrain agency by disconnecting action from the designer's informed judgment. Furthermore, when designers are expected to defend or enact decisions they do not understand, they may become distanced from the reasoning behind decisions for which they remain accountable.

The concept of computational reliabilism redirects technical investigation toward the epistemic conditions under which designers form beliefs about AI system outputs. Rather than striving solely for full transparency, AI tools can support designers by providing indicators of reliability, such as model validation across comparable contexts, traceable performance history, and verifiable alignment with domain standards, that support justified epistemic trust. These indicators must be legible and meaningful within design workflows, where normative and contextual judgments are central.

4.2.3 Explainability

Barredo Arrieta et al. (2020) frame explainability as a critical condition for enabling responsible human–AI interaction. They distinguish between ante-hoc explainability, in which models are interpretable by design, and post-hoc methods that attempt to reconstruct logic from opaque systems. The challenge lies in the growing complexity of high-performing models, particularly in deep learning, where increased predictive power often comes at the cost of human interpretability. This tension is especially acute in strategic design, where understanding the reasoning behind a suggestion is as vital as the algorithms's suggestion itself.

Explainability is not a one-size-fits-all but is often discussed in relation to the cognitive and contextual demands of its audience. In design processes, opaque AI systems may encroach upon or replace elements of the designer's reflective and deliberative role, shifting elements of problem framing toward system-driven logic. This may diminish authenticity, if designers lose the ability to endorse or contest recommendations on ethical or experiential grounds. It may also impact agency, by limiting the conditions under which designers can modify, reject, or reinterpret AI suggestions in ways that align with broader project goals or stakeholder needs. Explainability may be more productively approached as a condition for context-sensitive epistemic alignment. Technical investigations should assess how explanations are generated, what assumptions they encode, and whether they enable designers to evaluate, contest, and reframe system outputs. This includes designing explanation interfaces that foreground the moral salience of decisions and preserve the interpretive authority of the human designer.

4.3 Risks to Designer Autonomy in the Age of AI

Across these three technical concerns, bias, trust, and explainability, a common theme reinforces the conceptual investigation findings that AI systems do not merely support decision-making but participate in shaping the epistemic and moral conditions of design practice. Each technical property affects whether the designer retains the ability to frame problems responsibly, critically assess alternatives, and justify outcomes. Preserving designer autonomy in AI-mediated contexts requires that tools be technically configured to expose, rather than obscure, the assumptions, limitations, and ethical stakes of their operation. This includes identifying embedded value hierarchies (bias), enabling justified epistemic trust (computational reliabilism), and designing for actionable, context-aware explanations (explainability). Only by investigating these properties as directly impacting design decision-making can designer autonomy be meaningfully safeguarded.

With this technical understanding, the risks to the two dimensions of autonomy can be articulated in relation to the actions of possible AI behaviour. Prunkl (2022) articulates of the risks AI poses to human autonomy and explains them in accordance with the potential capabilities of artificial intelligence visible in common practice. These risks, if contextualised to this project's focus, offer a conceptual foundation for identifying how autonomous systems may distort a designer's moral and epistemic capacity in practice. This section first outlines these general risks as defined by Prunkl before reframing them as per the established understanding of designer autonomy.

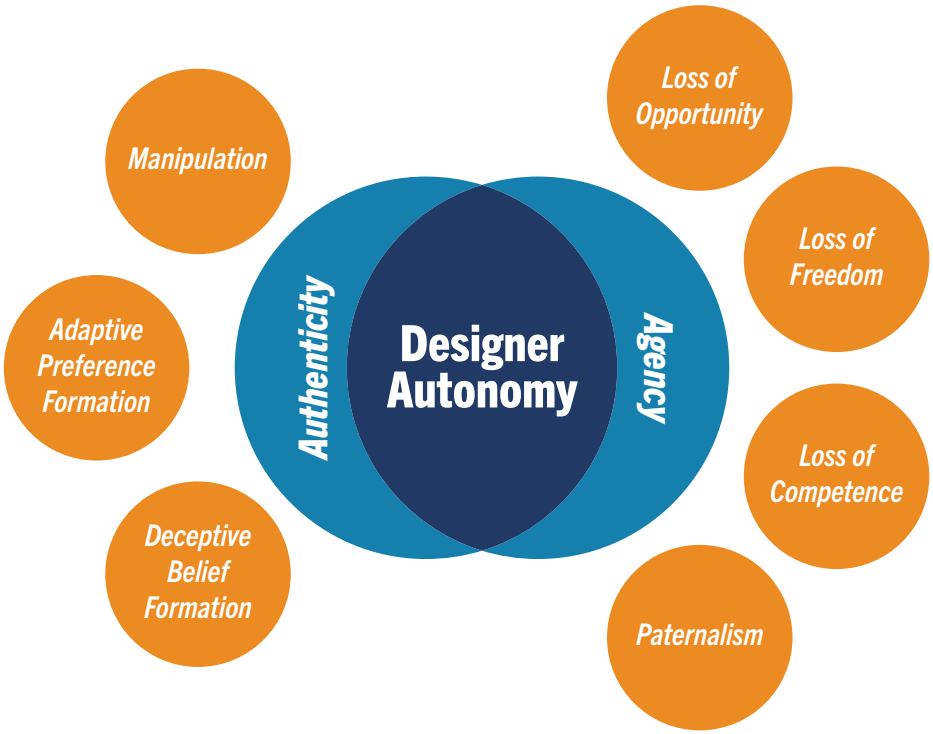


Fig. 21.: The risks to the dimensions of Designer Autonomy in the Age of AI. *(Adapted from Prunkl, 2022)*

Risks to Authenticity

Manipulation: AI systems (e.g., recommender systems) exploit cognitive vulnerabilities, nudging users' decisions without their awareness or consent.

Adaptive Preference Formation: Individuals adjust their desires to match the limited set of options presented by AI, often unconsciously, reinforcing the status quo or system-generated constraints.

Deceptive or Adaptive Belief Formation: AI can shape users' worldviews through algorithmic amplification of information (e.g., conspiracy content), reducing the user's ability to evaluate information independently.

Risks to Agency

Loss of Opportunity: Biased algorithms can systematically limit access to opportunities (e.g., in hiring, healthcare), reducing individuals' ability to realise their goals.

Loss of Freedom: Surveillance, predictive control, and automation can restrict behavioural options, either directly or through anticipatory constraint.

Loss of Competence: Routine reliance on AI can atrophy human decision-making skills, leading to diminished capacity for reflection, judgment, or ethical action.
Paternalism: AI systems may override individual preferences "for their own good," often without consent or recourse, undermining self-determination even when intentions are benevolent.

Building on the previous discussion of how algorithmic systems generate and justify knowledge and how technical properties like bias, trust, and explainability shape design conditions, this section focuses on the specific risks posed by AI to designer autonomy. These risks are not reducible to isolated failures of a system but rather emerge from the structural logic of AI tools as they mediate decision-making processes. Grounded in Prunkl's bifurcation of autonomy into authenticity and agency, the section delineates how AI systems may subtly influence designers' beliefs, preferences, and capacities for moral judgment. By mapping these risks against the ethical concerns identified by Mittelstadt et al. (2016), this section offers a conceptual framework to understand how epistemic and normative system behaviours jointly contribute to the erosion of autonomy. The aim is to show that these risks are interdependent and cumulative, operating beneath the surface of design activity while conditioning what is visible, thinkable, and actionable in practice.

Risks to Designer Authenticity

Manipulation: When AI tools subtly steer the designer's focus through prompts, defaults, or pre-structured options, the resulting framing risks reflecting system biases rather than the designer's own beliefs. The designer may unknowingly internalise these framings, leading to decisions that reflect system logics more than critically examined personal judgments. This undermines authenticity by eroding the connection between decision and self-authored judgment.

Adaptive Preference Formation: As designers work within AI systems that consistently favour certain styles, problems, or outcomes, they may begin to adapt their preferences to align with what the system rewards or enables. Over time, this narrows the designer's interpretive repertoire and gradually aligns their epistemic stance with dominant system logics. Authenticity is compromised when the designer's sense of what "feels right" is shaped less by reflection and more by machine-aligned habit.

Deceptive or Adaptive Belief Formation: When the information environment is algorithmically curated (e.g., AI-generated insights, trend predictions), designers may build project justifications or moral reasoning atop distorted or biased informational foundations. The risk is not outright falsehood, but a subtle detachment from the designer's own values and epistemic awareness. Authenticity demands ownership over one's beliefs, which may be challenged when belief formation is shaped by the algorithmic structure of design tools.

Risks to Designer Agency

Loss of Opportunity: AI systems that foreground certain framings, stakeholders, or problem types often do so by excluding others, narrowing what the designer can explore or propose. When tool defaults or client-facing outputs privilege efficiency, scale, or market fit, the designer's ability to propose alternative, value-driven directions diminish. This limits agency by making counter-framing structurally difficult to articulate or justify.

Loss of Freedom: Automation within the design process (e.g., automated wireframing, trend-based concept generation) can reduce the flexibility to pause, question, or redirect. Designers may become operators of workflows in which deviation feels inefficient or unjustifiable. The freedom to shape not only what is made, but what is seen as possible, becomes compressed, narrowing the space for political and moral reflexivity in framing decisions.

Loss of Competence: As designers grow accustomed to AI-assisted decision-making, their capacity to independently frame, argue, and justify design decisions may diminish. Reflection-in-action is offloaded, and the space for skilled intuition or abductive reasoning is filled by automated suggestion. This weakens agency by dulling the critical skills necessary for value-aware action in ambiguous contexts.

Paternalism: Some AI design tools present best practices or model outputs as authoritative, leaving little room for contestation. Even when well-intentioned, these systems may override the designer's reflective judgment "for their own good", assuming that deviation from normed standards is error rather than ethical divergence. Agency is reduced when designers feel discouraged or unable to go against automated advice.

A conclusive list of meanings of risks, specific to the context of autonomy of a designer is as follows:

Risk	Meaning
Manipulation	The subtle steering of a designer's framing decisions through pre-structured tool logic, such that system biases replace reflective judgment.
Adaptive Preference Formation	The gradual internalisation of system-favoured styles or outputs, narrowing the designer's interpretive repertoire over time.
Deceptive/Adaptive Belief Formation	The shaping of beliefs and justifications by algorithmically curated information, resulting in detachment from personal epistemic grounding.
Loss of Opportunity	The narrowing of design possibilities due to system defaults privileging specific framings, values, or stakeholders.
Loss of Freedom	The reduction in reflective flexibility as automation makes deviation feel inefficient or unjustifiable within constrained workflows.
Loss of Competence	The atrophy of framing and decision-making skills due to reliance on automated suggestions in ambiguous contexts.
Paternalism	The undermining of reflective judgment through the system's presentation of normative outputs as authoritative or non-contestable.

Table. 4.: Table of meanings of the identified risks of AI to Designer Autonomy.

To better understand how systemic properties of AI tools affect designer autonomy, the specific risks can be mapped to the previously explored six types of algorithmic concerns identified by Mittelstadt et al. (2016). Each risk is presented in relation to the autonomy components of authenticity and agency. This mapping reveals how both epistemic and normative concerns may undercut either component. The relationship is such that each risk may be realised by one or more underlying ethical concerns.

This matrix (table 4) reflects the distributed nature of influence within AI-mediated design environments. Each risk to designer autonomy emerges from multiple of system-level conditions, not a single ethical failure. Notably, the mapping reveals that the conventional division between epistemic concerns (e.g., inscrutability, inconclusiveness) and normative concerns (e.g., unfair outcomes, transformative effects) does not hold cleanly when considering how designers interact with AI. Risks to authenticity, such as belief formation or preference drift, are not always enabled by epistemic opacity alone but are often shaped by normative system behaviour over time (e.g., optimisation loops, behavioural nudges). Conversely, constraints on agency, such as paternalism or loss of framing freedom, may originate not only in normative overreach but also in epistemic limitations, for instance, when justification chains become non-traceable.

	Misguided Evidence	Inscrutable Evidence	Inconclusive Evidence	Unfair Outcomes	Transformative Effects	Traceability
Manipulation	●	●			●	
Adaptive Preference Formation			●		●	
Deceptive/Adaptive Belief Formation	●	●	●			●
Loss of Opportunity				●	●	
Loss of Freedom					●	
Loss of Competence					●	
Paternalism	●				●	●

Table. 5.: Matrix of identified risks of AI to Designer Autonomy and ethical concerns of algorithms.

This interdependency suggests that the two dimensions of autonomy should not be treated as isolated targets of system effects but are co-implicated in the same infrastructural logics. Design decisions shaped within AI systems do not merely reflect biased or opaque information; they are conditioned by how systems allocate attention, constrain interpretive space, and normalise certain framings over others. Autonomy, then, is not eroded through overt coercion or technical failure, but through the cumulative effect of value-laden structures embedded in tool logic, user interface, and data orientation. These structures are often subtle, operating beneath the level of explicit intention, yet shaping the space of design decisions in meaningful ways. Because of this, the erosion of autonomy is rarely self-evident; it tends to manifest through the gradual misalignment between a designer's intent and the conditions that guide action.

Awareness of autonomy as a situated and structurally conditioned capacity is therefore critical, not only for understanding how it is shaped, but for maintaining the reflexive ability to detect when it is being displaced. This reflexivity does not guarantee its preservation, but it renders autonomy a perceptible property. This awareness creates the potential to detect its erosion and can consequently lead to actions aimed at its maintenance or reclamation.

CHAPTER 5

CONCLUSION AND REFLECTION

Throughout this research, the nature of designer autonomy has been gradually developed as a dynamically enacted value, shaped within the complex interplay of human intention and technological mediation. Adopting the Value Sensitive Design approach, the project has moved through conceptual, empirical, and technical investigations to explore how autonomy is understood, exercised, and conditioned in AI-mediated design practice, specifically within the value-laden act of framing. By situating the inquiry in the context of strategic design and introducing the notion of co-framing, the study has sought to make visible the often tacit negotiations between designer and system, where beliefs, agency, and responsibility are constantly at stake.

The empirical investigation, an autoethnography of the novel conceptualisation of collaborative framing, reinforces that autonomy is neither simply given nor taken, but practiced through a combination of driving factors of the designer's actions. Factors which can be either supported or compromised by AI tool behaviour. These factors do not operate in isolation but interact to shape the designer's capacity to steer the process authentically and with agency. Simultaneously, the technical investigation highlighted how identified systemic properties of AI of bias, trustworthiness, and explainability structure the possibilities for meaningful human intervention, presenting both subtle and overt risks to autonomy across epistemic and normative dimensions of ethical concerns.

What emerges from this exploration is a richer, more practical understanding of autonomy as a relational and context dependent value. It becomes clear that preserving or enhancing designer autonomy in the age of AI is not a matter of resisting technological integration, but of cultivating the conceptual, practical, and technical conditions that enable critical, reflective engagement within increasingly mediated design environments. This section concludes the project by reflecting on the learnings about designer autonomy, answers to the three research questions, followed by broader implications for design practice, ethics, and future research.

5.1 Framing Practice as a Regime of Designer Autonomy

Joel Anderson's "regimes of autonomy" framework (2014) reframes autonomy not as a fixed essence but as a normative social order in which both the criteria for autonomy ("what gets you the deontic status") and the licensed privileges ("what autonomy gets you") are contingent on institutional practices and justificatory rationales. A regime, in Anderson's terms, is an interconnected package of (1) qualifications for autonomy, (2) institutionalized procedures for attributing, and contesting, that status, and (3) the underlying moral and political justifications for both criteria and implementation. By foregrounding regimes, Anderson shifts the focus from abstract definitions to practical stakes, showing that any claim "this actor is autonomous" carries embedded assumptions about thresholds of competence, permissible interventions, and broader social values. While the findings of the project are rooted in descriptive ethics, discussing them with the lens of a regime of autonomy allows for a glimpse in the potential application towards prescriptive research.

Design framing, the interpretive work by which designers define "what the problem is" and "what matters" has all three elements to qualify as a regime. First, what gets you autonomy in framing is the designer's demonstrated capacity for reflective judgment, the ability to interrogate assumptions, resist surface level cues, and reinterpret AI generated suggestions. Second, framing practice is institutionalized through professional norms, design education, and tooling workflows that prescribe who may frame, how frames evolve, and how authority is allocated between human and machine. Third, the justification for this regime is rooted in the political and ethical imperative to shape futures responsibly, recognising that framing choices embed value-laden exclusions and priorities . Drawing on the three-fold Value Sensitive Design investigations, we can now articulate the dual dimensions of autonomy in AI-mediated framing:

Fig. 22.: Two key aspects which define one's autonomy within the bounds of a specific context called "regime". *(Anderson, 2014)*

What Gets You Autonomy

Conceptual Foundations: Autonomy is constructed as a value blending authenticity (acting from critically examined beliefs) and agency (shaping problem frames intentionally) within dynamic, value laden interactions .

Empirical Enactments: The autoethnography revealed four enabling conditions: controlling the pace of activity, a sense of role and responsibility, digital literacy about relevant AI tool, and the political lens, each of which designers must not just cultivate, but continuously reflect on to claim autonomy in the co framing process.

Technical Preconditions: AI characteristics that support autonomy include transparent reliability indicators (computational reliabilism), context adaptive explainability, and interface affordances that invite critical interrogation rather than passive acceptance.

What Autonomy Gets You

Epistemic Authority: Designers are empowered to endorse, contest, or deviate from AI suggestions, maintaining interpretive control over problem definitions.

Moral Responsibility: Autonomy grants the right, as well as the responsibility to justify framing choices in light of broader societal and ethical commitments, ensuring accountability for design trajectories.

Creative Latitude: With autonomy, designers can expand the solution space, injecting systemic or socio political concerns into AI mediated workflows rather than deferring entirely to default model logics.

Institutional Recognition: Claims of autonomy license designers to negotiate project scope, resource allocation, and stakeholder priorities, shaping both process and outcome at the organizational level.

Together, these elements show that designer autonomy in collaborative AI practice is neither a given nor a simple absence of constraint; it is a situated capability, a regime dependent on cultivated competencies, supportive technical features, and institutionalized acknowledgments of epistemic and moral agency.

5.2 Answers to the Research Questions

Why is it important to understand designer autonomy as a value?

We learned that a designer's actions should be morally, epistemically, and politically responsible, therefore their autonomy must be linked to their capacity to approach problems and develop solutions in ways with reflect those ideals. It is critical because design is not a neutral, solution-oriented task but a value-laden practice of framing problems and shaping futures. Designers act as mediators of conflicting value systems across individual, organisational, systemic levels. Their framing decisions determine which perspectives are legitimized and which are rendered invisible. In an AI-augmented context, where algorithmic systems can influence, manipulate, or override human judgment, a clear understanding of autonomy is a prerequisite for ethical practice. It ensures designers can maintain accountability and justify their decisions, preventing a drift towards outcomes that are misaligned with human values and reflective commitments. Without this understanding, designers risk becoming executors of system logic rather than reflective, responsible agents.

We defined **Designer Autonomy** as the value of the designer to approach problems and develop solutions in morally, epistemically and politically responsible ways. It is enacted through the reflective negotiation of values across individual, organisational, and systemic frameworks, maintaining authenticity in belief and agency in actions in value-laden design environments. It comprises of 2 distinct yet dependent dimensions of authenticity and agency, explicated as follows:

Designer authenticity refers to the capacity of a designer to critically reflect upon and act in accordance with their own beliefs, values, and ethical commitments within the design process.

Designer agency refers to the ability of a designer to exert meaningful influence over the design direction to act intentionally, challenge constraints, and reshape frames in ways aligned with their reflective commitments.

How is designer autonomy is shaped in collaborative design activity with AI?

We learned that the site where autonomy of a designer is exercised, its effects are impacting the nature of design decisions and outcomes, and where it can be influenced by artificial intelligence is framing practice. The increasing importance of framing practice, both as a cognitive model and a prescriptive methodology in practice encapsulates the widening scope of the designer's responsibilities towards who they impact with their solutioning. A lack a theoretical model of collaborative reflective design practice, where the non-human actors are explicitly participating was identified as a gap. This was addressed by developing a model of collaborative framing practice, or co-framing.

It is the collaborative construction and negotiation of value interpretations between human and artificially intelligent actors in a design sense-making activity.

Four driving factors emerged which shape how designers exercise their autonomy in co-framing activity. Together, they highlight the conditions that inform decision-making and influence the trajectory of outcomes.

Having active control of the pace of the interaction: The designer's organisation of AI interactions through tacit, experience-driven phases. This pacing creates opportunities for reflexivity, shaping whether outcomes follow deliberate judgment or AI's momentum.

Having an expanded sense of role as a designer as a value shaper: The designer's understanding of their professional responsibilities. This influences whether outcomes foreground ethical and societal concerns or settle into narrow, one-dimensional framings.

Having adequate digital literacy about artificially intelligent systems: The designer's level of awareness of AI's mechanics, biases, and limits. This determines whether outcomes are critically curated and contextualised, or unconsciously steered by the model's default patterns.

Having an explicit moral and political lens to experiences and knowledge: The political perspective to their breadth of knowledge that the designer brings to frames. This affects whether outcomes reproduce dominant assumptions or are reoriented toward equity, access, and systemic critique.

We also learned about the risks to both dimensions of autonomy in an interaction.

The risks to Designer Authenticity are:
1. Manipulation: The subtle steering of decisions through embedded tool logic, replacing reflective judgment.

2.Adaptive Preference Formation: The unconscious internalisation of system-favored outputs, narrowing interpretive range.

3. Deceptive or Adaptive Belief Formation: The shaping of justifications by unfavourably biased information, causing epistemic detachment.

The risks to Designer Agency are:
1. Loss of Opportunity: The systematic narrowing of possible framings and exploratory paths of the co-evolution of problem-solution space.

2. Loss of Freedom: The reduction of reflective flexibility and capacity for deviation within design workflows.

3. Loss of Competence: The atrophy of independent framing and decision-making skills and potential.

4. Paternalism: The override of reflective judgment by authoritative system outputs posing as beneficial to the designer.

What technical characteristics of AI tools impact designer autonomy in an interaction?

While AI systems may not be autonomous actors in their own right, their capability and integration into processes means they exert valid influence over the designer's autonomy, therefore shaping the design decisions. To understand this influence, three interdependent characteristics of AI systems are identified to be understood by the designer.

1. Bias: Arises from human choices embedded in AI design that favour certain perspectives. This shapes outcomes by guiding which problem framings are visible, potentially obscuring alternatives.

2. Trustworthiness: Reflects justified confidence in the AI's reliability and accountability. This shapes outcomes by influencing whether designers critically engage with or simply follow AI suggestions.

3. Explainability: The extent to which AI reasoning and logic can be understood and interpreted. This shapes outcomes by enabling designers to evaluate, contest, or adapt suggestions, preserving agency in collaborative framing.

What are the broader implications of the conclusion of this project?

These answers, while providing a focused understanding of designer autonomy in AI-mediated framing practice, have wider implications to design theory and practice. They form a foundational descriptive account that gives way to future research considerations. Potential scope and implications are highlighted into five distinct recommendations.

Moving towards Prescriptive design research on Designer Autonomy

The descriptive account of designer autonomy developed in this project lays the groundwork for future prescriptive design research. By articulating autonomy as a situated value shaped through framing, this work enables more context-sensitive ethical prescriptions about how design tools and practices should be structured. Rather than abstract or universalised ethical claims, future frameworks can be built from the grounded understanding of how authenticity and agency are challenged in practice. This supports the development of normative guidance that is sensitive to both human reasoning and the socio-technical mediation of design decisions. The taxonomy of risks and the concept of co-framing offer concrete analytical tools to inform ethical audits, evaluative criteria, and method development for AI-integrated design contexts. Future work can be in the direction of exploring what are autonomous design methods, and how can an autonomous designer be judged based on their outcomes. There is strong potential in exploring design expertise to designer autonomy, thus informing what aspects of the role and responsibilities of the designer need to be further made explicit in an emerging socio-technical and cultural landscape.

Connecting concepts of Designer Autonomy to User Autonomy in solutions

While this project focuses on the designer as the central human agent, its findings also carry implications for how we conceptualise and protect user autonomy. Designers serve as value mediators between systems and users; thus, any compromise to the designer's autonomy can cascade downstream, limiting the designer's capacity to recognise, prioritise, or defend user values. If designers internalise algorithmic defaults or defer too readily to system framings, the autonomy of users, particularly those not well represented in data or design assumptions, may be structurally neglected. Strengthening designer autonomy therefore becomes a precondition for upholding user autonomy in socio-technical systems, positioning the designer as a crucial ethical actor in the translation of values into outcomes.

Studying designer autonomy across emerging methodologies

This project focuses on framing as the site where designer autonomy is most visibly negotiated, the cognitive model developed here can be extended to other stages of the design process that involve value negotiation and embedded interpretation. AI tools are increasingly integrated across the design workflow, from research synthesis and ideation to prototyping and evaluation. Each of which involves moments where decisions reflect implicit or explicit value judgments. Thus, the conceptual model of collaborative framing can be applied to methodologies across the design process. The co-framing framework, centred on how human and non-human agents influence meaning-making, offers a lens to study how autonomy is exercised or constrained in these phases. For instance user experience design, when AI systems suggest user segments, generate concept variations, or predict desirability scores, they shape the interpretive space in which the designer operates. Applying the model beyond initial framing can uncover how value-laden decisions accumulate over time, revealing autonomy not as a singular moment of authorship but as a continuous negotiation threaded through the entire design process.

Applying model to other design domains and process

Strategic design is studied here as it explicitly addresses the values to be embedded through an organisation's products or services. However, the conceptual model of designer autonomy and the co-framing framework are broadly applicable to other domains of design practice. In areas such as product design, service design, UX, and policy design, similar conditions of ambiguity, system mediation, and value conflict persist. The analytical lens developed here can be adapted to study framing practices and autonomy risks in those contexts, especially where generative AI tools are introduced. Doing so would allow researchers and practitioners to identify discipline-specific tensions, validate the robustness of the proposed framework, and refine it in response to different configurations of designer-system interaction.

Aligning Design Practice with Ethical and Political Frameworks

This project also contributes to ongoing efforts to align design practice with broader ethical and political frameworks. By demonstrating how AI systems mediate framing decisions and affect the designer's capacity to act with moral and epistemic responsibility, it calls attention to the systemic conditions that shape ethical agency. Rather than treating ethics as an external constraint applied post hoc, the findings encourage embedding ethical reasoning into the very structure of design processes. The concept of co-framing highlights how value hierarchies are enacted within tool logic and decision support systems, raising political questions about whose perspectives are privileged or marginalised. In doing so, this research supports a more critically engaged mode of design, which recognises its entanglement with governance, representation, and the distribution of power in socio-technical systems.

5.4 Reflection on the Project

This project began as an exploratory investigation into the nature of designer autonomy, inspired by observations of creative decision-making in complex, value-laden contexts. Over the course of the project, what started as a curiosity-driven inquiry evolved into a structured study of how designers navigate ethical, epistemic, and political responsibilities, particularly when collaborating with AI. The journey was iterative and reflective, requiring continuous adaptation to new insights, constraints, and methodological choices. In many ways, the project mirrored its own subject: exercising autonomy in the face of uncertainty and ambiguity.

Throughout the process, I sought to connect theoretical insights with practical reflection, exploring complex questions about designer autonomy in ways that allowed for both depth and flexibility. By combining theoretical research, autoethnographic reflection, and conceptual modeling, I engaged deeply with analytical thinking and synthesis of knowledge, addressing complex, interconnected issues in design practice. The project demanded personal responsibility, initiative, and ethical judgment, while exploring ways to translate abstract concepts into frameworks and recommendations relevant to design practice. My supervisory team's willingness to adopt an adapted evaluative approach allowed me to focus on designer behaviour rather than conventional product or service outcomes, supporting an alignment between my research ambitions and the educational goals of the program. Through this process, I gained clarity about what I want to explore further, even more than at the outset of the project. The insights generated through this project can inform organisational strategies, guide decision-making processes within organisations by identifying framing pressures, and support the development of frameworks that translate abstract principles into actionable design and business directions.

The project's nature created limitations, but they also presented opportunities that shaped the direction of inquiry. Methodological and time constraints prevented a real-world experimental study, leading to the adoption of autoethnography as appropriate approaches to study autonomy in practice. Certain aspects of designer autonomy, external to the bubble of the designer and the technology, such as the influence of time constraints, organizational pressures, and personal stakes, could not be examined, as did broader sociopolitical dynamics, including drivers of the cultural adoption of AI and technocratic influences. Recognizing these gaps, the project deliberately focused on dimensions of autonomy that could be rigorously explored within the chosen methods, describing what happens rather than prescribing what should be done. In doing so, limitations became a productive lens for refining research questions and deepening insight into the conditions that support responsible, reflective design.

The formal requirement of 840 hours for this project provides one measure of investment, yet the nature of design and creative research often extends beyond such quantified limits. I conclude this project with the feeling of knowing more about what I want to know more than I did when I started. There are always other ways it could have been done or new directions to explore. That balance between a sense of closure and the curiosity to continue is part of what makes any creative process rewarding.

REFERENCES

Anderson, J. (2014). Regimes of Autonomy. Ethical Theory and Moral Practice, 17(3), 355–368. https://doi.org/10.1007/s10677-013-9448-x

Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Christman, J. (2008). Autonomy in moral and political philosophy. In E. Zalta (Ed.), Stanford Encyclopedia of Philosophy.

Cooper, R., & Lilyea, B. (2022). I'm Interested in Autoethnography, but How Do I Do It? The Qualitative Report. https://doi.org/10.46743/2160-3715/2022.5288

Cross, N. (2001). Designerly Ways of Knowing: Design Discipline Versus Design Science. Design Issues, 17(3), 49–55. https://doi.org/10.1162/074793601750357196

Danks, D., & London, A. J. (2017). Algorithmic Bias in Autonomous Systems. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, 4691–4697. https://doi.org/10.24963/ijcai.2017/654

Davis, J., & Nathan, L. P. (2015). Value Sensitive Design: Applications, Adaptations, and Critiques. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains (pp. 11–40). Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_3

Dignum, V. (2019). Responsible Artificial Intelligence. Springer International Publishing. https://doi.org/10.1007/978-3-030-30371-6

Dorst, K. (2011). The core of 'design thinking' and its application. Design Studies, 32(6), 521–532. https://doi.org/10.1016/j.destud.2011.07.006

Dorst, K. (2015a). Frame Creation and Design in the Expanded Field. She Ji: The Journal of Design, Economics, and Innovation, 1(1), 22–33. https://doi.org/10.1016/j.sheji.2015.07.003

Dorst, K. (2015b). Lessons From Design. In Frame Innovation (pp. 41–72). The MIT Press. http://www.jstor.org/stable/j.ctt17kk9fn.8

Dorst, K., & Watson, R. (2023). There is no such thing as strategic design. Design Studies, 86, 101185. https://doi.org/10.1016/j.destud.2023.101185

Dunne, A., & Raby, F. (2013). Design, Fiction, and Social Dreaming. The MIT Press. http://www.jstor.org/stable/j.ctt9qf7j7

Durán, J. M. (2024). Beyond transparency: computational reliabilism as an externalist epistemology of algorithms.

Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). Writing Ethnographic Fieldnotes, Second Edition. University of Chicago Press. https://doi.org/10.7208/chicago/9780226206868.001.0001

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. Journal of Communication, 43(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Fetterman, D. M. (2020). Ethnography: Step-by-Step. SAGE Publications, Inc. https://doi.org/10.4135/9781071909874

Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review. https://doi.org/10.1162/99608f92.8cd550d1

Franssen, M. (2015). Design for Values and Operator Roles in Sociotechnical SystemSociotechnical systems. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains (pp. 117–149). Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_8

Friedman, B., Harbers, M., Hendry, D. G., van den Hoven, J., Jonker, C., & Logler, N. (2021). Eight grand challenges for value sensitive design from the 2016 Lorentz workshop. Ethics and Information Technology, 23(1), 5–16. https://doi.org/10.1007/s10676-021-09586-y

Friedman, B., Kahn, P. H., Borning, A., & Huldtgren, A. (2013). Value Sensitive Design and Information Systems. In N. Doorn, D. Schuurbiers, I. van de Poel, & M. E. Gorman (Eds.), Early engagement and new technologies: Opening up the laboratory (pp. 55–95). Springer Netherlands. https://doi.org/10.1007/978-94-007-7844-3_4

Gibbs, G. R. (2008). Analyzing qualitative data. SAGE Publications. https://uk.sagepub.com/en-gb/eur/analysing-qualitative-data/book225075

Goffman, E. (1974). Frame Analysis: An Essay on the Organization of Experience. Philosophy and Phenomenological Research, 39(4), 601–602. https://doi.org/10.2307/2106908

Goss, H., Tromp, N., & Schifferstein, H. N. J. (2025). Framing System Dynamics for Designers Innovating in Transitions. International Journal of Design, 19(1), 77–94. https://doi.org/10.57698/v19i1.05

Janesick, V. J. (2010). Oral history for the qualitative researcher : choreographing the story. Guilford Press.

Kelly, N., & Gero, J. S. (2022). Reviewing the concept of design frames towards a cognitive model. Design Science, 8, e30. https://doi.org/10.1017/dsj.2022.25

Knutzen, J. (2020). The Trouble with Formal Views of Autonomy. Journal of Ethics and Social Philosophy, 18(2). https://doi.org/10.26556/jesp.v18i2.852
Lawson, B. (2006). How Designers Think – The Design Process Demystified. University Press, Cambridge.

Manzini, E., & Coad, R. (2015). An Introduction to Design for Social Innovation. The MIT Press. http://www.jstor.org/stable/j.ctt17kk7sv

Melnyk, A., Edmonds, B., Ghorbani, A., & van de Poel, I. (2024). Editorial: Modelling Values in Social, Technical, and Ecological Systems. Journal of Artificial Societies and Social Simulation, 27(1). https://doi.org/10.18564/jasss.5361

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2). https://doi.org/10.1177/2053951716679679

Olasik, M. (2023). "Good morning, ChatGPT, Can We Become Friends?" An Interdisciplinary Scholar's Experience of 'Getting Acquainted' with the OpenAI's Chat GPT: An Auto Ethnographical Report. EUROPEAN RESEARCH STUDIES JOURNAL, XXVI(Issue 2), 269–284. https://doi.org/10.35808/ersj/3168

Pasmore, W., Winby, S., Mohrman, S. A., & Vanasse, R. (2019). Reflections: Sociotechnical Systems Design and Organization Change. Journal of Change Management, 19(2), 67–85. https://doi.org/10.1080/14697017.2018.1553761

Poel, I., & Royakkers, L. (2011). Ethics, Technology and Engineering.
Prendeville, S., Syperek, P., & Santamaria, L. (2022). On the Politics of Design Framing Practices. Design Issues, 38(3), 71–84. https://doi.org/10.1162/desi_a_00692

Prunkl, C. (2022). Human autonomy in the age of artificial intelligence. Nature Machine Intelligence, 4(2), 99–101. https://doi.org/10.1038/s42256-022-00449-9

Riessman, C. (2008). Thematic analysis. Narrative Methods For the Human Sciences, 53–76.

Rittel, H. W. J., & Webber, M. M. (1973). Dilemmas in a general theory of planning. Policy Sciences, 4(2), 155–169. https://doi.org/10.1007/BF01405730

Schön, D. A. (1983). The reflective practitioner. How professionals think in action. Basic Books.

Schön, D. A. (1984). Problems, frames and perspectives on designing. Design Studies, 5(3), 132–136. https://doi.org/10.1016/0142-694X(84)90002-4

Schön, D. A. (1988). Designing: Rules, types and worlds. Design Studies, 9(3), 181–190. https://doi.org/10.1016/0142-694X(88)90047-6

Shi, Y., Gao, T., Jiao, X., & Cao, N. (2023). Understanding Design Collaboration Between Designers and Artificial Intelligence: A Systematic Literature Review. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW2), 1–35. https://doi.org/10.1145/3610217

Simon, H. A. (1969). The Sciences of the Artificial. M.I.T. Press.

Strategic Design. (2016). In G. Calabretta, G. Gemser, & I. O. Karpen (Eds.), Eight Essential Practices Every Strategic Designer Must Master. BIS Publishers.

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. Internet Policy Review, 8(2). https://doi.org/10.14763/2019.2.1410

Valkenburg, R., & Dorst, K. (1998). The reflective practice of design teams. Design Studies, 19(3), 249–271. https://doi.org/10.1016/S0142-694X(98)00011-8

van den Hoven, J., Vermaas, P. E., & van de Poel, I. (2015). Design for Values: An Introduction. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains (pp. 1–7). Springer Netherlands. https://doi.org/10.1007/978-94-007-6970-0_40

van der Bijl-Brouwer, M. (2019). Problem Framing Expertise in Public and Social Innovation. She Ji, 5, 29–43. https://doi.org/10.1016/j.sheji.2019.01.003
Verganti, R., Vendraminelli, L., & Iansiti, M. (2020). Innovation and Design in the Age of Artificial Intelligence. Journal of Product Innovation Management, 37(3), 212–227. https://doi.org/10.1111/jpim.12523

What is design? (n.d.). https://www.theicod.org/en/professional-design/what-is-design/what-is-design

Wilkie, A., & Ward, M. (2009). Made in Criticalland: Designing Matters of Concern. Networks of Design, 118–123. https://research.gold.ac.uk/id/eprint/4657/

Williams, B. (1986). Ethics and the Limits of Philosophy. The Journal of Philosophy, 83(6), 351. https://doi.org/10.2307/2026097

Zahedi, M., Heaton, L., Guité, M., Paoli, G., & Reumont, M. (2016). Exploring framing within a team of industrial design students. https://doi.org/10.21606/drs.2016.284

APPENDIX

Problem Definition

What problem do you want to solve in the context described in the introduction, and within the available time frame of 100 working days? (= Master Graduation Project of 30 EC). What opportunities do you see to create added value for the described stakeholders? Substantiate your choice.
(max 200 words)

The identified gap is the absence of specific and explicit understanding of autonomy applicable for design practice. The central tension lies in the shaping of responsibilities when collaborating with automation technologies, rooted in the ambiguity surrounding the value of autonomy. While strategic design positions itself as societally responsive and ethically committed, a lack of clarity around what designer autonomy entails may lead to fragmented individual interpretations, particularly as autonomy is increasingly negotiated between human and machine agents across the design process.

As Prunkl (2022) explains, autonomy comprises both authenticity, which is one's own unmanipulated beliefs, and agency, which is the ability to act independently on those beliefs. Both are at risk of dilution when artificially intelligent systems mediate decisions. Since designers themselves are subject to influence in their framing and reasoning by algorithmic systems, understanding designer autonomy is vital for defining the human role in future design practice. This study seeks to explain designers' role to act meaningfully in increasingly AI-mediated, complex socio-technical systems. The descriptive findings will contribute toward building prescriptive design theory, by informing the ethical interpretation of methodologies, supporting responsible professional practice as well as design education.

Assignment

This is the most important part of the project brief because it will give a clear direction of what you are heading for. Formulate an assignment to yourself regarding what you expect to deliver as result at the end of your project. (1 sentence) As you graduate as an industrial design engineer, your assignment will start with a verb (Design/Investigate/Validate/Create), and you may use the green text format:

Develop an understanding of Designer Autonomy that will explicate its meaning, components, and vulnerabilities for design framing practice in an AI-augmented future.

Then explain your project approach to carrying out your graduation project and what research and design methods you plan to use to generate your design solution (max 150 words)

This project adopts a Value Sensitive Design approach to investigate the nature of designer autonomy in the age of artificial intelligence. The conceptual investigation draws from literature on ethics, philosophy, and design theory to define designer autonomy and identify risks posed by both human and technological factors. The empirical component focuses on Strategic Designers, where ethical and political stakes in framing practices are central to the discipline's aims. Participants will collaborate with an AI agent for a problem framing task. The task is designed to observe expressions of agency through the dynamics of the human–AI interaction, followed by a post-task interview to understand the authenticity behind participants' decisions. A concurrent technical investigation focuses on how the AI system influences participants' perceived autonomy. The data generated will be coded inductively, thematically, and reflexively analysed to develop an understanding of designer autonomy for strategic designers in the age of artificial intelligence.

CHECK ON STUDY PROGRESS

To be filled in **by SSC E&SA** (Shared Service Centre, Education & Student Affairs), after approval of the project brief by the chair.
The study progress will be checked for a 2nd time just before the green light meeting.

Master electives no. of EC accumulated in total _____ EC

Of which, taking conditional requirements into account, can be part of the exam programme _____ EC

★	**YES**	all 1st year master courses passed
	NO	missing 1st year courses

Comments:

Sign for approval (SSC E&SA)

G. Janse Digitally signed by G. Janse Date: 2025.06.12 15:15:42 +02'00'

Name G. Janse Date 12 Jun 2025 Signature

APPROVAL OF BOARD OF EXAMINERS IDE on SUPERVISORY TEAM -> to be checked and filled in by IDE's Board of Examiners

Does the composition of the Supervisory Team comply with regulations?

YES	★	Supervisory Team approved
NO		Supervisory Team not approved

Comments:

Based on study progress, students is ...

★	**ALLOWED** to start the graduation project
	NOT allowed to start the graduation project

Comments:

Sign for approval (BoEx)

Monique von Morgen Digitally signed by Monique von Morgen Date: 2025.07.01 10:17:35 +02'00'

Name Monique von Morgen Date 1 Jul 2025 Signature

Personal Project Brief – IDE Master Graduation Project

Name student	Ameya Sawant	**Student number**	5,950,899

PROJECT TITLE, INTRODUCTION, PROBLEM DEFINITION and ASSIGNMENT

Complete all fields, keep information clear, specific and concise

Project title On the Nature of Designer Autonomy: Strategic Framing Practice in the Age of AI

Please state the title of your graduation project (above). Keep the title compact and simple. Do not use abbreviations. The remainder of this document allows you to define and clarify your graduation project.

Introduction

Describe the context of your project here; What is the domain in which your project takes place? Who are the main stakeholders and what interests are at stake? Describe the opportunities (and limitations) in this domain to better serve the stakeholder interests. (max 250 words)

The drive for autonomous technologies comes with both opportunities for societal transformation and risks of undermining fundamental human values, one of which is autonomy. Human autonomy is a central tenet in frameworks guiding responsible development and use of AI, yet inconsistent definitions of the value may limit their effectiveness in practice (Prunkl, 2022). As artificial intelligence increasingly shapes professional domains, answers to how human roles shift must be grounded in clear conceptualisations of human values within domain-specific practices (Dignum, 2019).

With the evolved understanding of the socio-political role of designers, framing practice emerges as the site where value structures are either reproduced or reconfigured through the designer's interpretive decisions (Prendeville et al., 2023). For the domain of Strategic Design, which is the shaping of organisational strategies through design approaches and practices for the benefit of people, organisations, and the planet alike (Calabretta et al., 2016), this interpretive role becomes particularly sensitive, as designers shape the future through values and meanings embedded in solutions. Thus, growing algorithmic mediation in design processes calls for renewed attention to autonomy as a condition for ethical and responsible design practice.

This project chooses design students as the primary stakeholders, as they navigate emerging methodologies while still developing their skills as designers. Their interest lies in maintaining autonomy while ensuring ethical outcomes in future design practice. This project proposes to study what autonomy means for designers, and how it is exercised and influenced when working with AI as a mediating technology in design framing tasks.

➔ *space available for images / figures on next page*

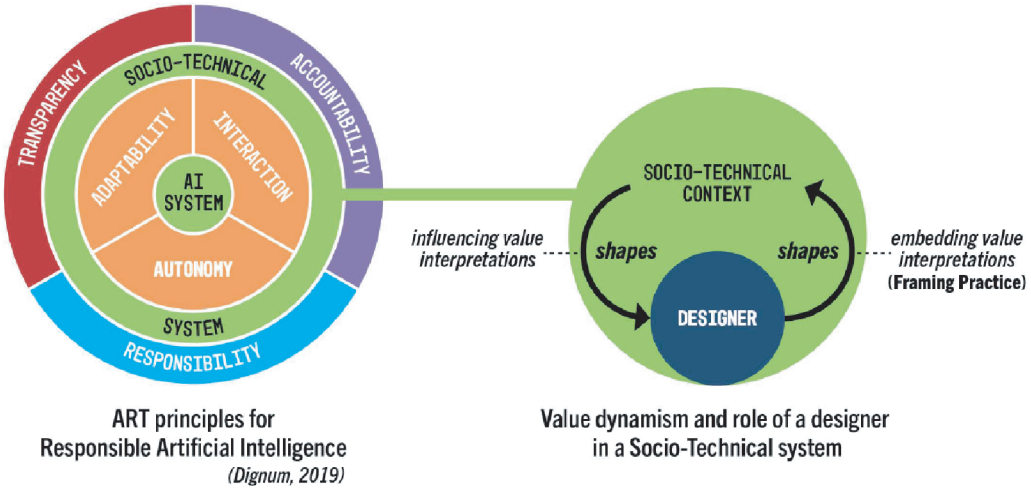


image / figure 1 Relationship between design frameworks and value dynamism in socio-technical contexts of design

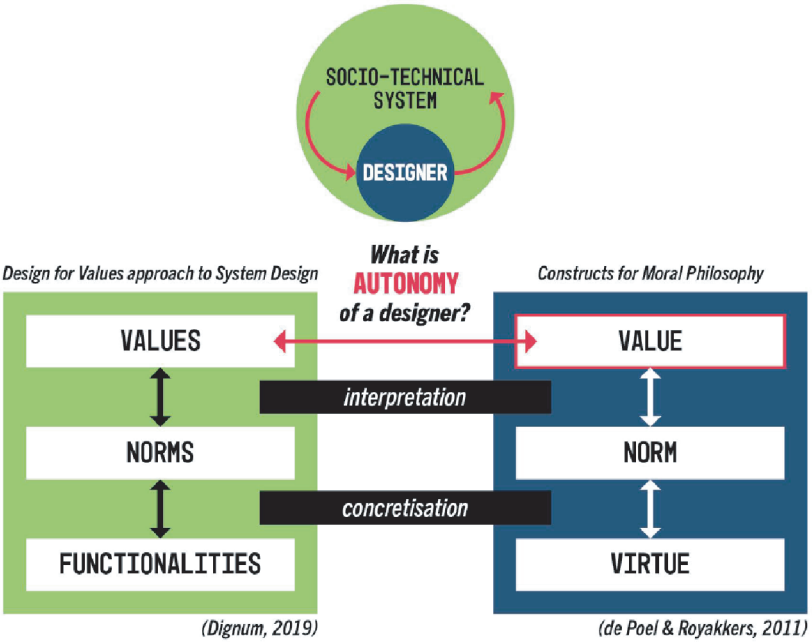


image / figure 2 Intersecting design and philosophy to explicate Designer Autonomy in a dynamic socio-technical context

Project planning and key moments

*To make visible how you plan to spend your time, you must make a planning for the full project. You are advised to use a Gantt chart format to show the different phases of your project, deliverables you have in mind, meetings and in-between deadlines. Keep in mind that all activities should fit within the given run time of 100 working days. Your planning should include a **kick-off meeting**, **mid-term evaluation meeting**, **green light meeting** and **graduation ceremony**. Please indicate periods of part-time activities and/or periods of not spending time on your graduation project, if any (for instance because of holidays or parallel course activities).*

Make sure to attach the full plan to this project brief.
The four key moment dates must be filled in below

Kick off meeting	14 Apr 2025
Mid-term evaluation	4 Jun 2025
Green light meeting	21 Jul 2025
Graduation ceremony	29 Aug 2025

In exceptional cases (part of) the Graduation Project may need to be scheduled part-time. Indicate here if such applies to your project

Part of project scheduled part-time	
For how many project weeks	
Number of project days per week	

Comments:

Motivation and personal ambitions

Explain why you wish to start this project, what competencies you want to prove or develop (e.g. competencies acquired in your MSc programme, electives, extra-curricular activities or other).

Optionally, describe whether you have some personal learning ambitions which you explicitly want to address in this project, on top of the learning objectives of the Graduation Project itself. You might think of e.g. acquiring in depth knowledge on a specific subject, broadening your competencies or experimenting with a specific tool or methodology. Personal learning ambitions are limited to a maximum number of five.
(200 words max)

This project stems from my interest in exploring how design theory should adapt to maintain designer roles and responsibilities in an increasingly technologically mediated world, leading to develop a value-sensitive study on the autonomy of designers.

I aim to demonstrate my ability to develop and conduct design research that translates ethical concerns into structured descriptions of design practice. This includes proving analytical depth in interpreting complex theoretical inputs, synthesising them into a coherent design research structure, and showing awareness of the social, political, and ethical implications of design practice in emerging technology landscapes. As a Strategic Design student, I want to contribute to the discipline by examining how underlying value drivers shape framing decisions and influence strategic design outcomes, with designer autonomy as the critical site where these dynamics are negotiated. Personally, I want to learn how to meaningfully understand philosophical concepts and connect them to the domain of design. I want to improve my ability to visualise complex systems, relationships, and arguments graphically, without oversimplifying nuance or loosing critical depth. My goal is also to build confidence in expressing my perspective as a designer and strengthen my communication skills across verbal, visual, and written formats.