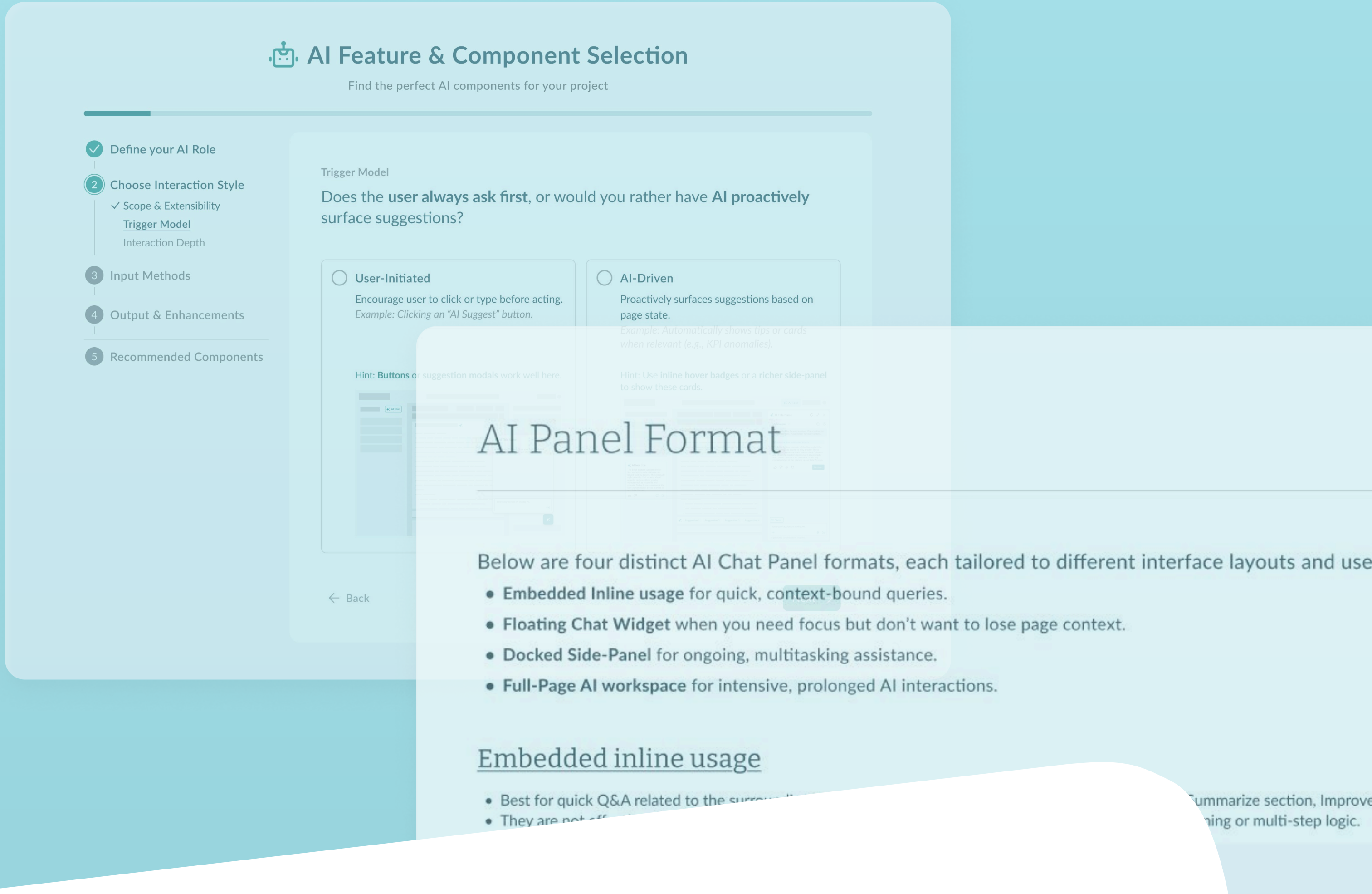


Designing for Scalable **AI** Interactions in Financial Workflows

A Framework for Internal AI Tool Integration at Van Lanschot Kempen

Graduation Project
Msc. Design for interaction
Ziyue Lu

Chair - Evangelos Niforatos
Mentor - Rachel Chen
Company Mentor - Paulina Meraza Farfan

TU Delft

VAN LANSCHOT KEMPEN

Designing for Scalable AI Interactions in Financial Workflows

A Framework for Internal AI Tool Integration at Van Lanschot Kempen

Graduation Project

Author
Ziyue Lu

Supervisory Committee
Evangelos Niforatos — Chair
Rachel Chen — Mentor
Paulina Meraza Farfan — Company Mentor

Institution
Master of Science in Design for Interaction
Faculty of Industrial Design Engineering
Delft University of Technology (TU Delft)

In Collaboration with
Van Lanschot Kempen

Acknowledgements

How long is six months?

It is enough time to witness the passing of seasons, to see ideas grow and transform, and to watch colleagues gradually become friends. Yet it is also short enough that the project's start still feels vivid, carried forward by deadlines and small urgencies.

I could not have known every turn the work would take, but I began this project confident that it was achievable, because I had a supportive environment and generous people around me.

My deepest thanks go to my supervisor, Evangelos Niforatos and Rachel Chen, for steady guidance and rigorous feedback. The thoughtful critique, patient encouragement, and practical advice provided the shape and discipline this work needed. Those conversations repeatedly helped turn loose ideas into clearer questions and usable design outcomes.

I am grateful to the colleagues at Van Lanschot Kempen, especially my teammates in the UX Design & Research Team — Paulina Meraza Farfan, Xumeng Zhang, Martijn de Jongh and Elvira Tretiakova, who opened their calendars, shared their time and expertise, and let me into their workflows. They have offered tremendous support throughout this project and helped me far more than I could have asked for.

Thanks are also due to product owners, bankers and other colleagues for giving invaluable input during research phase, their candid perspectives grounded the research and made the prototypes realistic and relevant. I would also like to thank my manager, Lisanne Smink and Geert Samplonius, whose support, trust, and care made this project possible. Without the opportunity they provided, this journey would not have taken shape. Their cooperation, suggestions, and everyday kindness made fieldwork a true pleasure.

To my friends and family, I am deeply thankful for the balance they brought to life. Their meals, jokes, and reminders to rest made it possible to sustain focus across the busy months. Their support was the steady current beneath this project.

Finally, a quiet acknowledgement to myself. Completing this project required persistence, late nights, and many small decisions. I am proud of the effort and the learning it required. Thank you for the discipline and curiosity that brought this thesis to life.

To everyone who read, tested, advised, or cheered me on along the way: thank you!

Abstract

This project investigates how to design a scalable, human-centered AI interaction framework for internal wealth management workflows at Van Lanschot Kempen. The project responds to operational fragmentation caused by multiple independently developed AI features, inconsistent interaction logic across tools, and varied trust levels among staff. The research phase combined an exploratory field study consisting of eleven semi-structured interviews with private bankers, relationship managers and investment advisors, thematic analysis of workflow pain points, and scenario mapping of core tasks such as text writing, information lookup and advice generation.

The iterative design process, combining research synthesis, prototyping, and refinement, translated these insights into three interlocking outcomes. A set of design principles emphasizes human control, transparency, clarity and actionable feedback. An atomic component library provides reusable UI elements and panel formats to ensure consistent interaction patterns across features. An AI Feature and Component Selection Wizard guides designers and product owners through role definition, interaction style, input and output configuration, and container selection, producing a concrete specification that supports cross-functional handoff.

A formative evaluation engaged three designers and two product owners in moderated sessions. Results show strong willingness to adopt the framework, increased confidence in design decisions and practical value for aligning product and design stakeholders. Recommendations for refinement include clearer terminology, richer visual previews and improved mapping to the Figma component library.

Contributions of the thesis include a method for operationalizing human-centered AI principles into a domain tuned design system for regulated environments and a practical tooling approach that shortens the path from concept to implementation.

Key words: human-centered AI, interaction design, design system, generative AI, financial workflow

Structure

Table of content

1. Introduction

This chapter lays the foundation for our exploration of AI in a wealth management context.

It firstly explores the broader evolution of generative AI and its implications for financial services. It then examines Van Lanschot Kempen's internal adoption of AI, highlighting both the promise and the fragmentation of current tools. Finally, the chapter defines the project's scope and research intent, detailing how user-centered insights will inform a scalable framework of reusable components and interaction patterns. Together, these sections establish the context, case study, and methodological roadmap for developing a unified AI interaction design system.

1.1 Generative AI in financial context

Over the past few years, Generative Artificial Intelligence (GenAI) has undergone a remarkable evolution, moving well beyond simple text-completion tasks into an era of truly multi-modal capabilities. Models such as OpenAI's GPT series, Anthropic's Claude, and Google's Gemini now power not only natural language generation but also code synthesis, chart creation, and even in-depth document analysis. This rapid advancement illustrates GenAI's potential to augment human expertise across diverse domains and has spurred a shift from consumer-focused chatbots toward enterprise-grade integrations. Platforms like Microsoft Copilot, Notion AI, and SAP Joule are embedding these underlying AI engines directly into knowledge-work applications, enabling professionals to invoke AI assistance without leaving their primary workflows .

In financial services, this trend is particularly pronounced. Recognizing that AI can automate compliance checks, generate executive summaries of lengthy reports, and surface actionable insights from vast, structured datasets (Balakrishnan, 2024). Yet, unlike many other sectors, finance faces heightened demands for accuracy and explainability. Even a minor error in a risk assessment or transaction summary can trigger regulatory penalties or reputational damage. This sensitivity has driven banks and financial institutions to approach AI deployment with caution, balancing the promise of efficiency gains against the need for rigorous oversight (Sharma et al., 2024).

Moreover, financial operations often involve complex, tightly defined processes. From Know Your Customer (KYC) checks to portfolio rebalancing, it require handling highly structured data and strict audit trails. Traditional machine-learning systems continue to play an important role in this context, with it's value focused on automating repetitive, rules-based tasks, freeing human experts to focus on judgment-driven activities. Generative AI, by contrast, extends automation into unstructured, high-value knowledge work. Large language models can summarize long meeting notes, draft client communications, synthesize across documents, and power conversational assistants that bridge human judgment and dispersed data sources. This expanded capability makes LLMs particularly useful for tasks of sense-making and drafting that traditional models address less naturally (Lewis et al., 2021).

However, GenAI models also introduce risks and constraints that have direct design implications for financial tools. LLMs can produce fluent but ungrounded outputs, a phenomenon widely discussed as hallucination, which is unacceptable in high-stakes contexts unless mitigated (Huang et al., 2025). Practical technical mitigations include retrieval-augmented generation to anchor outputs in source documents and domain-specific fine-tuning to improve relevance (Gururangan et al., 2020). From an interaction design standpoint, these technical realities translate into concrete requirements: interfaces must surface sources and confidence, support easy verification and correction, and preserve clear human oversight for high-impact actions. Such priorities shaped the interaction patterns and component choices developed in this project. Without careful thoughts on integration, GenAI tools risk becoming siloed experiments rather than embedded capabilities that support everyday work (Cao, 2022).

Finally, the industry's relationship-driven nature sets it apart. As a result of information technology's pervasiveness in today's society, many banking companies have grown to depend on AI as a means to simplify their operations, provide superior service to customers, and strengthen their communication channels (P et al., 2023). However, private bankers, relationship managers, and investment advisors build trust through personal interaction, empathy, and domain expertise, which are qualities that AI cannot fully replicate. Instead, GenAI must act as a collaborative partner, augmenting human decision-making rather than replacing it. Research has shown that when GenAI is designed to complement human strengths, providing data-driven suggestions while leaving final judgment in human hands, the combined human-AI team consistently outperforms either working alone (Kahn et al., 2020). These factors, the race toward multi-modal GenAI, the enterprise integration wave, and the unique demands of financial services, together form the backdrop against which this project seeks to design a coherent, trustworthy, and scalable GenAI interaction framework for Van Lanschot Kempen.

1.2 Van Lanschot Kempen: a case of internal AI adoption

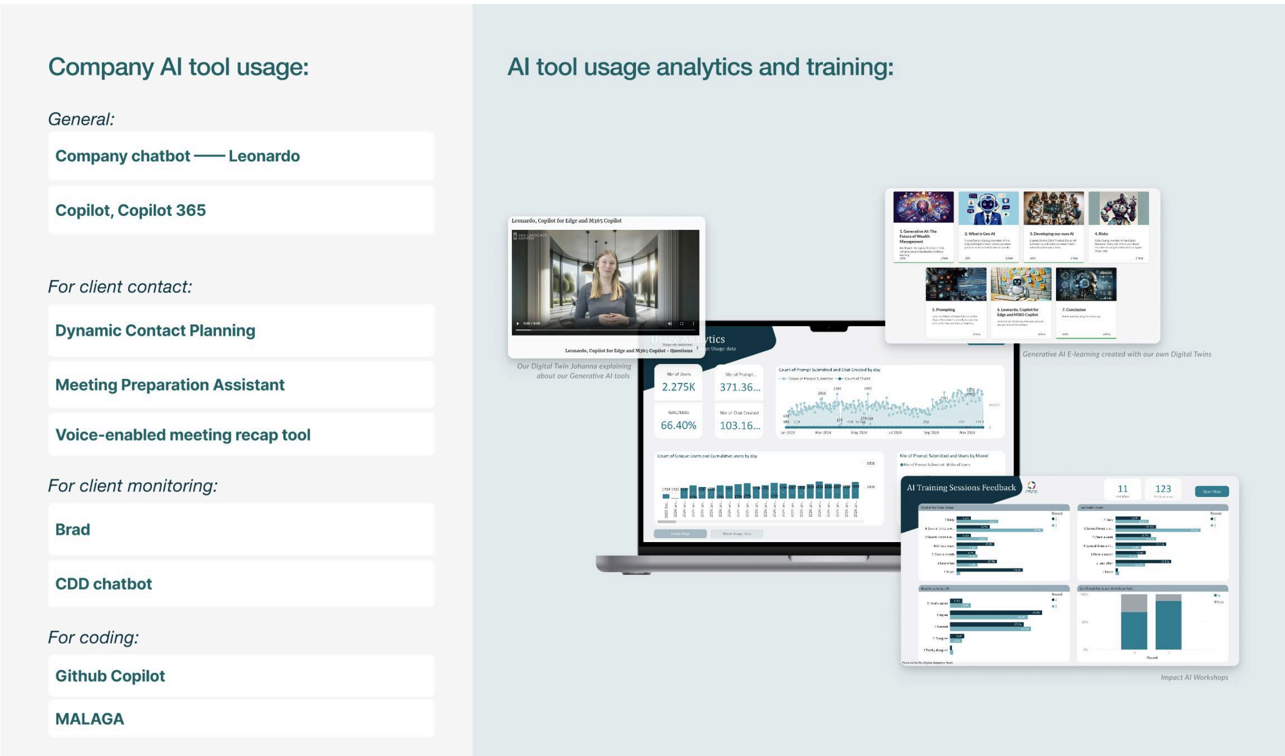


Figure 1: AI adoption in Van Lanschot Kempen

Van Lanschot Kempen is a Dutch private bank with over 200 years of history, specializing in wealth management, investment advice, and family office services. As a mid-sized institution, it combines traditional relationship-driven banking with a digital-first strategy. In recent years, the bank has invested in artificial intelligence to streamline internal processes, both to reduce the administrative burden on bankers and to surface actionable insights more quickly. These efforts are coordinated by a small central team in the Knowledge & Intelligence Design department, alongside distributed initiatives led by product owners in various divisions.

Today, several AI-driven features are embedded within Van Lanschot Kempen's primary internal platform, ClientCenter. These include:

- **AI Summary**: automatically condenses past meeting notes into concise highlights;
- **Meeting Note Documentation**: transcribes and structures spoken inputs via microphone into formal notes;
- **Actions & Suggestions**: surfaces time-sensitive reminders (e.g., upcoming deadlines, client birthdays) directly in users' task lists.

While each tool delivers clear value, they were developed independently by different teams. As a result, interaction patterns, visual styling, and control flows vary widely across features, creating a steep learning curve for employees who must switch contexts frequently. Moreover, inconsistencies in labeling, confirmation dialogs, and feedback mechanisms can undermine trust in the AI outputs, especially when tasks involve sensitive financial data or compliance requirements.

This fragmentation highlights the need for a unified AI interaction framework: one that preserves the benefits of localized innovation while ensuring consistency, transparency, and scalability across the entire organization. By focusing on company's unique combination of relationship-driven services and digital aspirations, this project seeks to develop a design system that empowers both the designers and product owners who build AI features, and the bankers and analysts who rely on them in their day-to-day work.

1.3 Project goal and approach

1.3.1 Project goal

This graduation project is dedicated to addressing the fragmentation of AI tooling at Van Lanschot Kempen by developing a scalable AI interaction design framework tailored for internal use. Rather than creating another standalone AI application, the project scopes a meta-level solution: a system of reusable components, interaction patterns, and practical guidelines that designers and product owners can leverage to build coherent, transparent, and trustworthy AI features across ClientCenter and future platforms. By standardizing how AI is surfaced, controlled, and explained while preserving the flexibility required by different teams, this framework aims to reduce redundant effort, shorten onboarding curves, and strengthen user confidence in AI-powered workflows.

As illustrated in Figure 2, the universe of possible AI interaction patterns is vast (the outer ring), but only a subset of these patterns are truly applicable within structured professional workflows (the middle ring). Of those workflow-focused patterns, Van Lanschot Kempen's current AI tools occupy an even smaller part (the inner circle), centered on text generation, information lookup, and advisory prompts.

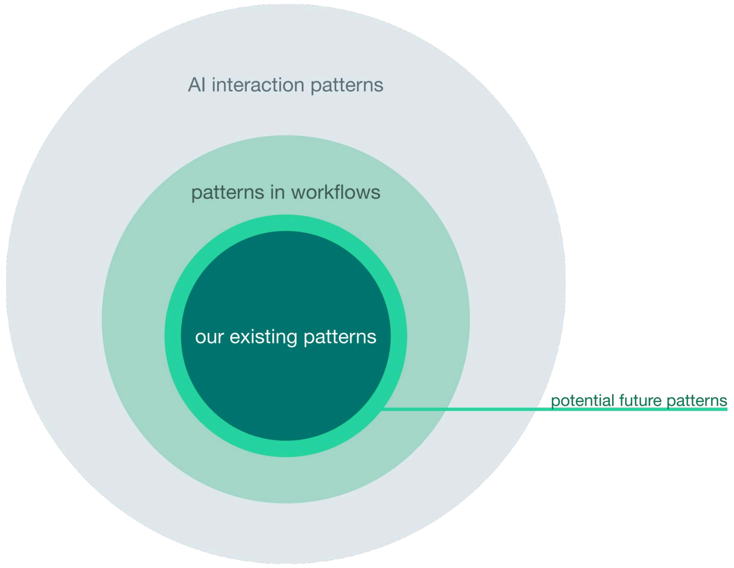


Figure 2: Visualization of AI interaction pattern scope within Van Lanschot Kempen workflows

Building on this foundation, the project deliberately seeks to broaden the spectrum of workflow-applicable patterns, surfacing new interaction paradigms that remain coherent with the bank's process constraints. By expanding this middle layer with potential future patterns, we lay the groundwork for more efficient, integrated AI tools that can evolve alongside the company's workflow needs without sacrificing consistency or user trust.

1.3.2 Project scope

The primary users of this framework are designers and select product owners in the company, who will directly engage with its component library and supporting selection tools. Their work shapes every new AI feature, from quick note summarization widgets to more complex generative assistants. Although these individuals are the framework's direct beneficiaries, the ultimate impact extends to all employees who rely on AI-enabled tools, improving their daily experience by ensuring consistency in labels, controls, and interaction logic.

1.3.3 Project approach

To ground the framework in real organizational needs and user expectations, the project is twofold. Firstly, it seeks to understand current practices, pain points, and attitudes through semi-structured interviews (45–60 minutes) with bankers, relationship managers, and investment advisors and thematic analysis of AI usage in primary scenarios.

Secondly, it aims to translate these insights into concrete design artifacts, including a component library organized via atomic design principles, extending the company's existing *Marvel* system with AI-specific UI elements. And an feature & component selection wizard, guiding designers through critical decision points to arrive at a tailored set of components.

By bridging empirical research and practical design, this approach ensures that future AI integration at Van Lanschot Kempen is both user-centered (aligned with the research questions outlined in Section 3.1), and strategically aligned with the design goal defined in Section 4.1. In doing so, the project delivers a replicable framework that supports the company's digital ambitions and contributes to the broader practice of AI interaction design in regulated enterprise environments.

2. Related work

To position this project within the broader academic and industry context, this chapter reviews existing research and design practices related to AI interaction in professional settings. As outlined in the introduction, the integration of AI into organizational workflows within complex, high-stakes environments like finance, poses unique challenges in trust, usability, and system consistency, requiring an understanding of both technical capabilities and human perceptions of AI as work-used tools.

The chapter first explores human-computer interaction (HCI) literature, emphasizing collaboration, interaction paradigms, and design principles for clarity, control, and explainability. It highlights the need for human-centered safeguards in decision-sensitive tasks and issues with generative models, which impact component design and workflows. Next, it examines industry practices, including vendor guidance and resources like Microsoft, ShapeOf.AI, and the People + AI Guidebook, which offer interaction modalities to aid adoption. However, case studies show generic resources are often insufficient; regulated organizations need domain-specific translations, governance tools, and cross-functional handoff mechanisms.

This review establishes the project's conceptual foundation and identifies a gap: limited work translates HCI theory into domain-specific design framework. By applying these insights to Van Lanschot Kempen's AI ecosystem, the project aims to translate theoretical knowledge and fragmented practices into a coherent, practical framework.

2.1 Human Computer Interaction(HCI) theory

Many evidence from empirical research and real-world applications shows that collaboration between humans and AI yields superior outcomes compared to scenarios that involve only humans or only AI (Kahn et al., 2020). Jiang et al. (2024) extend this concept in their article by discussing the symbiosis between humans and AI. They present it as the modern evolution of the man computer symbiosis first envisioned by Licklider in 1960. In this mutually beneficial relationship, the computational power and analytical capabilities of AI augment human information processing, problem solving and decision making while humans contribute contextual understanding, intuition, empathy and ethical judgment to enhance AI accuracy and adaptability.

*This perspective supports the foundational ambition of my project: AI within internal workflows should not aim to **replace professionals**, but to **complement** them, serving as a catalyst for smarter decisions while respecting human agency, judgment, and domain expertise.*

Yet achieving this harmonious collaboration remains challenging because many AI algorithms function as opaque black boxes. This lack of transparency often leads to low trust and acceptance, and users may experience communication breakdowns or encounter ethical concerns such as potential job displacement and loss of autonomy (Huang et al., 2023). To bridge this gap, the authors advocate for a human centered AI approach that integrates multidisciplinary research from fields such as human computer interaction, cognitive science and ethics. Their proposed practices include gathering feedback from humans in the loop, employing explainable AI techniques, designing user friendly interfaces that cater to individual preferences and implementing responsible AI frameworks that address fairness, privacy and security.

These principles directly inform the design goals of my AI interaction framework, especially in the context of financial organizations like Van Lanschot Kempen, where accuracy, clarity, and human oversight are essential. In designing components for generative AI tools, I place strong emphasis on interface clarity, user control, and explainability, enabling users to understand how the AI reaches conclusions, adjust its outputs, and confidently retain authority over critical decisions.

Building on the broader HCI research around human-AI interaction, Elshan et al. (2022) provide a systematic review of empirical studies examining the design elements that influence user acceptance of intelligent agents. The authors emphasize that the level of anthropomorphism or formality of an agent can significantly impact perceived trust and professionalism. In addition, they highlight the importance of behavioral proactiveness: AI systems that offer timely, context-aware suggestions tend to be more positively received, so long as they strike the right balance between assistance and autonomy. Communication clarity also plays a central role, especially in environments where transparency and explainability are crucial. It argues that users are more likely to accept and rely on AI agents when interactions are understandable and well-aligned with users' expectations and goals. Finally, the paper calls for more domain-specific, longitudinal research to examine how user trust in AI develops over time. This insight reinforces the relevance of studying AI interaction design within the organizational and regulatory context of finance, where long-term acceptance and internal adoption are key to sustainable integration.

This reinforces the decision to ground my research in the real, evolving workflows of wealth management professionals. By analyzing employees' preferences, trust levels, and adoption barriers across different AI use cases, I aim to build a scalable design system that reflects not just one-off usability concerns but the deeper behavioral dynamics of long-term AI adoption within regulated environments.

The theoretical insights above stress the need for AI interactions that are transparent, contextual, and aligned with users' mental models. Turning those principles into workable solutions, the next part examines two concrete case studies that illuminate how research and industry are approaching AI in financial settings. These more specific examples expose practical patterns, gaps, and governance challenges that directly inform the design priorities and implementation choices adopted in this project.

case study 1: domain-specific AI guidelines for financial services

A recent case study of a leading digital finance company in South Korea highlights a central lesson for HCI work on AI in regulated domains: generic human-AI guidelines are useful, but they do not fully address the situated constraints and stakeholder dynamics of specific industries (Cho et al., 2024). The authors report that practitioners found off-the-shelf resources to be helpful as starting points, yet insufficient when designing for banking workflows. This finding suggests that design theory for human-AI interaction needs an intermediate layer of domain-specific translation, which could be a set of patterns and component prescriptions that map high-level HCI principles to the operational realities of finance.

The case study also draws attention to the organizational dimension of AI design: guidelines are not only technical artifacts but socio-technical ones that must reflect inter-stakeholder dynamics within a company (Cho et al., 2024). In practice, product teams, compliance officers, data engineers and frontline bankers each hold different mental models, risk tolerances, and vocabularies for success. Consequently, domain-specific guidance should include not only UI patterns but also recommended governance practices, handoff artifacts, and communication templates that make trade-offs explicit across roles.

Methodologically, the study highlights the value of co-design and iterative

knowledge transfer between domain experts and design/engineering teams. Tools that help domain experts express tricky knowledge in structured, reusable forms can materially improve downstream model development and UI design (Park et al., 2021). For a financial context this means creating artifacts that capture domain concepts, typical edge cases, and justification patterns, which are materials that both inform model fine-tuning and drive UI affordances such as provenance links, confidence indicators, and escalation flows. Embedding such artifacts into the design system reduces the translation gap between "what the bank needs" and "what the model can safely produce."

Building from this starting point, my research can both validate the importance of domain-specific AI guidance and extend the academic and industry knowledge base in a way that helps the field progress more rapidly. By empirically investigating where generic guidelines fall short and domain constraints in practice, the work can provide concrete, evidence-backed prescriptions rather than only conceptual advice. Framing the guideline as a socio-technical artifact further requires moving beyond UI rules to think about how people will actually use them in their daily work: what contextual cues and scaffolds do users need, what kinds of in-tool guidance reduce ambiguity, and which governance or handoff documents will make adoption feasible across teams.

Accordingly, my research can probe not only interface preferences but also the organizational processes that enable integration, including how designers and other stakeholders should coordinate to move a feature from spec to production, what artifacts (checklists, templates, exportable specs) ease that handoff, and how to govern it to preserve auditability and trust. Finally, recognizing the framework as a living system implies design for evolution: the research can recommend practices for iterative maintenance so the framework remain current as models, data sources and regulatory expectations change.

Together with empirical grounding, socio-technical framing, cross-role integration, and planned evolution, these strands point to a research agenda that is both practically useful for firms and theoretically generative for HCI.

case study 2: Large Language Models(LLM) in finance

A survey of large language model applications in finance synthesizes how LLMs are being used, the technical affordances they enable, and the practical challenges that arise when these models are put into production in regulated settings. This survey maps a range of high-value use cases that are directly relevant to interaction design: automated synthesis of lengthy reports and meeting notes, draft generation for client communications, cross-document question answering, and conversational assistants that help staff retrieve contextualized information from dispersed systems (Li et al., 2024). These capabilities extend the reach of automation from structured, rules-based tasks into the messy, narrative work that bankers and advisors perform every day, including making sense of client histories, producing readable briefings, and turning unstructured conversation into actionable next steps.

Alongside use cases, the survey emphasizes a consistent set of technical limitations and risks that carry clear implications for UI design. Chief among these are the tendency of LLMs to generate fluent but potentially incorrect information (hallucination), sensitivity to stale or narrow training data, and difficulties in validating provenance when outputs recombine multiple sources. The literature points to practical mitigations: retrieval-augmented generation to ground responses in indexed documents, domain fine-tuning to improve relevance, and hybrid pipelines that route high-risk queries through rule-based checks or human review. These engineering patterns influence the design requirements, setting standard that interfaces must make source material visible, provide succinct uncertainty signals, and make verification and correction straightforward for users.

These survey findings point directly to actionable directions for this project. It can incorporate model-level constraints into the design materials by making model grounding and output provenance first-class concerns in the framework.

The survey's technical synthesis translates into several concrete component and workflow requirements. Components should surface provenance and clickable source links; response cards should include confidence or uncertainty indicators and "why" summaries that explain which documents or facts shaped a recommendation. Interfaces must also provide lightweight verification tools, such as quick links to the original documents, inline edit and accept/reject controls, and an obvious escalation path to a human expert when outputs are uncertain or high-impact. From a systems perspective, the survey suggests that certain outputs should require mandatory grounding and model configuration choices should be exposed as configurable and safeguarded options for product teams. Future work can test whether these defaults and UI affordances actually increase users' ability to detect and correct errors, and whether they influence adoption and trust in real workflows.

Beyond immediate UI prescriptions, the survey brings a broader methodological stance: design work must be informed by the evolving technical stack. The project can therefore treat the design framework not only as static artefacts but as a place for cross-functional conversations about suitability, acceptability, and feasibility. Embedding these model-aware decisions into design artifacts helps bridge the gap between what models can produce and what practitioners can safely deploy, and sets the stage for iterative updates as model capabilities, data sources and regulatory expectations evolve.

2.2 Industry practices

While academic research has laid important theoretical foundations for human-AI collaboration, industry has also started exploring practical frameworks for building better AI user experiences. These efforts offer valuable reference points for interaction design, especially in the context of generative AI tools.

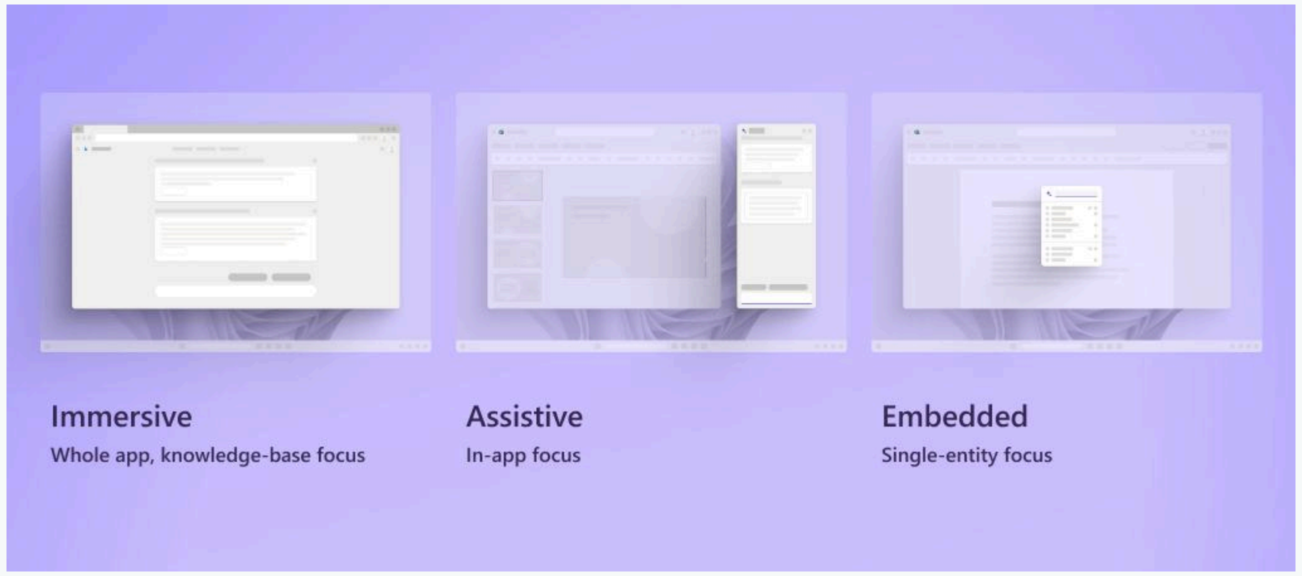


Figure 3: Microsoft - Three framework variations for building custom copilot experiences

For example, Microsoft has begun to translate HCI principles into concrete design strategies for AI-driven interfaces. In '*Creating a Dynamic UX: Guidance for Generative AI Applications*' (miglaros, n.d.), Microsoft outlines how to structure "copilot" experiences through three complementary frameworks: immersive, for focused full-screen workflows; assistive, which weaves AI support directly into existing applications; and embedded, where AI capabilities surface contextually at individual touch-points. These frameworks are supported by strong input/output design, adaptive collaboration, and the flexibility to blend different interaction modes depending on task and context.

These classifications are directly helpful as a starting point for the framework's own container and interaction categories, and the project will adopt this structural logic when organizing component formats and guidance.

At the same time, the Microsoft guidance is intentionally general-purpose and aimed at broad enterprise scenarios; it does not solve domain-specific constraints that are critical in financial services. In the context of banking and wealth management, , where precision, risk sensitivity, and regulatory compliance play a central role, additional requirements must be layered on top of these general patterns.

Similarly, Emily Campbell's '*ShapeOf.AI*' (The Shape of AI | UX Patterns for Artificial Intelligence Design, n.d.)is a community-driven library that catalogs and organizes reusable interaction patterns for AI-powered experiences. It introduces a shared vocabulary across five categories: identifiers, way-finders, inputs & prompts, tuners, and trust indicators, each addressing a specific facet of the user's journey when interacting with AI. The site aims to support more coherent and human-centered AI design by guiding users in how to start, control, and understand their interactions with intelligent systems.

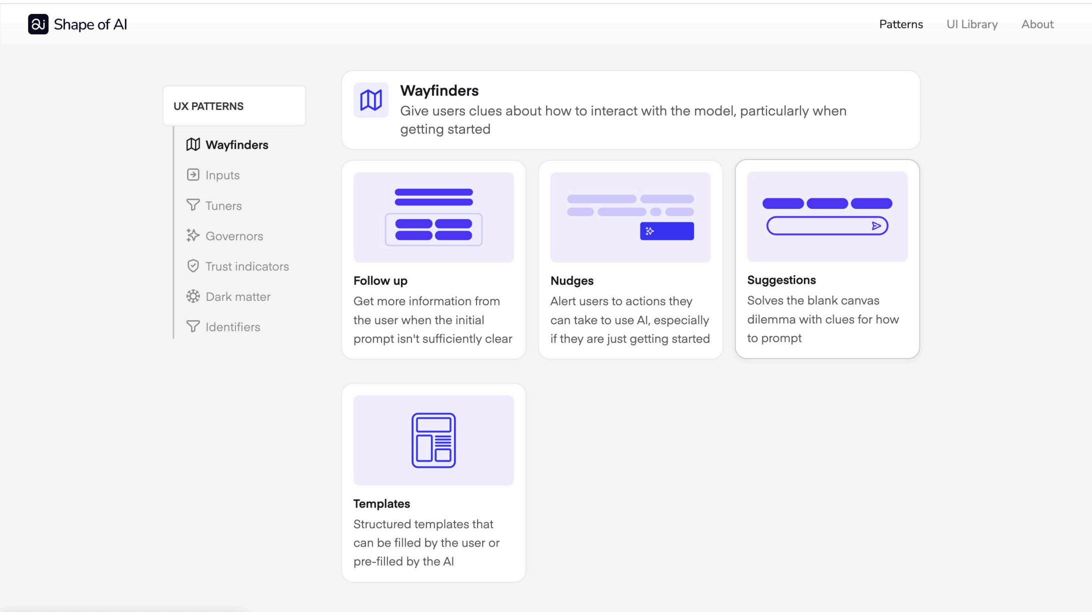


Figure 4: Shape of AI website interface

Their pattern-first approach is valuable because it emphasises composability: small, well-documented components and usage rules can be combined into richer interaction templates, which in turn supports faster, more consistent feature development. It provides a modular mindset and useful checklist for what a modern AI UI pattern set should contain and how those parts can be mixed and matched.

Nevertheless, ShapeOf.AI's corpus tends to focus on consumer and creative scenarios such as chat-based assistants, generative creative tools, and open-ended exploration, so several categories typical of financial workflows are under-represented. The resource therefore functions best for this project as a gap-finder and inspiration source. The project will reference ShapeOf.AI's modular classification and prompt-mode distinctions, but will supplement them with finance-specific elements.

case study 3: using the People + AI Guidebook in practice

Another important industry resource is the People + AI Guidebook, a living collection of practitioner-facing recommendations for designing human-centered AI (People + AI Guidebook, n.d.).

A recent research has directly investigated how practitioners actually use such resources in everyday product work. Yildirim and colleagues conducted interviews with 31 designers and product managers to study how the People + AI Guidebook is applied in industry settings and what gaps remain (Yildirim et al., 2023). Their analysis found that teams rely on the guidebook not only to solve concrete design problems, but also as an educational tool, a shared vocabulary for cross-functional conversations, and a scaffold for developing internal, organisation-specific resources and checklists.

The study's most salient finding is that practitioners' request for stronger support in early-phase ideation and problem formulation, which has direct implications for the current project. Whereas many guidelines are most useful in later design stages for shaping an interaction once an AI capability is already chosen, teams struggle earlier on to decide whether AI is the right solution for a given problem and how to frame success criteria. For a framework intended to guide internal AI feature development, this suggests the need to move beyond static guidance toward tools that actively help the upstream decisions: helping teams define goals, identify what inputs and data are required, and decide on acceptable levels of automation or human oversight. The People + AI Guidebook study highlights the value of packaging high-level recommendations into action-oriented artifacts that support early judgments along with late-stage designs.

For this project, firstly, the People + AI Guidebook and its study validate the usefulness of consolidated, practitioner-oriented guidance and its core patterns could be incorporated as reference points in the design. Secondly, the empirical gap identified by Yildirim et al. motivates concrete design commitments for the project. The framework should deliberately foreground early problem framing and success metrics, provide designs that teams can adopt to translate general guidance into local practice. In short, this research treats external guidelines as starting points to be extended, and embedded into the workflows of a finance organisation.

Taken together, these examples demonstrate that while there is growing attention on improving human-AI interactions from both academic and industry perspectives, current approaches often lack the specificity required for domains. My project seeks to address this gap by investigating how generative AI tools can be integrated into internal workflows within a financial organization, and by developing a tailored interaction framework grounded in real employee needs, constraints, and usage contexts. Through a combination of empirical research and design synthesis, this work contributes a domain-specific perspective to the growing field of AI interaction design, bridging the space between general UX recommendations and the practical demands of professional, high-stakes environments. It contributes not only to the academic discourse on AI interaction design but also to the evolving design toolkit for enterprise AI integration.

2.3 Design system methodologies

To ground the framework in established practice, this project draws on contemporary design-system thinking and a modular component methodology. A design system is an operational artifact that codifies appearance, interaction, patterns, and governance to manage design at scale. Practitioners typically distinguish three interdependent parts of a mature system: a style guide, a component library, and a pattern or template layer. These elements together help teams produce consistent interfaces and accelerate cross-team reuse (Nielsen Norman Group, n.d.).

There are two primary references that informed this project's design approach: Nielsen Norman Group's Design Systems 101 (Nielsen Norman Group, n.d.) and Brad Frost's Atomic Design (Frost, 2016).

Design System 101

According to Nielsen Norman Group, a comprehensive design system comprises a style guide, a component library, and a pattern library.

- **Style guide / foundations**
  Contains visual tokens (color, typography, spacing), voice/tone guidance, accessibility rules, and high-level design principles that act as an evaluative lens for design decisions. As the "north star," this layer ensures coherence across disparate products and teams.
- **Component library**
  A documented catalogue of atomic UI elements (buttons, fields, chips), mid-level assemblies (forms, cards), and ready-to-use organisms (navigation bars, tool panels). Each component entry should include purpose, states, accessibility notes, and implementation guidance to enable both designers and engineers to reuse components reliably.
- **Pattern library** (templates & workflows)
  Patterns capture common layout-level solutions—how components combine to support recurring tasks or workflows (e.g., review flow, approval panel, or AI-assisted drafting). Pattern documentation reduces ad-hoc design decisions and makes it easier to reason about interaction behavior at the workflow level.

In practice, design systems also require governance mechanisms, versioning, and tooling integrations—such as design tokens, automated documentation, and code pipelines—to ensure system consistency, facilitate team adoption, and reduce drift between design and implementation.

Atomic Design

Brad Frost's Atomic Design methodology complements this approach by providing a hierarchical structure—atoms, molecules, organisms, templates, and pages—that supports both bottom-up and top-down thinking in interface construction (Frost, 2016). This hierarchy makes it easier to identify reusable interaction elements at varying levels of granularity, allowing components to scale across different contexts while maintaining predictable behaviour. For AI interface design, an atomic approach is particularly useful for structuring modular, adaptable components, such as input fields, response cards, and workflow templates, and for documenting their states, behaviours, and constraints. By pairing atomic decomposition with workflow-level patterns, designers can ensure consistency, traceability, and user-centered outcomes, while also embedding domain-specific considerations such as transparency, reversibility, and trust in AI outputs (Frost, 2016).

Together, these methodologies provide the theoretical and practical foundation for structuring the AI interaction framework in this project. They inform the creation of a component library, the definition of reusable interaction patterns, and the implementation of design governance practices that enable scalable, consistent, and transparent AI integration in financial workflows.

The concrete application of these principles is detailed in Section 4.2 Design Process, describing how atomic components and pattern libraries are instantiated and adapted to the project design.

3. Explorative study

As the theoretical and industry frameworks have established the foundational principles and patterns for human–AI interaction, this chapter presents the core of this project's empirical work: an explorative study of AI usage within Van Lanschot Kempen's internal workflows. Drawing on semi-structured interviews with 11 practitioners across three roles (private bankers, relationship managers, and investment advisors), the study delves deeply into three dimensions of their AI experience: their current workflows and pain points, their usage patterns and attitudes toward existing AI tools, and their detailed interaction preferences across key scenarios.

The outcomes of this explorative study largely align with established HCI theory and industry practice, reaffirming the need for transparency, modularity, and user agency. Crucially, they also uncover domain-specific insights, such as the critical need for reversibility in AI actions, the importance of minimal-step interactions, and the nuanced trust dynamics around generative features. By systematically mapping these real-world observations into structured findings, the study lays a solid evidence-based foundation for the component library and interaction framework that follow. In doing so, it ensures that every element of the design system directly addresses the lived realities, expectations, and constraints of VLK's professional users.

3.1 Research questions

To better understand how AI can meaningfully support employees in the financial sector, this explorative study investigates the current workflows, needs, and attitudes of professionals using internal AI-enabled tools.

The focus of this work is *ClientCenter*, a core internal application used daily by bankers, relationship managers, and investment advisors at Van Lanschot Kempen. The platform supports client communication, meeting planning, and day-to-day task management, and already incorporates several generative-AI features: meeting-note summarization, voice-based documentation, and contextual suggestions (e.g., reminders for client birthdays or deadlines). *ClientCenter* is a custom tool developed in-house by Van Lanschot Kempen to support the firm's employees' workflows. While other financial institutions may have similar banker-facing platforms, those systems are independently developed and are not the same as *ClientCenter*. A representative screenshot of the *ClientCenter* interface is shown in Figure 5.

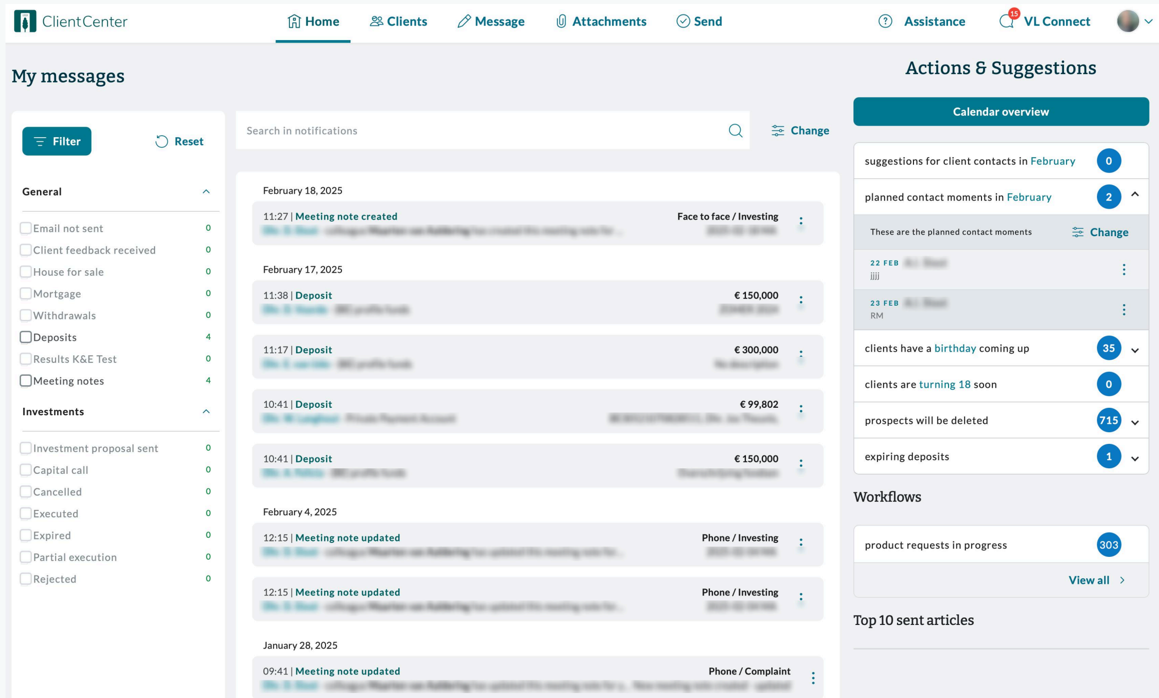


Figure 5: ClientCenter dashboard screenshot

The goal of this study is to gain insight into how financial professionals interact with these existing AI functions, how these tools fit (or fail to fit) into their daily work, and how they imagine the role of AI expanding in the future. With this in mind, the research is guided by the following questions:

a. What are the typical workflows and key scenarios in which financial professionals use ClientCenter?

This question aims to map out the everyday tasks, routines, and pain points of users in order to identify opportunities for AI to augment or streamline work.

b. How are current AI features within ClientCenter perceived and experienced by users?

This includes understanding how users interact with tools like AI-generated summaries or action suggestions, and whether these features are seen as helpful, disruptive, or neutral.

c. What are users' attitudes toward integrating AI into their professional tools?

This question explores not only openness and trust, but also concerns, boundaries, or conditions for AI acceptance in a high-stakes, relationship-driven domain.

d. What interaction styles do users prefer when engaging with AI, and why?

The study seeks to uncover user expectations for control, transparency, timing, and tone in AI communication, as well as preferences for passive vs. proactive behavior.

Through semi-structured interviews, the study aims to capture both present experiences and future expectations. These insights will inform the design of a domain-specific AI interaction framework that aligns with financial professionals' needs and values.

3.2 Research set-up

3.2.1 Research method

To investigate the workflows, attitudes, and interaction preferences of financial professionals, this study employed semi-structured interviews as the primary method. The interviews were conducted via *Microsoft Teams* in the form of online video calls, with each session lasting around 45–60 minutes.

The structure of the interviews broadly followed the four themes defined in the research questions: current workflow, AI usage and attitudes, and interaction preferences. A script was made and used to guide the conversations around the objectives of the research questions outlined in Section 3.1. This approach allowed for open dialogue while ensuring that central themes were consistently addressed across participants. In some parts of the interviews, visual materials were shown to participants to illustrate scenarios and interaction possibilities, which facilitated richer discussion. The full interview script and visuals shown during the interview is included in the Appendix.

To support more natural reflection and future-oriented thinking, the question flow was adapted based on the path of expression model from *Convivial Toolbox* (Sanders & Stappers, 2012) shown in Figure 6. The interview sequence was organized to move from present experiences to past reflections, and finally to speculative visions of the future. This approach was chosen to help participants build on their lived experience before articulating how they imagine AI might support them going forward.

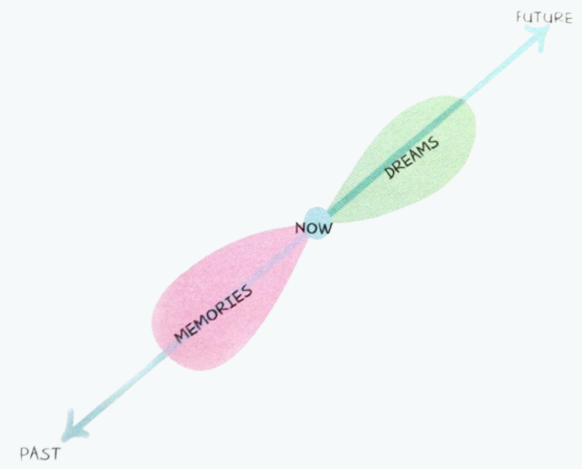


Figure 6: Path of Expression Model from Convivial Toolbox

To explore participants' preferences regarding AI interactions more concretely, three distinct usage scenarios were introduced during the interviews:

Text writing	Looking up information	Advice from AI
• Rewrite article to send to client • Improve email writing • Convert draft to readable text	• Summary and create overview • Find document • AI meeting summary	• Ask about administration tasks • Ask about professional knowledge • Suggestions about tasks

Table 1: distinct AI usage scenarios

The choice of these three scenarios was grounded in two considerations.

Firstly, insights from prior research within the organization indicated that text writing and looking up information are among the most common tasks performed in ClientCenter. Advice from AI was selected as a evolution of the second scenario, representing a case where AI does not merely retrieve information, but synthesizes and interprets it for decision-making.

Secondly, the three scenarios were purposefully chosen to align with three distinct interaction quadrants derived from the *Shape of AI's Modes of Prompt framework* (AI UX Patterns | Filters | ShapeofAI.Com, n.d.). As shown in Figure 7, this framework categorizes interaction scenarios along two axes: the user's clarity about the input (from unknown to known) and the user's clarity about the desired output (from goal unknown to known). By covering three separate regions of this interaction space, the study was better positioned to identify patterns in user preferences across a range of cognitive modes.

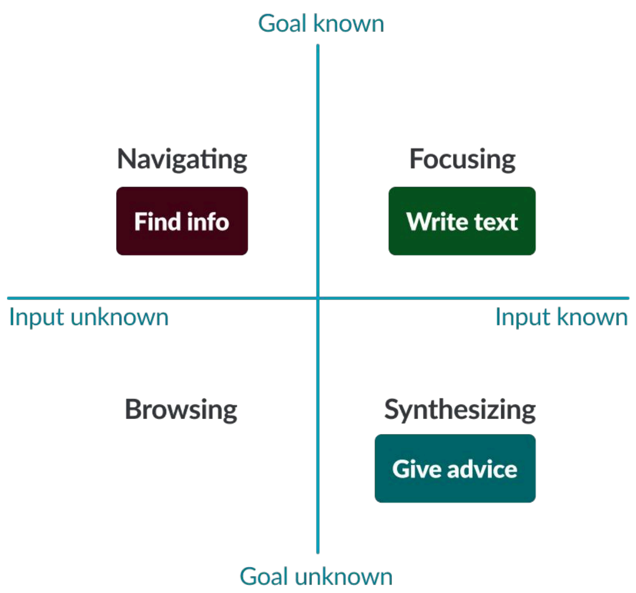


Figure 7: 3 scenarios fall under the modes of prompt.

In each scenario, participants were presented with three prototype interaction concepts, each representing a different interaction way. These visual examples link abstract conversations with tangible possibilities and prompt richer feedback and comparison.

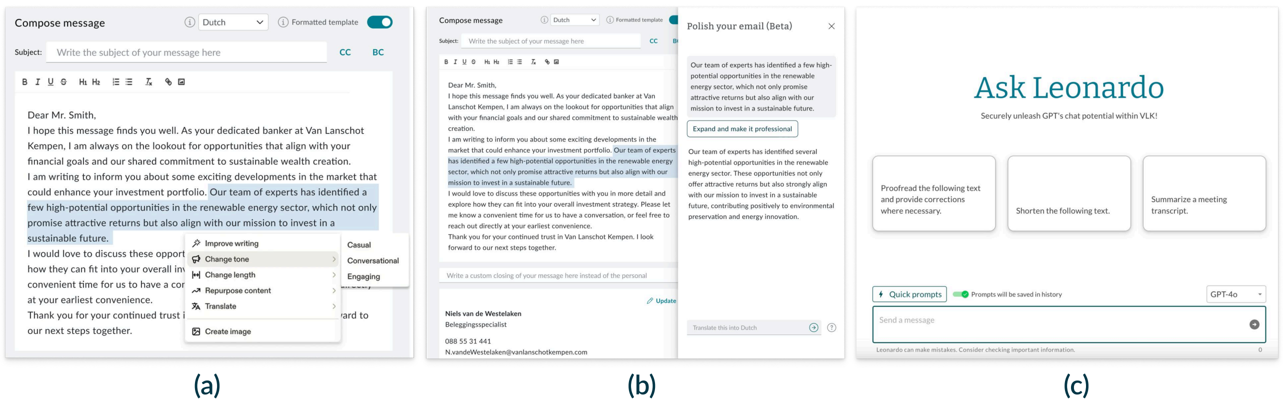


Figure 8: text writing scenario visual

For instance, as shown in Figure 8, in the text-writing scenario, participants explored three options: (a) an AI tool integrated into the text editor via a drop-down menu; (b) an AI tool displayed in a side panel; and (c) a standalone AI tool interface, such as a dedicated company GPT, *Leonardo*. To ensure unbiased responses, these options were presented in a randomized order for each participant, mitigating any influence from presentation sequence. These interaction concepts drew inspiration from established, industry-leading design patterns observed in cutting-edge AI tools, grounding the prototypes in practical and proven frameworks.

Together, these methodological choices allowed for both depth and flexibility, enabling participants to share rich insights grounded in their real-world workflows while also reflecting on speculative futures and evaluating concrete interaction models.

3.2.2 Participants demographic

SAMPLING

Quota Sampling

1. Role
2. Experience
3. AI acceptance
4. ClientCenter usage

Reached out through email.

PARTICIPANTS

11 Participants

5 Private Banker,
4 Relationship Manager,
2 Investment Advisor

45-60 minutes per participant
Interviews are conducted between **27 Mar - 17 Apr 2025**

ANALYSIS SUFFICIENCY

Participants covered:

1. 3 different roles
2. Senior, junior
3. Low, High AI acceptance
4. Medium, High ClientCenter usage

To ensure diversity and relevance among participants, I adopted quota sampling strategy (Rukmana, 2014), guided by four key criteria: user role, years of experience, frequency of ClientCenter usage, and familiarity with AI tools. I first contacted the product owner of ClientCenter to obtain a name list of employees who actively use ClientCenter and internal AI features. From this list, I selected individuals across three main user roles (private bankers, relationship managers, and investment advisors), ensuring representation across different levels of AI adoption (low, medium, high) and ClientCenter usage (medium, high). Each candidate was approached individually via email, where I shared the purpose, scope, and setup of the study (see Appendix for the email template).

In total, 34 invitations were sent out, and 11 participants accepted and completed the interview. The final participant group consisted of 5 private bankers, 4 relationship managers, and 2 investment advisors, covering a range of seniority levels and tool usage patterns. This sample provided a well-balanced perspective that aligns with the study's objectives, as it captures the viewpoints of three user roles with complementary responsibilities, both junior and senior professionals, low to high AI tool acceptance, and medium to high frequency of ClientCenter usage.

Invitation emails were sent starting March 20th, followed by two pilot interviews conducted with in-house designers to refine the interview structure. Formal interviews took place between March 27th and April 17th, 2025.

3.2.3 Analysis process

To analyze the interview data, I followed the *Thematic Analysis* (Clarke & Braun, 2014) method, which provides a flexible yet structured approach to identifying meaningful patterns across qualitative data.

After completing the interviews, I began by transcribing the full recordings of all eleven sessions in *Loop*, a company used documentation tool. This allowed me to become familiar with the data and ensuring accuracy in later analysis.

Once transcribed, I organized the material by cleaning up the language, segmenting the responses according to topic areas, and annotating them for clarity.

The next phase involved coding the data. To enhance objectivity and depth, I invited two designer colleagues to independently code the transcripts alongside me. After coding was complete, we reviewed the material collaboratively and began synthesizing recurring themes, looking for patterns across roles, tools, and attitudes.

To validate our findings and mitigate individual bias, I organized a workshop on April 23rd with the same two colleagues. During the session, we cross-examined each other's interpretations, discussed ambiguities, and refined the thematic groupings together. This collaborative step helped broaden the analytical perspective and added confidence to the insights we later carried forward.

During the analysis, I experimented with different tools. Initially, I attempted to use *ATLAS.ti*, a professional qualitative analysis tool. However, I encountered several limitations. The process was time-consuming using it, and the tool lacked flexibility in merging related quotes across multiple parts of a transcript. Meanwhile, my colleagues did not have access to *ATLAS.ti*, which restricted collaborative analysis. Given these constraints, I opted to conduct the analysis using *Figma*, a tool I am proficient in and that allowed for flexible structuring and visual collaboration see in Figure 9 (detailed analysis process in Appendix). Despite the change in tooling, I strictly followed the thematic analysis methodology throughout the process.

The analysis began on April 17th and concluded by May 2nd, when I presented the preliminary research findings during the midterm evaluation.

Figure 9: analysis process

3.3 Research insights

Based on the collected data and subsequent analysis, the findings from this research are organized into three key perspectives that together offer a comprehensive understanding of how financial professionals interact with AI in their everyday work.

Firstly, it maps users' current workflows, delineating how tasks are organized, which tools are employed, and where friction arises. This analysis confirms anticipated pain points, also uncovers less obvious patterns of cross-role coordination and ad-hoc tool adaptations that had not emerged in prior academic literature.

Secondly, it examines AI usage and attitudes, exploring which AI features are embraced, which are resisted, and why. While we expected participants to appreciate time-saving tools for drafting text and summarizing data, we also identified a nuanced spectrum of trust levels and adoption drivers. Notably, users demonstrated a sophisticated understanding of data quality's impact on AI outputs and expressed a strong preference for features that offer clear provenance and reversible actions, insights that extend existing HCI findings into the domain of financial operations.

Thirdly, it probes interaction preferences under three representative scenarios. Beyond validating recognized interaction modes, we observed emerging desires for richer interactions that go beyond current tool capabilities. These findings suggest untapped opportunities for designing workflow-embedded AI features that balance autonomy with user control.

Together, these findings form the foundation for identifying both design principles and practical requirements for responsible, effective AI integration in internal banking tools.

3.3.1 Current workflow

The financial professionals interviewed in this study operate in closely connected, role-specific teams that serve approximately 200 clients group per member. As illustrated in Figure 10, each client is typically supported by a team of three to four people: a private banker, an investment advisors and one or two relationship manager. While their responsibilities are distinct, their work is highly interdependent, and communication between roles is frequent.

Private bankers serve as the primary client-facing contact and the central coordinator of the team's relationship with the client. On a day-to-day basis they schedule and run client meetings, synthesize client goals and life events into financial needs, and translate those needs into actionable plans that may involve banking products, wealth planning and investment strategies. Private bankers are responsible for identifying cross-sell opportunities and expanding the client relationship, which requires both commercial judgement and deep knowledge of the client's preferences and history. In practice they rely heavily on ClientCenter and personal notes to track meeting outcomes, upcoming tasks and client preferences.

Investment advisors bring domain-specific expertise to the team and focus on technical, market-facing tasks. Their role ranges from providing targeted recommendations, such as buy/sell actions or alternative investment selections, to producing in-depth analyses (risk assessments, modeling scenarios) when client needs demand specialist input. Investment advisors often take proactive outreach during market volatility, contacting clients with time-sensitive commentary or execution advice. They use both internal research systems and shared team resources in ClientCenter to gather data; when interacting with AI features they are primarily concerned about the factual soundness, source provenance, and reproducibility of any suggested actions.

Relationship managers play a coordination and operational support role that keeps the team functioning smoothly. Their activities include preparing materials and briefs for PB-led meetings, maintaining client communications for less formal touch-points (for example birthday notes or appointment reminders), and routing incoming client requests to the right specialist. They often handle the administrative backbone of client servicing: updating records, following up on documentation requests, and ensuring compliance-related items are filed correctly.

Together, these roles form a collaborative ecosystem where responsibility is distributed but tightly aligned around client service and financial strategy.

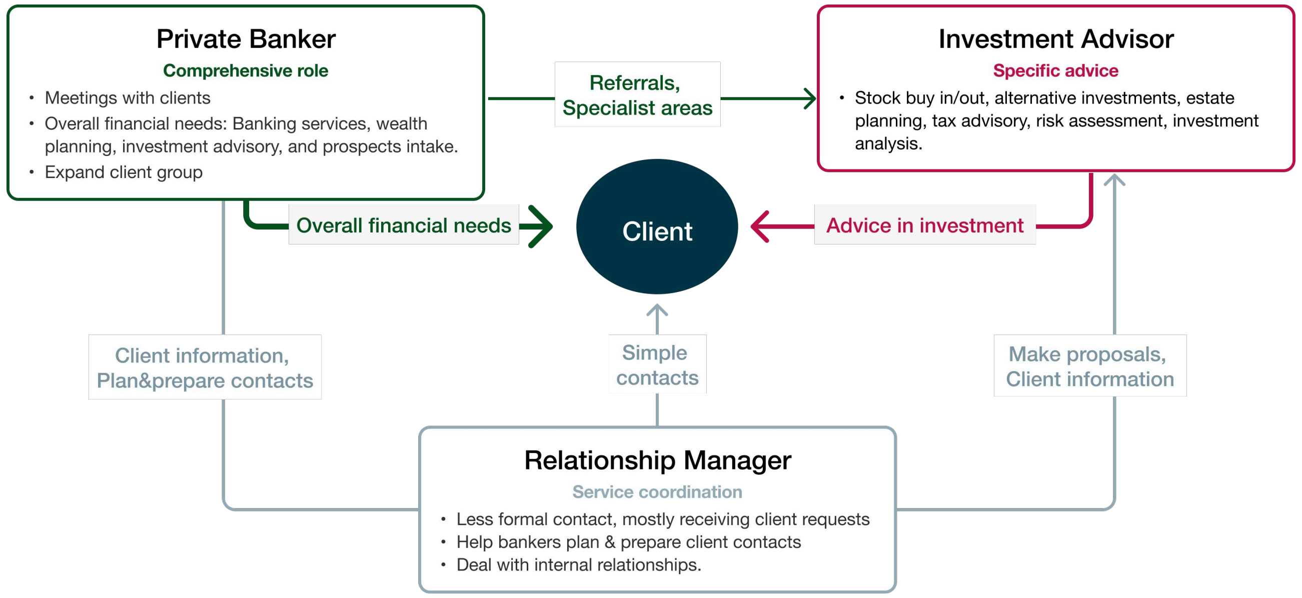


Figure 10: private banker, invest advisor and relationship manager's role responsibility

Despite the differences in their day-to-day tasks, all three roles share a similar routine at the start of the workday: checking messages via Outlook or ClientCenter, reviewing their agenda (typically managed across Outlook, Excel, or third-party tools), and reading financial news online. Interestingly, some tech-savvy employees have already begun experimenting with AI tools to streamline the news-reading process, signaling openness to AI support in information-heavy tasks.

However, when asked about the most time-consuming aspects of their job, regardless of role, all participants consistently pointed to compliance and administrative work. Tasks such as Customer Due Diligence (CDD) and Know Your Customer (KYC) procedures, which involve verifying client identities and assessing risk profiles, were seen as especially burdensome. Similarly, meeting note-taking was cited as repetitive and low-value. Both task types are characterized by high volumes of manual data entry across systems, back-and-forth clarifications with clients, a high sensitivity to errors, and little perceived added value. These shared pain points highlight clear opportunities where AI could step in, not to replace expertise, but to reduce friction and free up professionals to focus on strategic and relational work.

3.3.2 AI usage & attitude

Overall, participants demonstrated a positive mindset toward using AI in their daily workflows. Most expressed a willingness to learn and actively engage with AI tools, especially when the tools could help them save time or improve the quality of their output. Many users agreed that AI not only supports efficiency, particularly in repetitive or low-value tasks, but also helps them improve communication with clients by offering refined phrasing or alternative ways of expressing ideas. Importantly, participants acknowledged that high-quality AI output depends on well-structured data, and several mentioned that they see it as their responsibility to

"keep the source clean" for better AI performance. Even those unfamiliar with AI tools showed interest in learning, though they expressed a strong need for more structured guidance and education from the company.

Despite this openness, several technical and usability challenges were noted. Participants frequently mentioned a desire for broader, more relevant, and up-to-date data sources. Many also struggled with finding or accessing AI tools, expressing a wish for AI to be more seamlessly integrated across internal systems. Another recurring challenge was prompt creation, users often lacked confidence in how to formulate effective prompts and requested features like prompt history, reusable templates, and example libraries to improve their usage.

Interestingly, participants showed a strategic awareness in how they selected and used different AI tools. As shown in Figure 11, they were able to articulate the trade-offs between available tools and tended to adapt their tool choice to the specific scenario or task. This level of discernment reflects not only a growing literacy around AI capabilities but also a desire for tools that are tailored to their domain-specific needs.

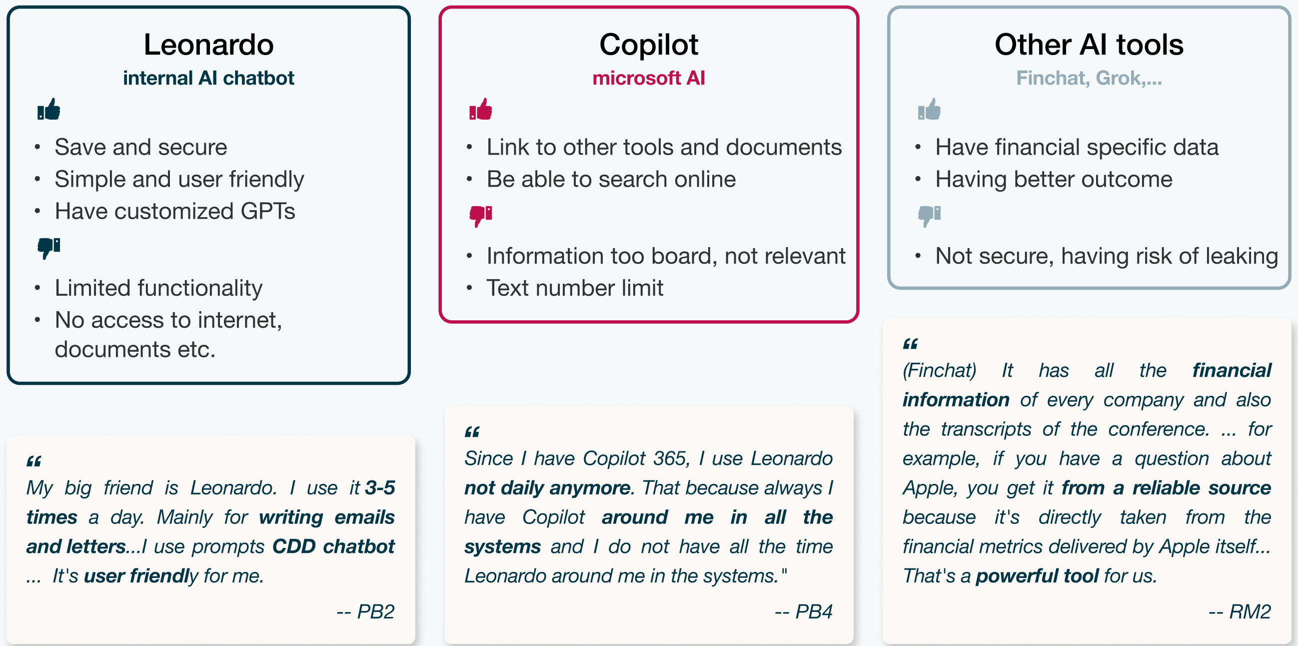


Figure 11: participants' reasons of using different AI tools

Trust in existing AI features varied by tool and scenario. Participants tended to trust internal, company-managed tools more than external services, partly because internal tools often have good sources and make provenance visible. Because the current AI capabilities in ClientCenter are not positioned to replace human decision-making but rather to assist with drafting and summarization, most users reported limited anxiety about AI making consequential mistakes at this stage. More nuanced, scenario-specific patterns of trust are described in detail in the scenario analyses that follow.

When asked what kinds of AI features they would like to see in the future, participants commonly mentioned functionalities that align with their current pain points and workflows. For example, many expressed interest in features that could automate document generation, summarize client histories, or support internal process navigation. More respondents mentioned the items listed in Table 2, ordered by frequency of mention.

	Feature name	Description
1	Searching	• Filter based on client profile, list most relevant clients to send messages, plan calls, invite to events, etc. • Answer clients' specific questions
2	Suggesting	• Prioritize signals, list most important tasks • Suggest on commercial opportunities • Actively help remember stuff (birthday etc.)
3	Planning	• Fit contact moment in agenda • Plan in advance, give client more time • Plan internal meetings with specialists
4	Writing	• Adapt to personal writing habit and tone • Based on the meeting record, generate meeting note • Having pre-set email template for common things
5	Summarizing	• Summarize articles and reports • Summarize client information for preparing meeting
6	Making	• Making company style slides from report

Table 2: expected future AI use case

Beyond these overall desires, we compared how participants viewed AI across three core scenarios(text writing, information lookup, and advice generation), in terms of their familiarity, trust, and desire to use AI:

Use AI to help with text writing

Most participants shared that they already use AI for text-writing tasks. This is largely because the technology in this area is mature, widely adopted, and has proven to be reliable. The outcomes are predictable and carry low risk, allowing users to retain editorial control while benefiting from AI's language generation capabilities. As a result, text writing stands out as the most familiar AI-supported scenario among users and earns a relatively high level of trust.

Use AI to look up information

While participants familiarity with this kind of AI support is somewhat lower than with text writing, many users still showed moderate to high trust, especially when the AI output could be verified (for example, by checking cited sources, cross-referencing internal records, or consulting a colleague/expert) or when the task was relatively low stakes. Moreover, this scenario was rated as the one with the highest desire to use AI, reflecting the practical benefits users anticipate in reducing time spent searching, compiling, and summarizing content.

Get advice from AI

Compared to the first two, this scenario elicited much more skepticism. Participants emphasized that while such advice could be helpful in theory, they currently do not trust AI enough to rely on it in high-stakes or complex decision-making contexts. The skepticism stems from the novelty of the capability and a lack of successful use cases: participants have little evidence that AI advice will be reliably correct. They also pointed out a practical concern: if an AI recommendation fails, human user ultimately needs to bear responsibility for the consequence. They worried that incorrect guidance, especially around compliance or client-facing actions, could have serious consequences. As such, trust in this scenario is low, and while some users acknowledged its potential, the desire to adopt AI in this capacity remains limited.

Across these three scenarios, clear patterns emerged in terms of familiarity, trust, and desire to adopt. Participants saw text writing as the most familiar and trustworthy, followed by information lookup. Advice-giving was seen as the least trustworthy and desirable. Through detailed analysis of their responses, three main factors were found to influence their attitudes:

- **Users tend not to trust new AI functions by default**, where these broadly refer to any unfamiliar tools/feature introduced into their workflow, rather than a single specific feature. In a conservative, high-stakes environment such as finance, unfamiliarity itself creates caution, and trust must therefore be earned through repeated exposure and consistently reliable outcomes.
- **Autonomy and control are essential to building trust.** Participants expressed higher confidence in AI when they could easily modify, undo, or ignore its suggestions.
- **The ability to judge output quality plays a crucial role**. Users feel more comfortable when they can quickly evaluate whether the AI's result is right or wrong, especially in domains where incorrect information could cause reputational or regulatory harm.

Interestingly, while text writing ranked highest in terms of trust and familiarity, information lookup was rated most desirable. Drawing on the *Shape of AI framework*'s prompt typology (Figure 7), information lookup tasks typically fall into the quadrant where the user has a clear goal but lacks the method to reach it ("goal known, input unknown"), a space where AI support is most valued. In contrast, advice-giving functions, while potentially powerful, remain least desirable due to risk aversion, trust barriers, and the requirement to change long-standing work habits.

Several barriers to desire were identified across scenarios. These include: the need to significantly adjust existing workflows, the pressure of AI overstepping human judgment, and the fear that functional errors could have serious consequences.

Interview responses also suggested that these perceptions are not uniform across experience levels. Newer or more tech-savvy employees tended to express greater openness to trying GenAI tools, likely because they are less tied to traditional workflows and more comfortable experimenting with novel tools. More senior staff were generally more cautious and comfortable with their existing workflow: they demanded clearer provenance and undo controls, though several noted they would adopt tools once reliability had been demonstrated repeatedly.

Finally, each scenario is also accompanied by its scenario-specific pain points, which are summarized in Table 3. These points offer further insight into what constrains adoption and where targeted design improvements could unlock new value.

Scenario	Pain points	Quote
Text writing	• Hard to make correct prompt • Combine previous text incorrectly • Doesn't adapt to my writing style	*"If you just say 'write this', it doesn't work that well. It often **takes longer** to **write a good prompt** than the mail itself." -- PB5*
Looking up information	• Source too general • Cannot extract information from pdf • Don't have wide historical data source	*"If I have a document of 40-50 pages I will let AI make a summary of it... saves me a lot of time... Then there's only one question -- **is the summary good enough**?" -- IA1*
Advice from AI	• Only for advise, can't do actual tasks • Question the accuracy, afraid of missing important things	*"But for really practical daily work like administration, I can't make it do anything... it's more like the **communication stuff** ... I just ask for improvement or how can I do this... But **not really actual tasks** to do." -- RM4*

Table 3: pain points under each scenarios

3.3.3 Preferred interaction

Building upon the three scenarios introduced above, this section continues the analysis of the semi-structured interviews described in section 3.2.1. In addition to discussing their attitudes and trust levels, participants were also asked to respond to three alternative interaction prototypes per scenario. Their reflections on these prototypes provide further insight into what makes AI interactions feel usable, trustworthy, and aligned with their mental models. The feedback reveals not only scenario-specific preferences but also deeper values regarding control, clarity, and efficiency.

Use AI to write email

Expectations were highest in this scenario. Participants were familiar with AI writing tools and saw strong potential to improve the efficiency and consistency of email communication. They preferred interaction models that were template-driven, allowing them to quickly select a purpose and have the AI generate an initial draft accordingly. Many also expressed a desire for features that could match their writing tone, suggesting that AI should adapt to their previous messages or allow fine-tuning for more personalized results.

This scenario was perceived as particularly useful for standardized communication, such as sending articles to clients, announcing new services, or summarizing event invitations. Some participants also mentioned they would benefit from prompt libraries, where frequently used requests could be saved and reused across the team.

Despite the enthusiasm, several concerns were raised regarding loss of voice and content fidelity. Users worry about AI-generated text that misaligned with their personal style. They also feared accidentally overwriting their original draft, especially if there was no way to revert changes. Another repeated friction point was the effort required to write a good prompt. Even when the tool offered intelligent suggestions, users found themselves spending time rephrasing requests, which offset the time-saving benefits they expected.

Use AI to get client information

In this scenario, the preferred interaction pattern was a two-step structure: starting with a concise overview and then offering the option to explore deeper information through follow-up queries. This interaction style helped users gain situational awareness quickly without being overwhelmed by details.

This preference aligns with how users typically prepare for client interactions, especially when meeting newly transferred clients, identifying connections to expand their client network, or checking for updates from client-related news.

Participants appreciated when the AI could surface useful context efficiently, such as summarizing a client's recent portfolio changes or linking relevant external articles.

However, several interaction-specific concerns emerged. A major one was the difficulty of writing effective prompts. While users were open to querying the AI for more information, many were unsure how to phrase their questions to get useful results. Some feared AI hallucinations or data fabrication, particularly if the AI didn't clearly indicate its source or confidence level. A few participants also mentioned that such features felt less useful when applied to clients they already knew well, saying, "*I've been working with this client for years, already know everything.*"

In addition, some raised concerns about privacy and data security when external tools such as Microsoft Copilot are involved. They worried that sensitive client information might be processed outside the firm's controlled environment, underscoring the need for strong guarantees around data handling, storage, and regulatory compliance.

Get advice from AI

In this scenario, the option that presents a task list was generally seen as the most understandable and actionable. Users appreciated having a clear structure where they could see suggestions laid out explicitly and evaluate them at a glance. However, while the visual clarity of the task list format was praised, participants emphasized that they do not want AI to fully plan their day for them. Instead, they preferred to be in charge of their own planning, using AI on demand to answer specific questions or speed up prioritization. This distinction reflects an important balance between proactive vs. reactive interaction. While AI-initiated suggestions can be helpful, users want the freedom to accept, ignore, or modify them.

The scenarios where participants found this kind of AI advice most useful included:

- Planning tasks in advance, especially for heavy client days
- Filtering clients using multiple conditions (e.g., AUM, recent meetings) to identify who to contact
- Prioritizing urgent or overlooked tasks that might otherwise slip through

On the other side, several concerns were raised. Users feared missing important tasks if the AI failed to surface something crucial. They also expressed hesitation about losing control, particularly if AI suggestions appeared too assertive or acted automatically. Another repeated concern was the inclusion of irrelevant suggestions, especially when the AI failed to account for their specific role or responsibilities.

Across all three scenarios, several cross-cutting interaction preferences emerged, which are summarized in Table 4. These include a preference for minimal effort interactions that follow a clear "display → select → input" flow, a strong desire for AI functions to be contextually embedded in their primary workspace (rather than siloed in separate tools), and a need for inline guidance, such as examples, previews, or smart defaults, especially when engaging with open-ended input fields.

Less actions	Prefer displaying information directly rather than navigating through multiple steps.	" *If I open the screen, I see it immediately and I don't have to type.* -- RM4
Clear & Relevant	Displays only relevant information and clearly indicates the source.	" *I'd rather have 20 valuable things to act upon than to have a suggestion for every client, so less is more.* -- PB1
Easily undone	Humans retain control over AI, with the ability to quickly delete or modify AI actions.	" *Before changes are made ... I want to get back to the original message... I'm working in my original message and I'm scared of losing it.* -- PB5
Fast reaction	Impatient with loading; prefer visible progress updates.	" *The page is very slow, which causes me not to use it as often as I could.* -- RM4

Table 4: common preference in AI interaction among all scenarios

To summarize, the explorative study provides a layered view of how financial professionals engage with AI tools in their daily work. The analysis began by mapping workflows and identifying where frictions occur, then examined how AI is currently used and perceived, and finally introduced scenario-based probes to elicit preferences and spark further ideas. These points form a coherent inquiry, moving from present realities toward speculative possibilities.

These findings form the foundation for the design phase of the project, where insights will be translated into an interaction framework tailored to the financial context. The next chapter outlines this translation process in detail.

4. Framework design

Building on findings gathered from the related work and explorative study, this chapter shifts the focus to translating these insights into practical design guidance for AI feature development, presenting the framework for designing scalable AI interactions within Van Lanschot Kempen's internal workflows.

The framework is structured to support designers and product owners in creating AI-powered features that are transparent, modular, and aligned with user needs. It contains three interrelated dimensions: high-level design principles that codify user expectations and organizational values; a component library of reusable interface elements ranging from atomic inputs to full-page AI workspaces; and an interactive selection wizard that translates principles and components into concrete design decisions. By grounding each element in observed behaviors, workflow realities, and user preferences, the framework ensures that AI interactions remain actionable, trustworthy, and consistent across different tools and contexts.

The following sections detail the design goal (Section 4.1), the design process (Section 4.2), and the resulting design outcomes (Section 4.3), providing a comprehensive view of how the framework supports effective AI integration across VLK's internal systems.

4.1 Design goal

Building on the foundation laid through user research, the next stage of this project focuses on translating insights into design. The aim is to move from understanding current workflows, attitudes, and interaction preferences toward shaping future experiences with generative AI. This begins with establishing a clear and strategic design goal. The goal is to:

> Design a scalable framework for integrating generative AI into the workflows of employees at VLK, promoting consistent, transparent, and effective AI interactions.

This framework is not an end-user-facing product, but rather a design infrastructure intended to guide the creation of AI-powered features within financial workflows at Van Lanschot Kempen. The primary users of this framework are designers and some product owners who are responsible for designing and building AI features across various digital touch-points at Van Lanschot Kempen. While they are the ones who will directly engage with the framework to plan, prototype, and implement AI-driven interactions, the framework's ultimate impact extends to all employees who rely on these tools in their daily work.

To ensure this framework can be adopted and scaled across teams and contexts, it must address four key considerations:

Transparency
Users need to understand what the AI is doing, where the information comes from, and how to assess its accuracy. Clear feedback, source visibility, and output evaluability are essential for building trust.

Collaboration
Based on our research, bankers frequently work in teams, so the framework accounts for interactions where multiple team members may engage with AI collaboratively. This includes offering assistance, suggestions, and augmentation while keeping human users in control, ensuring that AI supports both individual and team workflows effectively.

Consistency
With different teams and departments experimenting with AI independently, there is a growing need for a shared set of interaction principles that create familiar, intuitive experiences across tools.

Scalability
The framework should support modular integration, allowing new AI components to be added over time, and adapted to different tasks and contexts within VLK's operations.

4.2 Design Process

The proposed framework can be understood as a specialized variation of a design system, tailored specifically to support AI interactions in internal financial tools. According to the design system methodologies reviewed in Section 2.3, comprehensive systems benefit from structured component libraries and modular design hierarchies. Drawing on these principles, alongside insights from the explorative study, I structured the design process into three main phases to develop a framework that is both scalable and grounded in real user workflows.

a. Define Design Principles

Using the qualitative insights from interviews and scenario testing, I began by extracting key behavioral and interaction themes that could be translated into design principles. These principles (detailed in Section 4.3.1) serve as high-level guidelines to steer future AI feature development, ensuring alignment with user needs and organizational values such as transparency, flexibility, and trust.

b. Build the Component Library Using Atomic Design

For the system architecture itself, I chose to adopt the Atomic Design method. It offers a clear, methodical structure that aligns well with the layered complexity of AI interactions. Moreover, It supports both bottom-up and top-down thinking, allowing me to start from detailed interface elements and also reason through how they come together into full workflows.

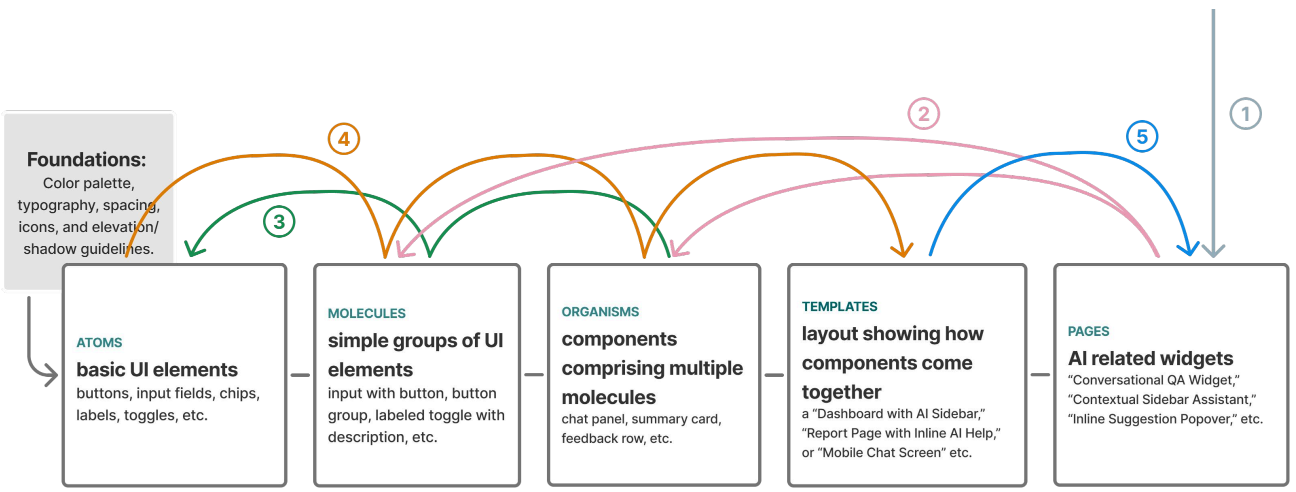


Figure 12: design process based on Atomic Design method

As shown in Figure 12, my design path began with the three user scenarios developed and validated during the research phase. Since participants had already evaluated different prototypes in these contexts, I started by refining and redesigning the page-level flows for each scenario based on their feedback.

In the second step, I deconstructed these flows to extract relevant molecules and organisms, such as prompt modules, AI-generated message cards, feedback widgets, and response tuners. These interaction units were then systematically documented and expanded.

The third step involved identifying any missing but essential interaction components not yet represented in the original scenarios. I supplemented the library accordingly, ensuring completeness and modularity across different potential AI use cases.

In the fourth step, I used these components to construct reusable templates, integrating multiple interaction units into layout patterns that could support entire tasks. To test the system's integrity, I applied these templates to new scenarios not originally part of the research, examining whether the existing components were sufficient to build pages for those use cases.

I intentionally chose this grounded, scenario-first approach instead of building the system from scratch or through abstract ideation. By starting from real user data and interface behavior, the resulting system remains more realistic, actionable, and attuned to actual team workflows, avoiding overly conceptual or theoretical models that may not fit real-world constraints. During this process, I also kept a record of my design considerations and reasoning, which would later support the formulation of a design guide.

c. Develop the Interaction Guide

In the final phase, I returned to the notes made during component development to identify what considerations and decision-making logic repeatedly influenced the design of effective AI interactions. These include, for example, when to offer proactive AI suggestions versus waiting for user input, how to express uncertainty in AI-generated content, and what forms of feedback help users feel in control. I then organized these into a stepwise framework (seen Figure 13) that outlines what factors must be considered, and in what order, when introducing new AI functionality into internal tools.

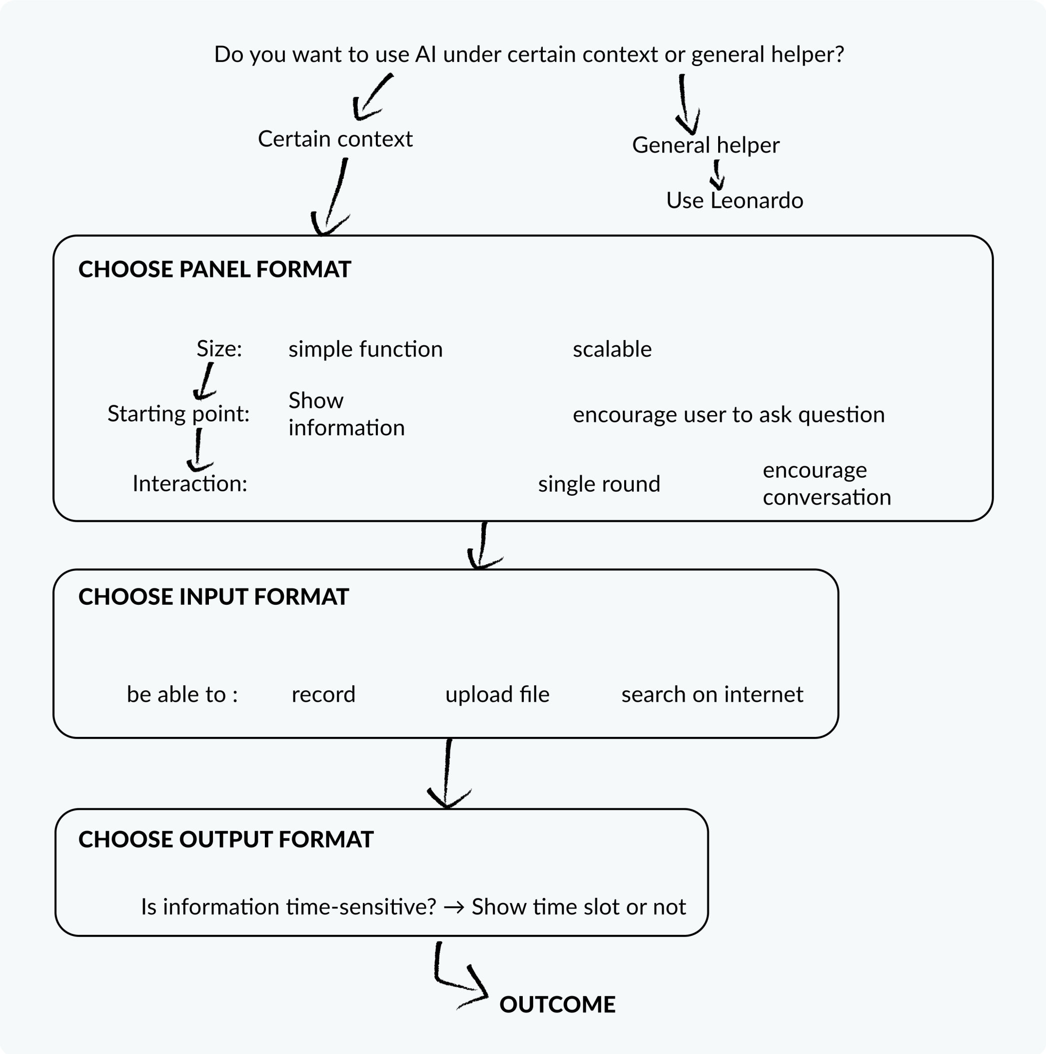


Figure 13: stepwise framework draft

At the same time, I began considering the format for delivering this framework. Given the need for scalability and usability across teams, I decided to present it in the form of a web-tool wizard, allowing interactive exploration and embedded logic. With algorithmic support, the system can eventually offer personalized recommendations, helping designers and product owners at Van Lanschot Kempen make informed decisions about AI interaction design in different use cases.

In the following sections, I will present the outcomes of each design phase in turn: the principles that guided the work (Section 4.3.1), the resulting interaction components and templates (Section 4.3.2), and the development of the interaction guide (Section 4.3.3).

4.3 Design outcome

Based on the design process and the guiding principles distilled from our research, the outcomes of this project are organized into three interlocking deliverables that together form a cohesive AI interaction framework.

First, the ***Design Principles*** *articulate the core values—human control, transparency, clarity, and actionable feedback—that underpin every component and interaction pattern. These principles serve as the ethical and practical foundation for all future AI feature development at VLK, ensuring that every new tool respects user agency and fosters trust.*

Second, the ***Component Library*** *provides a structured collection of reusable interface elements ranging from atomic inputs and chat bubbles through to full-page AI workspaces, organized according to atomic design. The majority of AI-specific components were designed by author for this project, with a set of foundational primitives adopted unchanged from Van Lanschot Kempen's existing design materials: fonts, buttons, icons, base input fields, tooltips, disclaimers, and modal window margins. This report highlights one representative slice of components as an illustrative example (see Section 4.3.2). The complete set, including variant states and usage guidelines, is available in the appendix.*

Third, the ***AI Feature & Component Selection Wizard*** *translates both the principles and the components into an interactive decision-support tool. By guiding designers and product owners step by step through role definition, interaction style, input/output configuration, and container selection, the wizard ensures that every new AI feature aligns with Van Lanschot Kempen's workflow realities and strategic goals. A demo link is also available in the appendix.*

Together, these three outputs provide a scalable, transparent, and practical system for building consistent AI experiences. By empowering designers and product owners, the framework ensures that wealth management professionals experience AI features that are trustworthy, coherent, and tailored to their daily tasks.

4.3.1 Design Principles

Based on the user research and design synthesis, I formulated a set of guiding principles to inform the design of AI interactions at VLK. These principles aim to ensure that AI features are not only functional, but also trustworthy, usable, and well-integrated into existing workflows. They address key concerns raised by users and provide practical guidance for designing future AI-powered components and experiences.

Human Control	• UI labels remind users that AI is an assistant • Provide obvious stop or undo options. Any AI edit should be reversible. Users should be able to edit AI outputs. • When AI suggests an action (like updating a record), require explicit user confirmation.
Transparency	• Use labels or icons on content AI generated. • Provide short explanations on how suggestions were formed (e.g. "Based on today's sales data, I suggest…"). • Avoid false precision or over-promising, the language should reflect if AI is unsure. • Display source links or citations for factual info to let users verify claims .
Clarity	• Use a clean, familiar UI layout (e.g. speech bubbles, timestamps, assistant avatars). Group related actions in menus or accordions. Use tooltips or expandable info icons to hide complexity (e.g. detailed settings or AI training notes) unless needed. • Prioritize direct information presentation, reducing the need for users to navigate through multiple steps. • All presented information is relevant and clearly sourced, allowing users to easily comprehend and trust the data.
Actionable	• Use prominent buttons or quick-reply chips to guide user flow (e.g. "Apply" or "Explain more") . Offer easy regenerate or edit prompt options when answers are unsatisfactory. • If the AI cannot answer, it should offer alternatives (like linking to help docs or human support). • Embed easy feedback mechanisms (thumbs, star rating, quick comments) right in the UI. Prompt users to rate answers or flag mistakes, and automatically generate a revised response when requested.

4.3.2 Component library

The component library forms the core of the design framework. Based on the scenarios and interface elements identified during research and following the logic of atomic design, I mapped out all AI-related UI components and grouped them into nine key overview pages. Each page introduces a cluster of elements according to their function, scale, and interaction role within the system. The structure of this component library is summarized in Table 5.

Templates	AI panel format
	Inline AI usage
	Floating/Side-panel widget
	Full-page AI workspace
	Start page
Organisms	Prompt input
	Chat bubble
Atoms	AI icon buttons
	Tags

Table 5: component library structure

The majority of AI-specific components were newly designed for this project, based on the research insights and validated scenarios, and then systematized into a coherent library. To preserve brand continuity and reduce adoption friction, some foundational primitives was intentionally adopted unchanged from company's existing design system: *fonts, buttons, most iconography, base input fields, tooltips, legal disclaimers, and modal window margins.*

As an extension to the company's existing design library, not every atomic or organism-level element was documented as a standalone page. This component sets extends and complements the current design library by focusing on AI-specific interaction elements. In particular, to reduce cognitive load and encourage reusability, I merged smaller decisions into the higher level and implemented them as component properties. Designers can toggle these options easily via sidebar controls, making the system more modular and efficient to use.

Given the scale of the system, this report highlights key structures and representative examples. The complete, detailed component documentation is available in the appendix.

a. Entry Layer: AI Panel Format

A central entry point of this system is the AI Panel Format, which outlines four primary ways generative AI interfaces can appear across different contexts:

- Embedded inline usage
- Floating chat widgets
- Docked side panels
- Full-page AI workspaces

Each format is defined with clear usage guidelines, scenario suitability, and best practices regarding placement and interaction flow.

For example, the Full-page AI workspace replaces the entire interface with a dedicated view focused on AI interaction, accessible via main navigation or deep links. This format supports immersive engagement with AI tools while keeping the user anchored in the task context. To preserve this contextually and differentiate it from generic chatbot experiences, a notable design strategy is to use gridded layouts that create temporary workspaces. These surfaces display relevant inputs and visual outputs (e.g., AI-generated summaries or client data), ensuring a clean visual hierarchy. This reduces the risk of losing key information in a linear chat history and supports a more structured workflow. (See Figure 14 for layout example.)

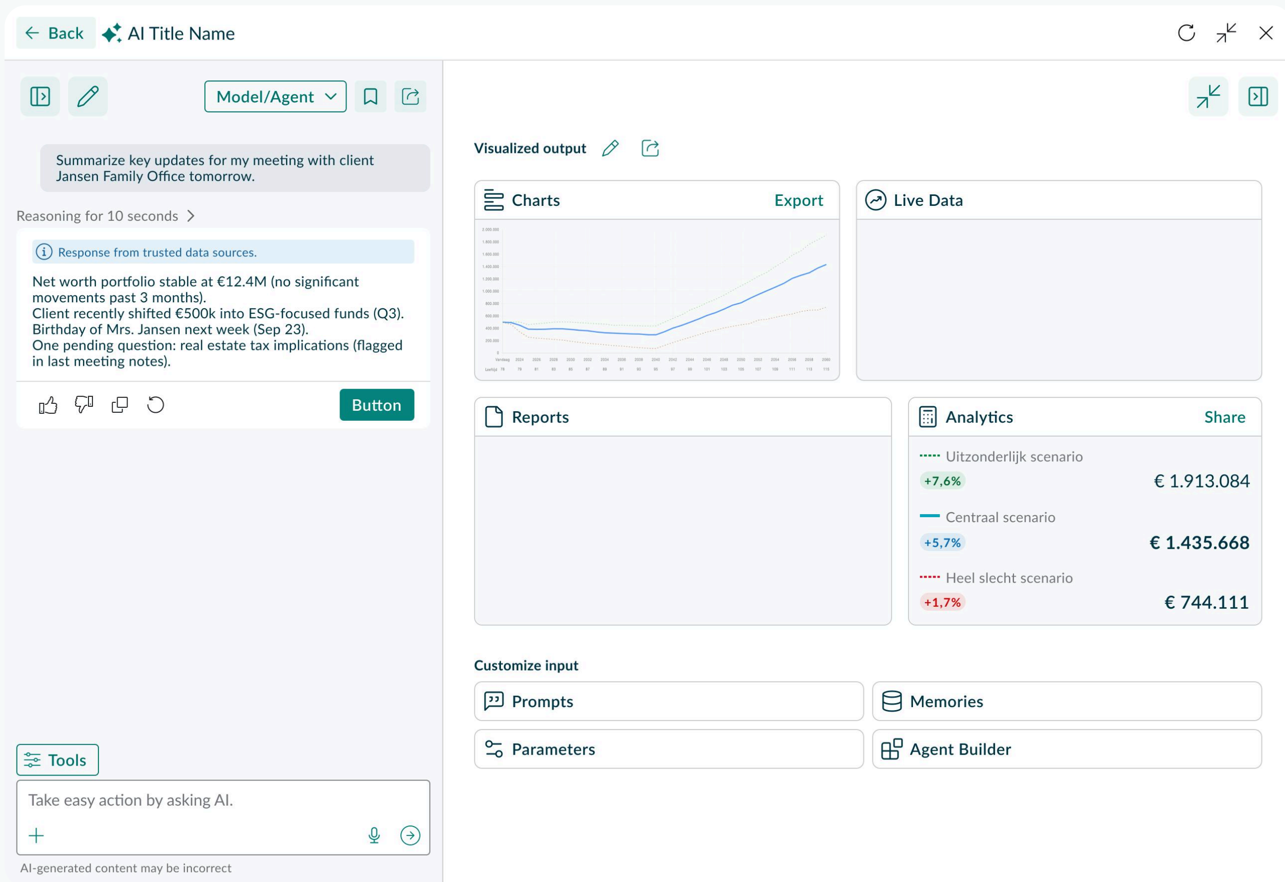


Figure 14: full-page AI workspace example

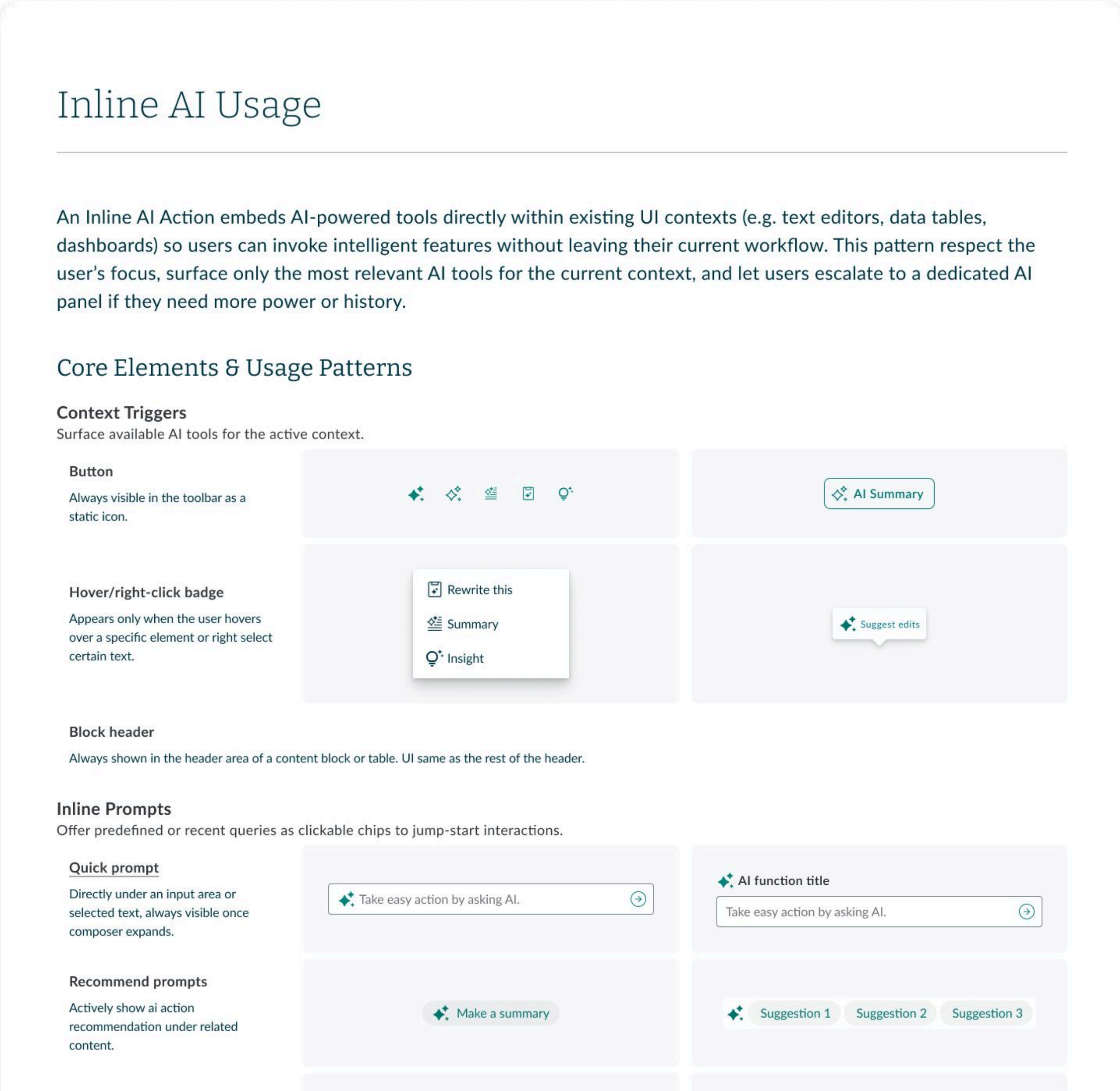
b. Pattern Deep Dive Example: Inline AI Usage

These panel types each have corresponding design documentation pages that outline their structural variations, do's and don'ts, and interaction best practices.

Taking the Embedded Inline Usage as an example (Figure 15), the page outlines its core elements and usage patterns, covering:

- Context triggers: buttons, hover states, right-click badges
- Prompt types: quick prompts, AI-generated suggestions, manual input fields
- Output behaviors: inline display, floating feedback, tooltip-style summaries
- Escalation links: "See more" buttons, contextual links to AI workspace

Each sub-pattern includes annotated examples and detailed dos and don'ts to guide consistent implementation and reduce user confusion.



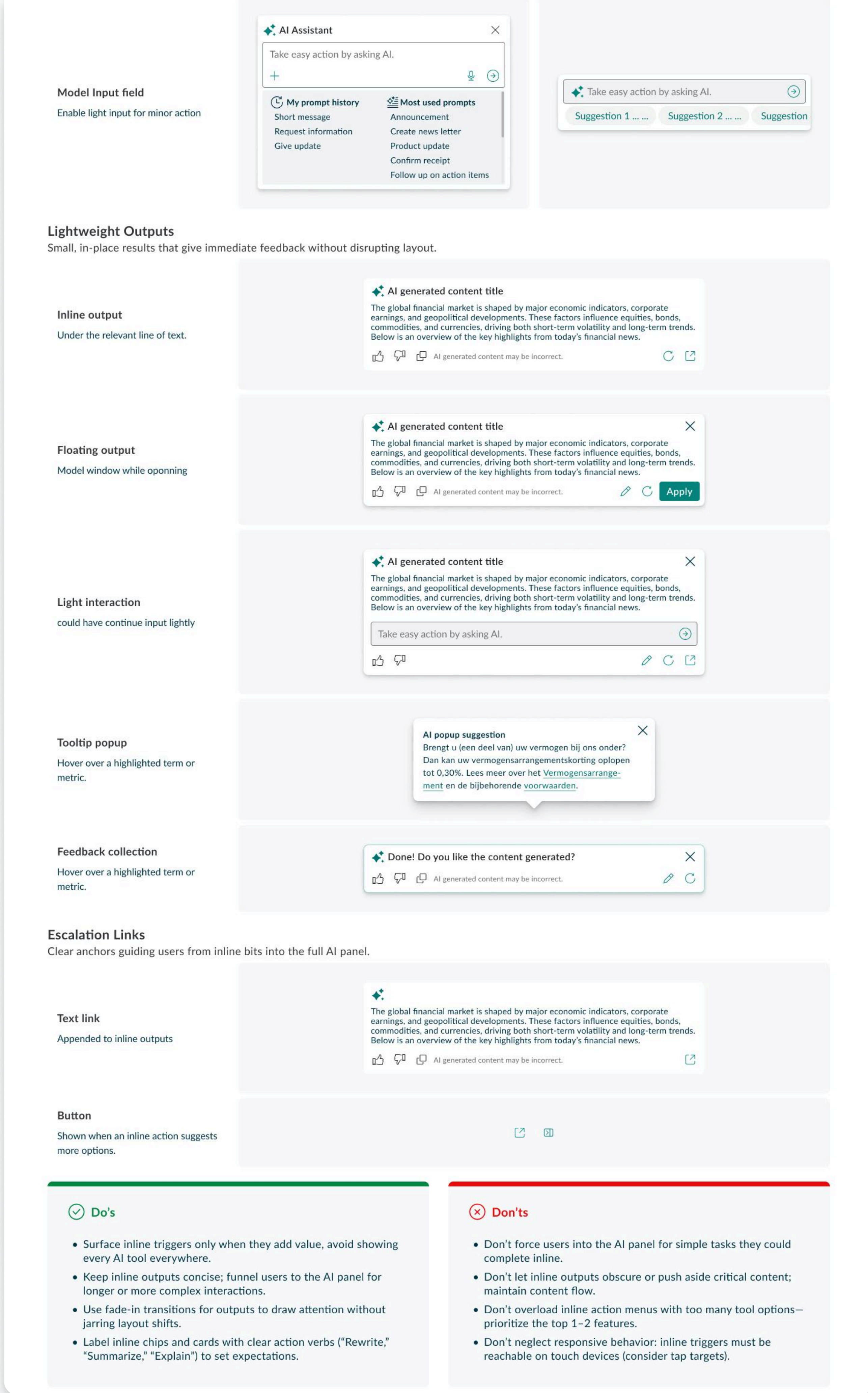


Figure 15: Template page: Inline AI usage

c. Organisms: Prompt Inputs and Chat Bubbles

At a more granular level, the two most critical component types in AI interaction are the Prompt Input and the Chat Bubble.

A Prompt Input is the interface container where users write, refine, and submit AI queries (Figure 16). Whether embedded, in a sidebar, or activated via voice or file upload, it adapts in complexity and layout. I listed all potential input types (text, file, audio, tool selection, etc.) and offered format variations optimized for different panel sizes. These considerations were especially important given that prompt construction is often a barrier for non-technical users, as surfaced in the interviews.

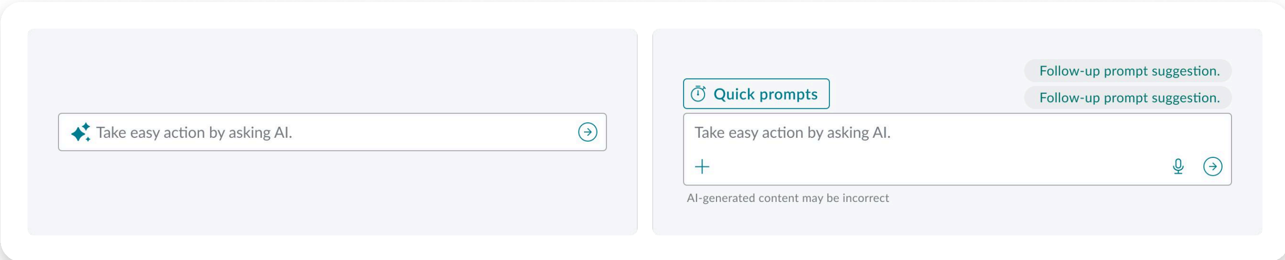


Figure 16: prompt input example

The Chat Bubble is more complex than prompt inputs. In designing it, I considered the different types of communication that occur in AI-assisted financial workflows. To address this, I created versions of the bubble for three message types:

- User Prompts: showing the user's queries or instructions
- AI Responses: summarizing, answering, or recommending actions based on system context
- System Notices: status indicators such as timestamps, loading animations, or fallback messages

By designing these as configurable versions rather than fixed templates, I ensured that designers using the framework could adapt the chat bubbles without rebuilding them from scratch. For instance, in the AI response bubble, they can toggle layouts directly in the component's property panel (see Figure 17). This approach makes the chat bubbles both predictable in structure and flexible enough to fit different financial use cases.

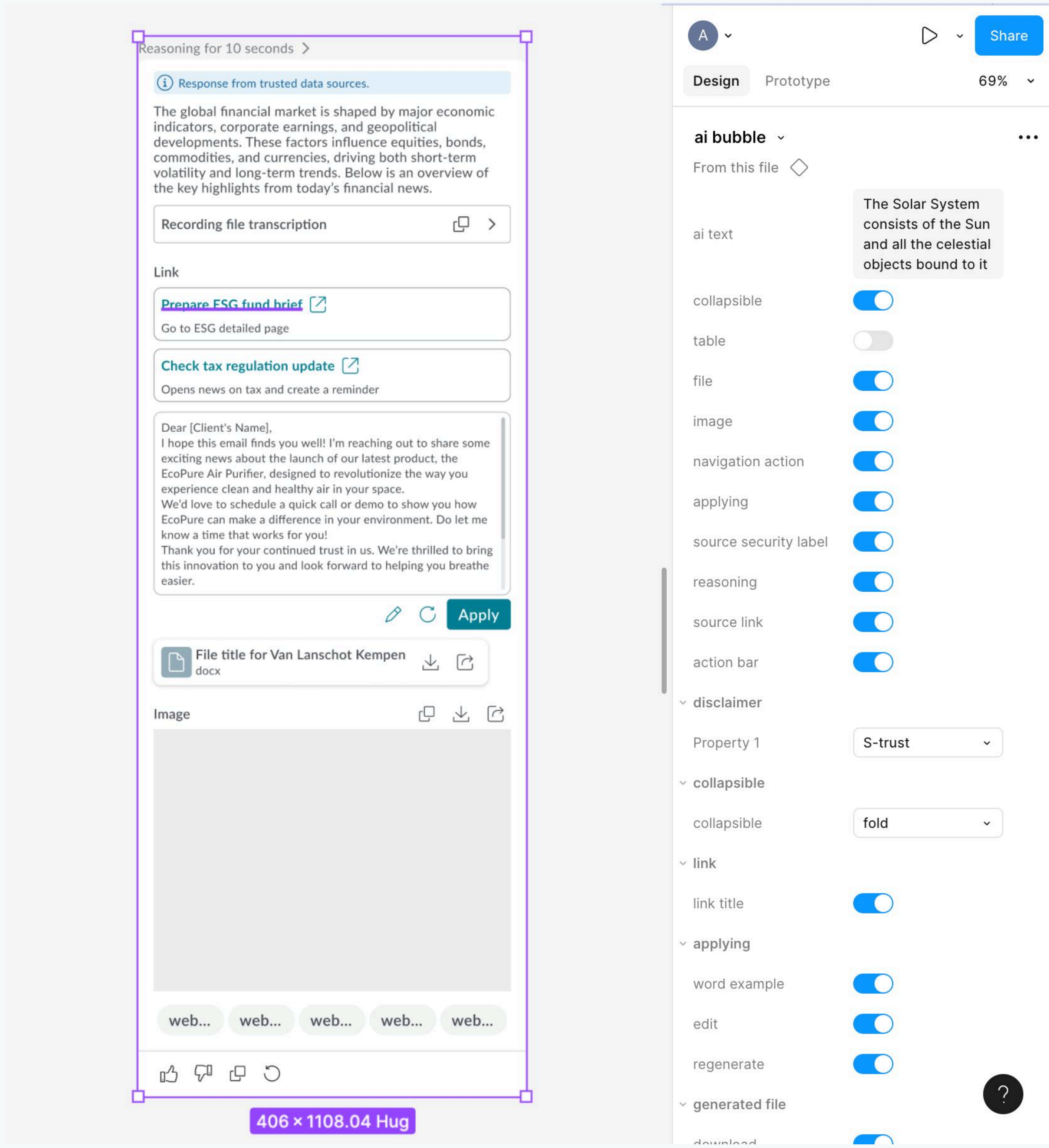


Figure 17: template AI Response widget & component's property panel

d. Atoms: Icon Buttons & Tool Tags

Beyond high-level patterns and mid-level components, I also designed and specified rules smaller assets that frequently appear across tools.

Icon buttons were sourced from company's existing design library and reused according to placement and task criticality. In less familiar contexts, such as dashboards or entry points to new AI tools, icon + label combinations should be used for clarity. Conversely, in focused task environments where meaning is already established, icon-only buttons are acceptable.

Newly designed tags , such as those for AI tools, names, or document categories, help users structure their queries and inputs more quickly. Consistent tagging also supports easier cross-page navigation and improves perceived control.

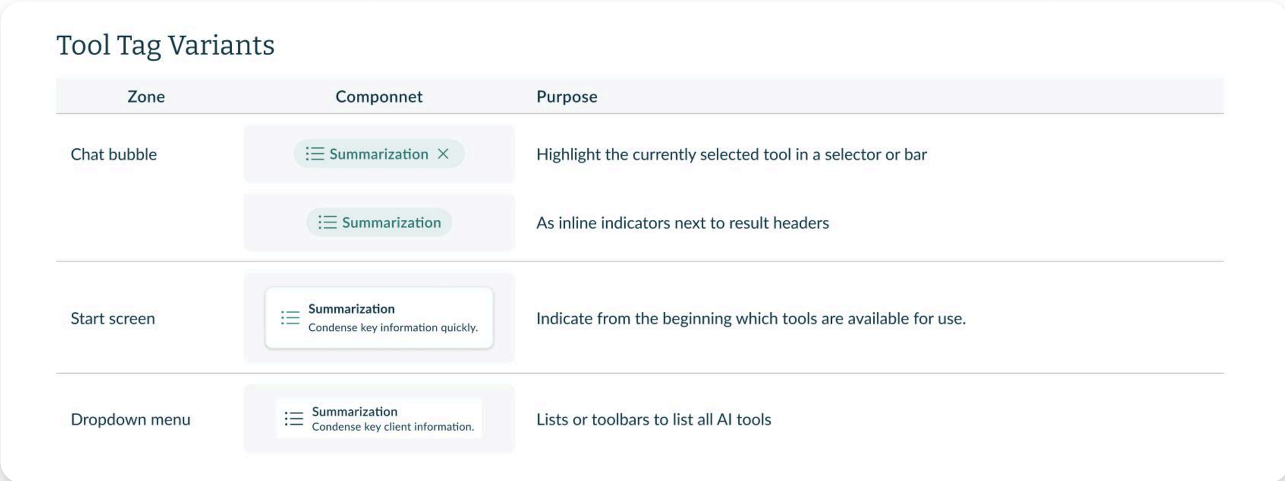


Figure 18: tool tag design

Altogether, this component library does more than simply document reusable pieces, it defines the shape of AI interactions within VLK's internal tooling. It responds directly to users' desires for transparency, contextual relevance, and clear next steps, as surfaced in the research. By grounding these components in real workflows and connecting them to existing design infrastructure, the system becomes scalable, practical, and immediately actionable. The next section will turn to the Design Guide, which builds on these components to offer decision-making principles and logic behind their application.

4.3.3 AI Feature & Component selection Wizard

To help designers and product owners navigate this extensive component library, I developed an AI Feature & Component Selection Wizard. Rather than expecting users to browse dozens of pages manually, the wizard walks them through the critical factors they must consider when introducing a new AI feature. By answering a short series of targeted questions, they arrive at a customized set of components and interaction patterns tailored to their specific use case. Under the hood, a simple scoring algorithm weights each answer and highlights the top recommendations—making the process both efficient and pedagogical.

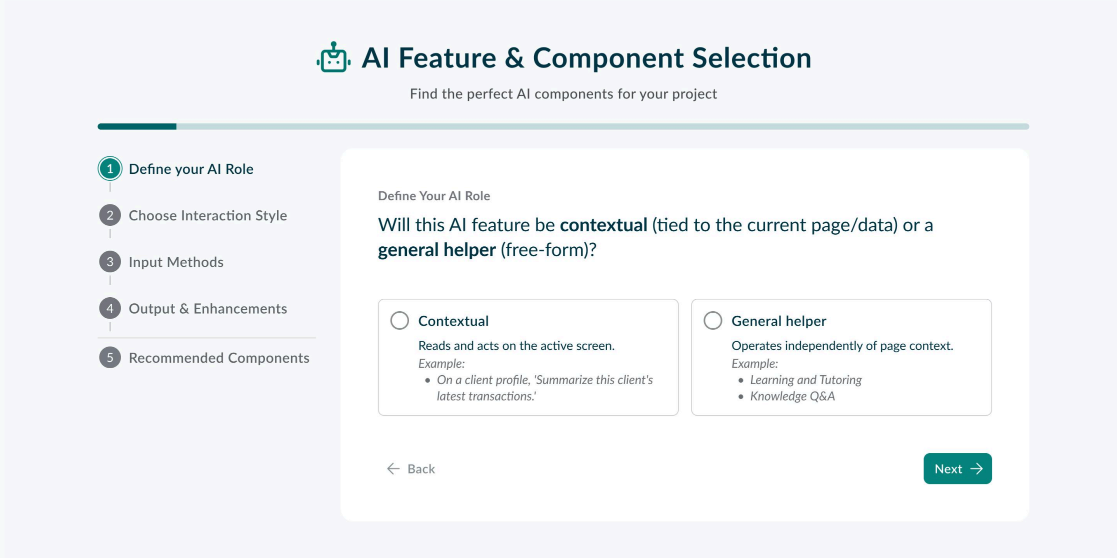


Figure 19: AI Feature & Component Selection Wizard - Step 1

Step1: Define Your AI Role

The wizard begins by asking whether the feature is contextual (tied directly to the current page or data), or a general helper that operates independently. For example, summarizing a client's recent transactions on their profile page is clearly contextual, whereas drafting an email about upcoming events is a free-form task. If the user selects "General helper," the wizard advises pointing product owners to existing standalone tools like Leonardo (internal ChatGPT). If "Contextual" is chosen, the wizard proceeds to the next step, ensuring only relevant panel formats and components are considered.

Determining whether an AI feature is contextual or general shapes every subsequent design decision. Contextual AI must integrate with existing data and interface, so designers need components that can read the current view and augment it in place. General helpers, by contrast, require standalone containers and a broader input/output mechanism. By asking this first, the wizard ensures that designers select the right container type from the start—and it prevents wasted effort building context-free features where deeper integration is required.

Step 2: Choose Interaction Style

In this multi-part step, designers clarify the scope and behavior of their AI feature:

2A. Single Function vs. Extensible Platform
Will this be a one-off function (triggered once with a single response (e.g., “Rephrase this paragraph”)) or an extensible platform supporting multi-round conversations, tool switching, and persistent history (e.g., a sidebar chat that follows users across pages)?

2B. Trigger Model
Should the AI wait for an explicit user prompt, or should it proactively surface tips and cards based on contextual triggers such as KPI anomalies? Choosing “User initiated” keeps the AI in the background until summoned, while “AI driven” enables timely in-context suggestions.

2C. Interaction Depth
After the AI delivers its first result, will users accept it and move on (single round), or will they engage in back-and-forth refinement (multi round)? Designers looking to foster deeper exploration will lean toward multi-round flows, whereas quick confirmation tasks may be better suited to single-round interactions.

At the end of Step 2, the wizard’s scoring algorithm tallies points for each answer, surfacing the four best-fitting UI containers in the next section.

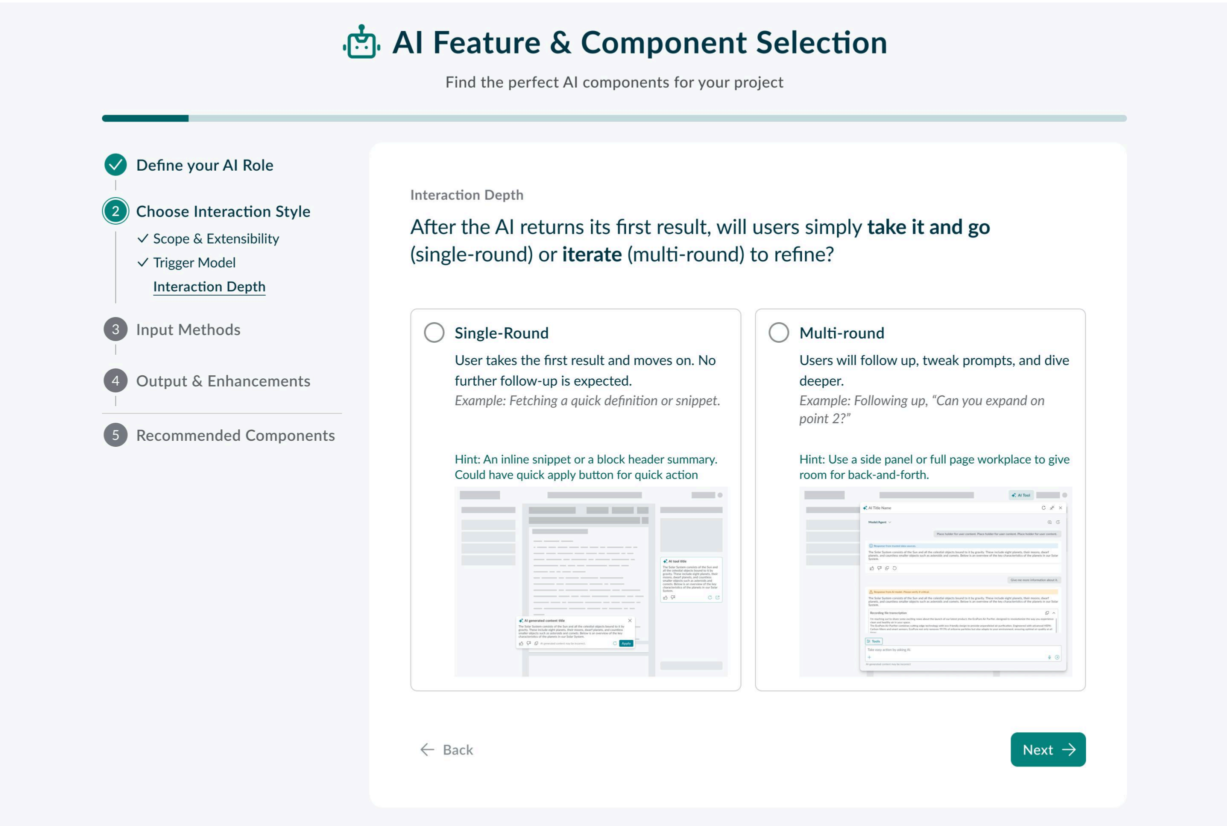


Figure 20: AI Feature & Component Selection Wizard - Step 2C

Select Your Container
Based on the interaction style choices, the wizard recommends up to four interface containers. Options include a floating chat widget for lightweight, asynchronous Q&A; floating modals (simple or extended) for focused, single or multi-tool sessions; docked side panels with persistent chat and history; and full-page workspaces for deep, immersive exploration. Each recommendation comes with a brief rationale, helping designers understand why, for example, an extended modal with history support best suits an extensible, multi-round platform.

The chosen container not only defines the visual layout but also sets the interaction dynamics: how much screen real estate AI can occupy, whether it follows the user across views, and how prominently suggestions appear. By narrowing to the top four options, designers avoid misaligning feature purpose with container form, and they gain a clear rationale for each recommendation.

Step 3: Configure Input Methods

Next, users specify how they intend to capture user inputs. Multiple selections are allowed (seen Table), covering standard text prompts, voice input, file uploads, live data feeds, tool/model selectors (e.g., choosing GPT-4 vs. a domain-specific agent), and presets & history for prompt reuse. This step ensures the chosen components include the correct input fields, buttons, and controls.

Text prompts	Standard text field.
Voice	Microphone icon for speech-to-text.
File Upload	Attachment button for documents, images, tables.
Live Data	Automatically pull page or backend metrics.
Tool/Model Selector	Dropdown or chips to pick GPT-4, FinanceBot, CustomAgent.
Presets & History	Buttons or chips for preset prompts, recent history, or recommended queries.

Table 6: input methods

Careful consideration of input modalities is fundamental to ensuring that the AI feature can function effectively within the given workflow. Different tasks necessitate different input types, such as free-text fields for writing tasks, document upload for analysis, or real-time data for contextual responses. By explicitly selecting *input methods, designers are not only guided toward assembling the appropriate UI components (e.g. input fields, icons, and controls), but are also prompted to reflect on the nature and format of information the AI will require. This step serves a dual purpose: it guarantees technical compatibility while also acting as a catalyst for designers and product owners to clarify the intended function, scope, and expected output of the AI feature, fostering a more intentional and informed design process.*

Step 4: Define Output & Enhancements

Users then pick from a wide range of output formats and auxiliary features, such as inline snippets, summary cards, charts and dashboards, auto-action buttons, loading indicators, AI reasoning panels, source labels, file/media embeds, source links, navigation actions, and timestamps. These options populate the final list of components needed to build the feature.

Inline Snippet	A one-line suggestion or rephrasing.
Summary Card	Titled card with summary text and action buttons.
Charts & Dashboard	Graphs or data-grid widgets.
Auto-Action Button	e.g. 'Generate Report,' 'Send Email.'
Loading Indicator	Spinner or skeleton during AI processing.
AI Reasoning Panel	Show the AI's brief 'chain-of-thought.'
Source Security Label	Indicate data is from online/offline source or secure.
File/Media Embed	Downloadable reports, image/video outputs.
Source Link	Link back to original document or knowledge base.
Navigation Action	Buttons to navigate elsewhere in the app ('View Client Details').
Timestamp	Display generation or update time for time-sensitive data.

Table 7: output formats

AI-generated outputs can take many forms. The selection of output format is not only a matter of UI composition, but also a means of shaping how users interpret and act upon AI-generated insights. This step ensures that designers deliberately consider what the AI should deliver and how users will engage with it. Importantly, by requiring teams to commit to specific visual and interactive output elements, this process forces a deeper reflection on what kind of value the AI is expected to provide. In other words, the output format becomes a constructive constraint, pushing designers and product owners to clarify the technological capabilities needed and the informational completeness required. The act of defining the output in concrete terms thus becomes a driver of critical thinking and design intentionality.

Step 5: Review Recommended Components

Finally, the wizard compiles your selections into a complete AI feature specification:

Aspect	Your selection
Container	From Step 2
Input Methods	From Step 3
Output & Enhancements	From Step 4

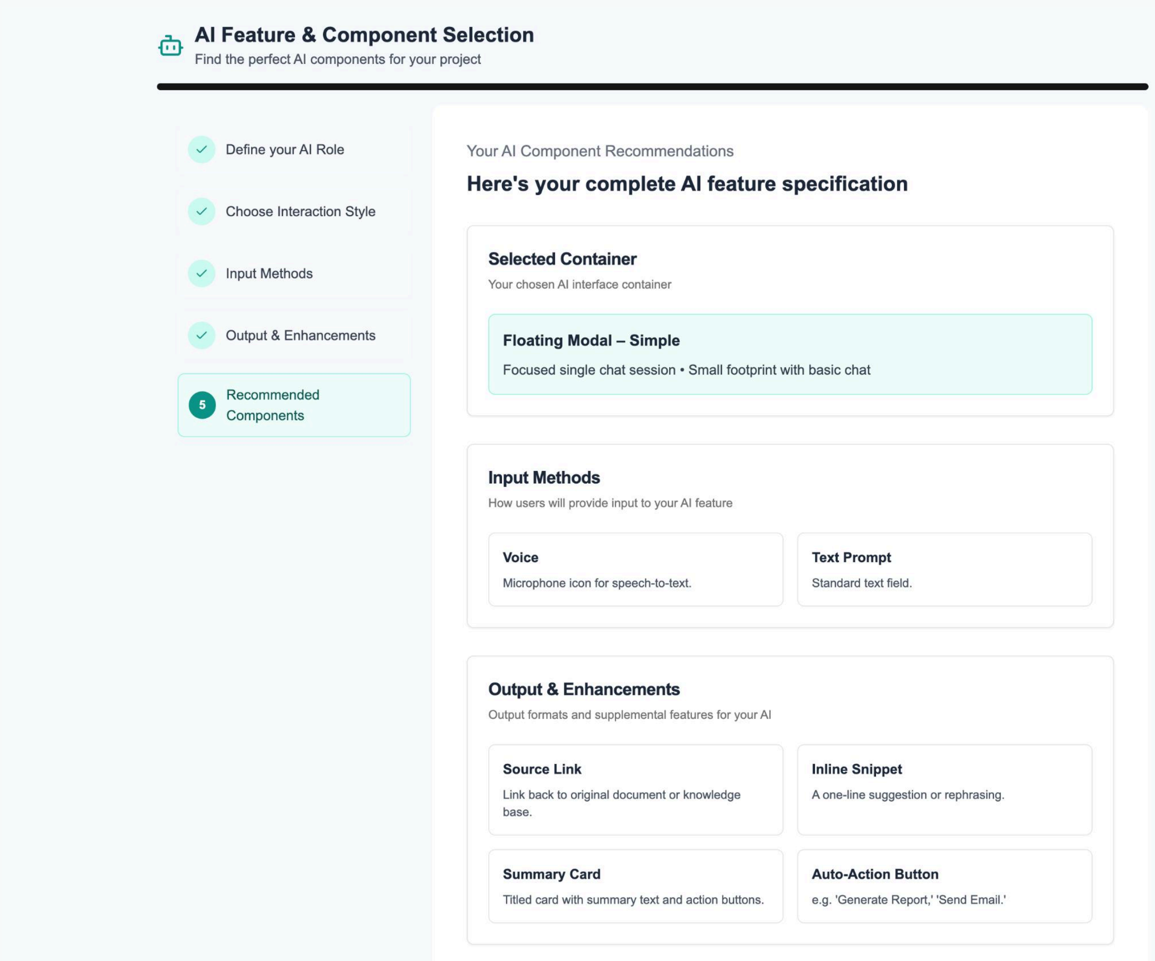


Figure 21: AI Feature & Component Selection Wizard - Recommended components

It also includes a "Getting Started" guide for Figma:

1. Access the Figma AI component library.
2. Open the "Chat Panel Format" page and copy your selected container into the draft page.
3. Use the right-hand properties panel to toggle input and output options until the design matches your requirements. Also look into other component files for further changes.
4. If you need further assistance or imaginative brainstorming, reach out to the "AI design superheroes" on your team for support.

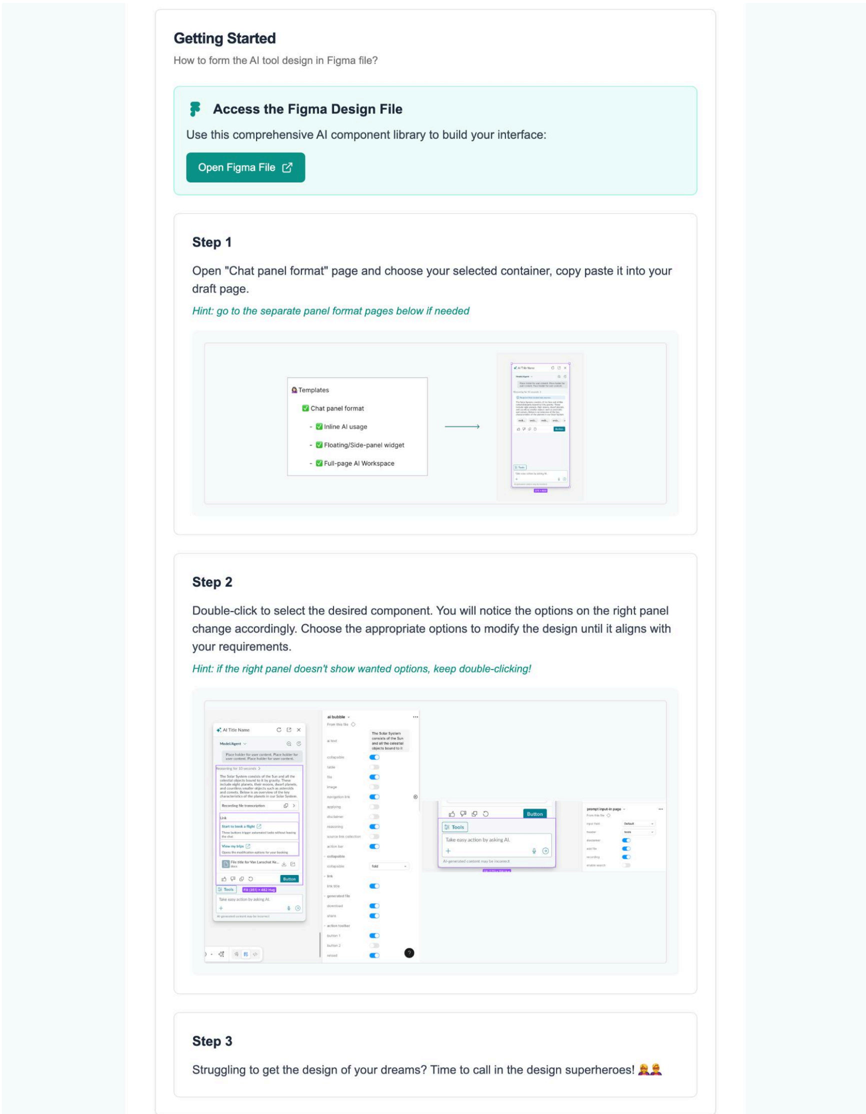


Figure 22: AI Feature & Component Selection Wizard - Getting started

A consolidated specification saves time and reduces error, giving designers a ready-made blueprint. The "Getting Started" instructions bridge the gap between selection and execution, showing exactly how to integrate the chosen components into a working prototype. By closing the loop from conceptual factors to concrete assets, this final step transforms the library into an actionable design partner.

Behind the scenes, the wizard leverages a simple algorithm: each Step 2A, 2B, and 2C choice is assigned a point value for each container type. The wizard then sums these points and highlights the top four scoring containers, with the highest one marked as the primary recommendation. This dynamic scoring ensures that container suggestions align closely with the intended feature scope and user expectations.

By guiding designers through these decision points (Defining Role, Interaction Style, Input Methods, and Output Formats) the AI Feature & Component Selection Wizard transforms the component library from a static reference into an interactive design partner. This approach bridges research insights and design execution, making it easier for VLK's design and product teams to deliver consistent, transparent, and scalable AI experiences.

5. Evaluation

To assess the practical value and usability of the proposed AI interaction design framework, a formative evaluation was conducted with five internal stakeholders at Van Lanschot Kempen: three designers and two product owners. The aim was to explore whether the framework could effectively support AI feature design across roles, promote consistent interaction patterns, and facilitate scalable implementation. By asking participants to use the AI Feature & Component Selection Wizard in a realistic task scenario, the evaluation probed how well the framework clarified design decisions, guided component selection, and supported individual workflows.

The results affirm the framework's foundational value. Participants expressed a clear willingness to adopt the tool in future projects and noted that it also stimulated more structured thinking about how AI tools should behave within specific work contexts. At the same time, the evaluation surfaced critical areas for refinement. These include clarifying AI-specific terminology, adding richer visual to support decision-making, and improving the transition from wizard output to component selection. Insights also revealed differences across roles, underscoring the importance of role-aware design.

Overall, the evaluation confirmed that the framework is not only usable and well-scoped, but also capable of supporting cross-disciplinary collaboration in AI tool development. The findings provide actionable guidance for enhancing its accessibility, scalability, and long-term integration into internal design and product workflows.

5.1 Evaluation goal

The primary goal of the evaluation phase is to assess the integrity, availability, and scalability of the AI interaction framework in real design scenarios at VLK. Specifically, we aim to verify that:

1. **Framework Logic Is Clear and Understandable**
   Designers and product owners should be able to grasp the overall structure without extensive training. We will measure this through task-based walkthroughs, observing whether participants can articulate how the framework's layers fit together and why each step is necessary.
2. **Ease of Use for Product Owners**
   Product owners, who may have limited design expertise, should be able to leverage the framework to produce valid AI feature specifications quickly. We will track the time and number of assistance requests required for them to complete a simple AI feature definition using the Wizard, as well as their subjective ratings of ease and confidence.
3. **Advanced Application by Designers**
   Experienced UX designers should be able to use both the component library and the underlying principles to create polished, high-fidelity prototypes for complex AI interactions. We will evaluate this by assigning designers a scenario that demands multi-step AI integration and reviewing their prototype for consistency with framework guidelines, component correctness, and creative adaptation of reusable elements.

By focusing on these three dimensions the evaluation will determine whether the framework can withstand typical usage, scale across teams, and adapt to increasing complexity as the company's AI initiatives evolve.

5.2 Evaluation set-up

5.2.1 Evaluation method

To assess how well the AI interaction framework supports real-world design work, we conducted a series of moderated user tests, each lasting approximately 30 minutes. Participants were asked to design and specify a new AI feature for the ClientCenter platform.

The feature brief was:

Task Goal
Proactively present AI suggestions for commercial opportunities to bankers.

Scenario
Imagine you've been asked to extend ClientCenter/Horizon with an AI module that analyzes a client's profile (including recent interactions, transactions, and product use) in real time and surfaces timely, relevant actions the banker can take to serve the client better and capture business opportunities. You will use the AI Feature & Component Selection Wizard to define the feature's role, interaction style, input/output methods, and UI container.

In each session, I began by outlining the design challenge and introducing the framework as the tool they would use. Participants were presented with the scenario and given a brief demonstration of the wizard's flow. This setup ensured everyone understood the task goals and the purpose of the framework before diving into the exercise.

Participants then worked through the wizard independently, making decisions about the feature's contextual role, interaction style, UI container, and input/output methods. As they progressed, I observed where they hesitated, asked for clarification, or experimented with different options. These moments revealed which questions were intuitive and which required more explanation, guiding potential refinements to the wizard's language and examples.

Upon completing the wizard, each participant reviewed the generated feature specification and component list. They evaluated whether the recommended container and components matched their mental model of the task, rated the overall clarity and completeness of the output, and assessed how confident they felt in using those recommendations to begin a real design. Their feedback captured both numerical ratings and candid comments about the framework's strengths and gaps.

The session closed with a reflective discussion tailored to each participant's role. Designers were asked whether the level of detail in the component properties supported advanced prototyping and how they might adapt the components for visual consistency. Product owners, on the other hand, focused on whether the process helped them articulate clear requirements before engaging design or development teams, and whether the wizard streamlined their own planning activities. These role-specific insights helped us understand how the framework serves different stakeholders and where further customization may be needed.

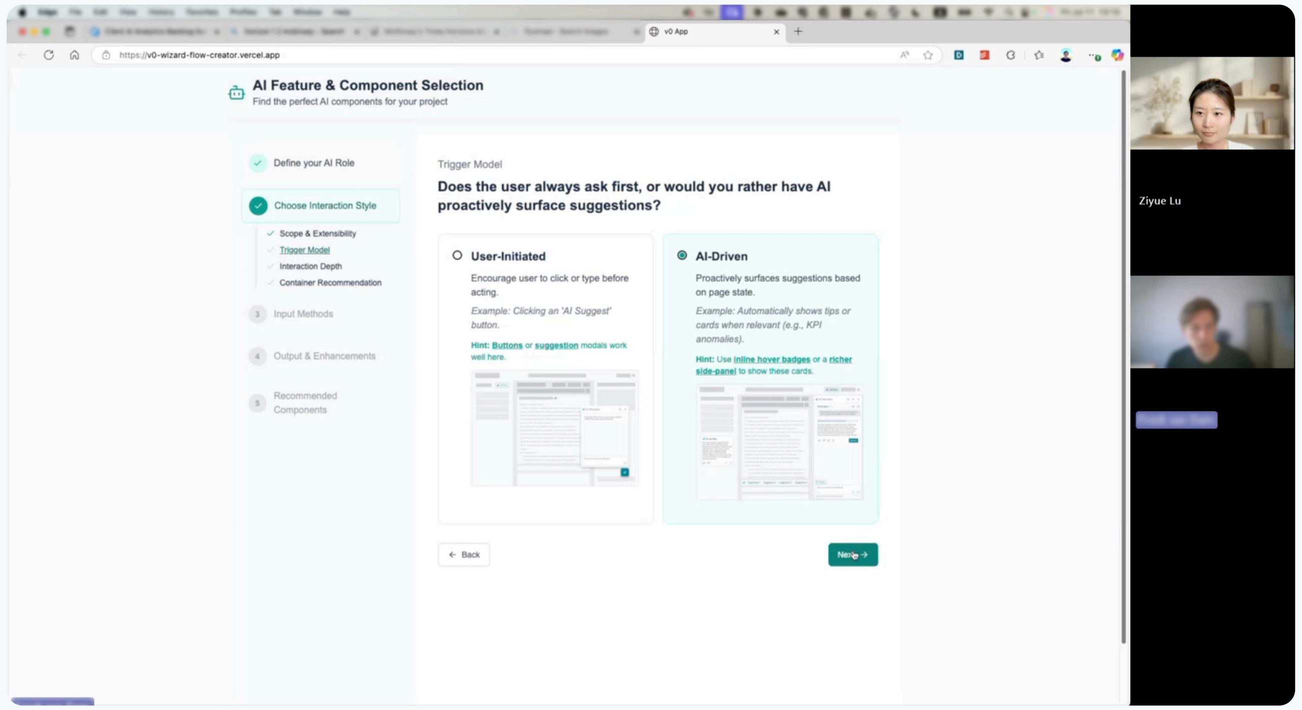


Figure 23: testing process with one of the product owners

5.2.2 Participants demographic

To ensure that the evaluation reflected the real dynamics of AI feature development at VLK, I employed a purposive sampling approach. 5 participants were recruited: 3 UX designers from UX Design & Research team and 2 product owners who regularly collaborate on internal tool initiatives. In the company, product owners frequently engage in preliminary design discussions and wire-framing, making their perspectives on the framework's usability and requirement-clarification process especially valuable.

Recruitment was conducted via direct messaging on the company's collaboration platform, with all 5 sent invitations resulting in participation. Each potential participant received a brief note explaining the study's objectives and 30-minute time commitment. All invitees responded positively, resulting in a diverse group that combined design expertise with product management insight. This mix allowed us to observe both how senior designers leveraged the component library and wizard for advanced prototyping, and how product owners used it to define clear, actionable AI feature requirements before handing off to design or development teams.

5.3 Evaluation results

*The evaluation revealed a generally **positive reception** of the AI interaction design framework across both designers and product owners, affirming its value as a structured yet flexible tool for supporting early-stage AI feature ideation. Participants highlighted the framework's clarity, ease of use, and ability to prompt critical thinking about key design decisions. Importantly, the framework was seen not only as a means of aligning with existing design systems, but also as a prompt for deeper reflection on what effective AI support should look like in a workflow context.*

*At the same time, feedback surfaced **opportunities to refine** the framework further. These included clarifying AI-specific terminology, expanding visual and contextual guidance and improving the connection between the wizard's output and the Figma component library. These gaps, while not undermining the tool's core value, point to ways it can be made more accessible and actionable for a wider range of users.*

*Finally, notable **differences** emerged in how **designers and product owners** engaged with the tool. Designers tended to use the component library as their primary resource, while POs leaned on the wizard to map out business goals and technical feasibility. This reinforces the importance of role-sensitive entry points and underscores the framework's potential not only as a design artifact, but as a shared planning tool across disciplines. Together, these insights validate the direction of the project while providing actionable guidance for its continued evolution.*

5.3.1 Overall comments

Across all five evaluation sessions, participants expressed a strongly positive attitude toward the AI interaction framework. Even those with limited prior exposure to structured design systems found the process intuitive and approachable. As one designer noted,

> "
> *Yes, it's quite interesting and it helped me, especially I like the input part because sometimes we think about what can be the input, but it's never clear until we are doing the design... that actually is the essential part when you are creating the design. So I like that part and the interaction style.*
> *-- Designer A*

This comment underscores how the framework not only guided their component selection but also prompted them to reflect on critical design decisions they might otherwise overlook.

Participants felt that the Wizard provided a clear starting point for defining new AI features, particularly for non-expert users. A product owner remarked,

> "
> *Really it will be really helpful. I think this is really good starting point for all the teams that are exploring and and thinking about how to integrate in a friendly way for our bankers.*
> *-- Product owner A*

This feedback highlights the framework's ability to make AI tooling clear for stakeholders beyond the UX team, streamlining collaboration and ensuring that everyone shares a common understanding of feature requirements.

The framework's flexibility was also praised. Designers appreciated that it could accommodate both simple, one-off functions and more complex, extensible platforms. One participant summed this up:

> "
> *As a developer, I think they will be able to see [how it works]; after exploring it for a few minutes, it should be quite easy.*
> *-- Designer C*

This remark speaks to the framework's balance of structure and adaptability—offering enough guidance to prevent decision paralysis, while remaining lightweight enough to support advanced customizations.

Finally, users reported feeling confident in their design choices after completing the Wizard, noting that the generated component list aligned closely with their mental models of the task. By forcing explicit consideration of inputs, outputs, and interaction flows, the framework helped participants think holistically about how an AI feature would function within the employees workflow rather than jumping straight into wire-framing.

Overall, these insights confirm that the framework is both usable and valuable, and that participants are willing to adopt it for future AI tool design at VLK.

5.3.2 Enhancement Opportunities

While participants found the framework highly valuable, several areas emerged where small refinements could greatly improve clarity and adoption.

a. terminology clarity

During the evaluation, several participants hesitated at terms such as “trigger” and “live data,” which were presented without contextual definitions. For example, they were not sure whether “trigger” referred to user actions, system events, or scheduled processes, and whether “live data” meant real-time streaming or periodic refresh. Without a clear understanding of what these terms signified, users were unsure how to apply these options to their feature design. This uncertainty not only slowed down their decision-making but also undermined their confidence in the wizard’s guidance.

Improvement
To address this gap, each technical term should be accompanied by an inline **tooltip** or “info” icon that, when hovered over or clicked, displays a concise definition and a concrete example.

For instance, hovering over “trigger” might show: “An event that initiates the AI feature, such as ‘User clicks “Generate Report” button’ or ‘System detects new transaction.’” Similarly, “live data” could be annotated as: “Automatically pulls updated account balances or recent transactions every five minutes.” (Figure 24)

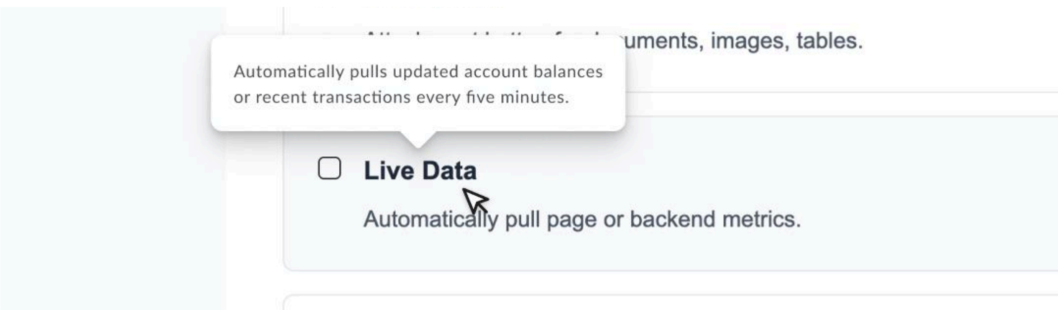


Figure 24: improvement on terminology clarity

By embedding these micro-explanations directly into the interface, users can quickly resolve confusion and proceed with greater clarity.

b. enhanced visual guidance

While the initial container selection screen provided helpful diagrams, the subsequent steps for selecting input and output options relied exclusively on text checkboxes. Participants reported that without visual examples, it was difficult to envision how summary cards, charts, or inline snippets would actually appear in the interface. As one product owner remarked,

> “
> *I want to see what this looks like before I commit to it, otherwise I’m just guessing.*
> *-- Product owner B*

Improvement
To make these decisions more intuitive, the wizard should integrate thumbnail previews alongside each input/output option. For example (seen Figure 25), when choosing a “Voice input,” a miniature mockup could illustrate its layout.

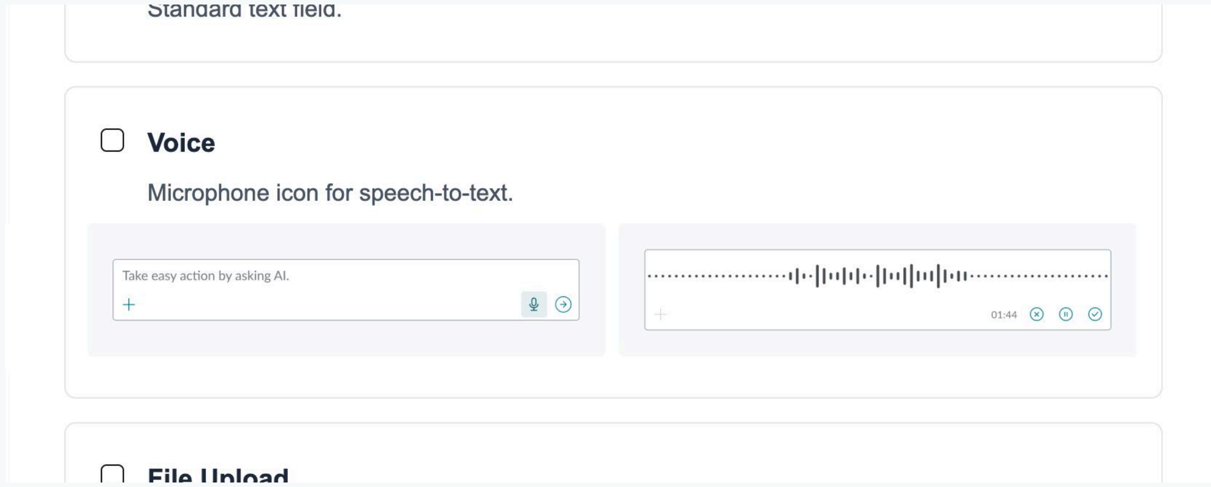


Figure 25: improvement on terminology clarity

Additionally, a preview pane could dynamically update to show the user’s current selections in a composite mockup, allowing side-by-side comparisons. Reordering the wizard so that visual previews appear earlier in the flow, and maintaining a persistent “live preview” area, will align more closely with users’ visual thinking processes and reduce uncertainty.

c. fluent handoff to Figma components

After completing the wizard, participants received a list of recommended components but struggled to locate them within VLK’s Figma library. In some cases, the wizard’s terminology did not exactly match the Figma page titles, and users were left guessing which library section contained the needed assets. This disconnect added additional search time and occasionally required moderator intervention.

Improvement
To streamline the transition from specification to implementation, each component in the final list should include an exact Figma reference tag. For instance, "Component: Summary Card – Figma Page: chat bubble/organisms ". These tags could function as clickable links in a digital export, directly opening the relevant Figma page. Consistently aligning the wizard's titles with the design system's file structure will eliminate ambiguity, ensuring that both designers and product owners can locate and apply the recommended elements immediately.

By incorporating these enhancements the framework can become even more intuitive, reducing cognitive friction and further empowering both designers and product owners to build AI features with confidence.

5.3.3 Role-Based Perspectives

Although all participants found value in the AI interaction framework, their priorities and usage patterns differed markedly by role.

UX Designers tended to focus on the component library, viewing the wizard as a helpful scoping tool primarily for less experienced colleagues. While they appreciated the structured flow of the wizard, experienced designers often moved quickly into Figma to begin hands-on work, relying on the library's adherence to established design system conventions and the availability of universal components and buttons. One designer noted that the wizard could serve as a "jump-start" for ideation sessions or onboarding new team members, but that real design refinement inevitably took place directly in the design tool, where they could manipulate layouts and styles with greater precision.

Product Owners, by contrast, were most interested in the wizard's end-to-end flow. They valued its ability to frame practical business scenarios and spark conversations around requirements, use cases, and technical feasibility. Product owners were realistic about the inherent complexity of AI features in a regulated environment and cautioned that the wizard's scope would need to be complemented by deeper cross-functional collaboration. As one product owner observed,

> "
> *This tool helps me map out capabilities and align stakeholders, but the real work of defining data pipelines, compliance checks, and technical integrations will come afterward.*
> *-- Product owner A*

For them, the framework functioned less as a design execution tool and more as a shared language for brainstorming and early-stage planning.

These role-based insights highlight the importance of balancing detailed component specification for designers with scenario-driven guidance for product owners. Future iterations of the framework may consider customizing views or entry points. Specifically, streamlining access to the component library for designers while emphasizing the wizard's flow and outcome summaries for product owners. By this way, it could to better meet each group's distinct needs.

6. Discussion

In this chapter, we step back from the concrete deliverables and empirical findings to examine what they mean in practice and how they might evolve over time.

We begin by assessing how successfully the design outcome address the goals laid out in the research phase, aligning user-centered principles with industry best practices and ensuring accessibility across roles.

Next, we explore the framework's capacity for growth: how it can be embedded into company's existing development processes, extended to mobile platforms, and tested in other industries to validate its broader relevance.

Then, we look ahead to emerging opportunities that will keep the framework responsive to future technological and organizational shifts.

Finally, we acknowledge the study's boundaries and consider how these factors shape both the strengths and blind spots of the current framework. Together, these reflections offer a roadmap for refining and scaling this AI interaction system in the years to come.

6.1 Bridging insights and design

In this section, the focus is on how exploratory research directly shaped each stage of the design process, and how effectively the resulting framework fulfills the original ambitions. The path is first traced from user insights to the establishment of design goals, principles, and artifacts. This is followed by an examination of how the evaluation results confirm that the system delivers on the promise of scalable, transparent, and human-centered AI integration. Lastly, it also discussed whether the outcome fit the

6.1.1 From research insights to design foundation

The design process was grounded at every step in research insights gathered from bankers, relationship managers, and investment advisors. Early in the project, key pain points including unclear ownership of AI actions and the cognitive burden of switching among tools were distilled into the design goals outlined in Section 4.1. These goals, emphasizing transparency, consistency, collaboration and scalability, served as a guiding reference throughout concept development.

Each major research finding was translated into the concrete Design Principle. For instance, the need for reversibility and oversight led to the Human Control principle, which also informed the inclusion of "undo" affordances in components. Similarly, struggles with understanding AI rationale inspired the Transparency principle, which formed the basis for labeling AI outputs and providing inline citations.

To operationalize these principles, atomic design was applied to structure the Component Library. Each component was designed to embody one or more principles: clarity in layout, tooltip explanations, or clearly marked AI-generated content. This systematic method ensured that all interaction patterns remained consistent with insights derived from interviews.

Finally, to address the challenge many designers faced in translating abstract principles and components into coherent features, the AI Feature & Component Selection Wizard was developed. Its task-based flow reflects the observed need for guided prompts to define interaction modes. By encouraging designers and product owners to articulate each decision explicitly, the wizard responds to a key research insight, which is that undefined prompts is the major barriers to AI feature adoption.

6.1.2 Accessing design goal fulfillment

4.1 Design goal set out to design a scalable framework for integrating generative AI into the workflows of employees at Van Lanschot Kempen, emphasizing transparency, consistency, and human-centered control. It's ambition lies in moving beyond point solutions and create an end-to-end system that guides the creation of AI features from first principles through polished implementation.

Clear evidence of success is seen in the evaluation. Participants across roles expressed a strong willingness to adopt the framework in future projects, citing confidence in both the guiding principles and the component recommendations. In particular, the wizard's structured, decision-driven flow ensured that each AI feature began with well-defined objectives, directly aligning with our commitment to workflow coherence. These outcomes demonstrate that this framework does more than document best practices, it embeds them into a practical, repeatable process that delivers scalable, transparent, and human-centered AI interactions in a financial context.

6.1.3 Positioning within scholarly and industry context

The AI interaction framework not only builds directly on the project's empirical insights, but also aligns closely with established HCI theory and extends leading industry practices.

From a theoretical standpoint, the framework echoes Shneiderman's (2022) principles of human-centered AI, which advocate for systems that empower users through clear feedback, reversible actions, and meaningful oversight. For example, the Human Control principle, mandating undo and explicit confirmation mechanisms, implements this call for preserving user agency. Similarly, Huang et al. (2023) stress the importance of explainable AI to mitigate "black box" concerns. The inline rationale and source citations in the component library, directly responds by revealing how suggestions are generated. Elshan et al. (2022) highlight behavioral proactiveness and communication clarity as key trust drivers, concepts that are operationalized through the Wizard's proactive trigger options and clarity-focused, tooltip-enabled UI components.

When compared to industry exemplars such as Microsoft's *Creating a Dynamic UX* guidelines for copilot experiences and the *ShapeOf.AI pattern library, this* framework exhibits several points of convergence. Like Microsoft, it offers multiple panel formats tailored to different task scopes. Similarly, the classification of input/output elements mirrors the identifiers, way-finders, and trust indicators championed by ShapeOf.AI. However, this project extends these industry references by grounding every component in domain-specific scenarios and workflows. Where Microsoft focuses on broad enterprise products, we add the financial context of high-stakes data and regulatory oversight. And where ShapeOf.AI provides a general pattern taxonomy, we embed those patterns into a domain-tuned component library and decision wizard.

By bridging these academic and industry perspectives, the framework demonstrates how HCI research can inform a practical design system, while grounding industry best practices in explorative user research. In doing so, the project contributes both to scholarly discourse by offering a replicable model for deriving design systems from qualitative research, and to the professional community by delivering an actionable system that can accelerate AI adoption.

6.2 Adaptability & scalability

This section examines how the framework can be scaled and adapted beyond its current instantiation, focusing on practical integration with team processes, technical extension across platforms, and transferability to other sectors. Each area identifies concrete considerations and suggested next steps, which align with prior HCI findings and industry case studies, that support wider adoption while preserving the framework's core principles.

6.2.1 Team & process integration

One practical way to embed the framework into product development lifecycle is to treat the wizard output as a input to sprint refinement, the generated specification can feed backlog creation, estimations, and acceptance criteria. This aligns with the literature that treats design systems as operational artifacts which must be connected to engineering and delivery pipelines (Nielsen Norman Group, n.d.; Frost, 2016). Linking the wizard to ticketing systems shortens the handoff from product definition to implementation and makes traceability explicit. Equally important is aligning the framework with review gates so that AI features are assessed against regulatory and quality requirements before development begins.

Successful integration also requires role-specific onboarding and support. Designers will benefit from example Figma files, component usage guides, and template starter pages; product owners need concise cheat-sheets that translate wizard outputs into business requirements and risk checkpoints. A layered training program such as hands-on workshops, recorded walkthroughs, and a searchable FAQ could help different roles reach competency quickly. Additionally, appointing internal champions or design librarians who can advise teams and curate the library encourages best-practice reuse. This emphasis on role-tailored documentation and internal champions is supported by empirical studies showing that practitioner-facing resources are most effective when complemented by organization-specific training and easy-to-consume artifacts (Yildirim et al., 2023; People + AI Guidebook, n.d.).

Finally, process integration should include governance and feedback loops. Domain-specific investigations further suggest that governance must explicitly reflect inter-stakeholder dynamics in regulated firms, so checklists, compliance gates, and handoff templates should be treated as socio-technical artifacts, not just UI rules (Cho et al., 2024). Instrumenting the wizard and component usage (e.g., tracking which templates are chosen most often and collecting post-release UX metrics) enables continuous improvement. Regular cross-functional reviews will surface friction points and allow timely updates to the library and question set. Embedding these practices into quarterly roadmaps ensures the framework remains aligned with organizational priorities and operational realities.

6.2.2 Multi-platform extension

In practice, bankers and relationship managers occasionally use mobile devices for specific, time-sensitive tasks. Typical scenarios include capturing meeting notes immediately after a client visit via voice input while commuting, sending quick client updates, reviewing urgent market alerts, or checking key metrics on the go. These "micro-task" scenarios inform mobile-specific design decisions, such as voice-to-text inputs, concise summaries, one-tap provenance links, and quick-action buttons.

Extending the framework beyond desktop web requires rethinking interaction assumptions, because constraints like screen size, input modality and session length materially change design trade-offs. On mobile, limited screen real estate favors condensed interactions, and the UI should priorities single-task completion with clear escalation paths to richer desktop experiences. Designs must also account for intermittent connectivity and performance: lightweight caching, progressive loading indicators, and conservative use of heavy visualizations will improve perceived responsiveness.

Platform extensions also raise implementation concerns: authentication flows, data permission scopes, and secure local storage differ between mobile and desktop. To manage these, platform-specific adaptation rules should be codified in the design guide. For example: "On devices with screen width < 480px, prefer Inline Snippet + 'View More' link; defer full-page workspace to desktop". LLM-focused literature stresses that model limitations (e.g., hallucination, stale data) require preserving provenance and verification affordances even in compact UIs, therefore the adaptation rules must explicitly require visible provenance and quick-verification actions on mobile as well as desktop (Li et al., 2024). Prototyping on each target platform will surface constraints that inform these rules and prevent a one-size-fits-all application of desktop patterns.

Furthermore, accessibility considerations are particularly crucial across platforms, keyboard navigation and screen-reader compatibility on desktop must map to clear auditory cues and concise language in voice interactions.

6.2.3 Cross-sector applicability

Certain aspects of the framework are not domain-related. Principles such as *Human Control*, *Transparency*, *Clarity*, and *Actionable Feedback* apply across industries. Also, components like prompt inputs, confirm dialogs, and source labels can be reused with minimal changes. This distinction between portable principles and sector-specific instantiations is mirrored in recent case studies: generic human-AI guidelines are valuable starting points but must be translated into domain-specific patterns and governance to be operational in regulated sectors (Cho et al., 2024). However, many elements require tailoring, for example, the nature of "live data," compliance checkpoints, domain vocabularies, and acceptable margins of error differ a lot between finance, healthcare, retail, and manufacturing.

Validating transferability begins with targeted pilot studies. A pragmatic approach is to run small, role-specific pilots in a new sector that mirror the original research methodology: conduct contextual interviews, map workflows, and run the wizard with representative stakeholders. This method follow methodological recommendations from industry surveys and help surface where model-level constraints change design requirements (Li et al., 2024). Pilot outcomes should be evaluated against domain-relevant criteria, such as safety in healthcare, supply-chain latency tolerance in manufacturing, or privacy consent flows in consumer retail. Based on those results, a curated "sector profile" can be added to the design guide that lists recommended components, required compliance checks, and typical container choices for that industry.

Finally, cross-sector scaling benefits from modular governance: keep the core principles and atomic components centralized, but enable "sector packs" that local teams can install and adapt. Practitioner studies of guidebook adoption show that packaging general guidance into scaffolded, organization-specific artifacts (cheat sheets, checklists, starter files) materially improves uptake (Yildirim et al., 2023). This preserves consistency where it matters while enabling necessary specialization.

By combining careful pilot validation with modular packaging and clear sector documentation, the framework can be adapted responsibly to new contexts without losing the design hygiene and trust mechanisms established for the financial domain.

6.3 Future directions

6.3.1 Cross-functional collaboration and governance

Sustained success depends on cross-disciplinary alignment. Regular workshops that bring designers, product owners, engineers, and compliance stakeholders together can use the framework as a common language for evaluating new AI proposals. These sessions should follow a structured agenda: feature framing via the wizard, risk assessment against compliance criteria, and a rapid prototyping sprint to produce a minimal viable interaction using library components.

Governance mechanisms should complement these workshops. A lightweight approval workflow combining automated checks with human sign-offs can ensure that new AI features meet organizational standards for transparency, data use, and user control. Additionally, appointing a "AI steward" role or a small design-ops cell to curate the component library and capture lessons from deployments will maintain quality and foster community ownership.

6.3.2 Longitudinal Adoption Studies

Short, formative evaluations capture immediate usability and perceived usefulness, but the dynamics of trust and dependence on AI evolve over weeks and months. Longitudinal studies should track representative user groups as they adopt AI features in daily operations, measuring metrics such as frequency of use, time saved per task, error rates, reliance tendencies, and self-reported trust and skill retention. These data will clarify whether the framework supports sustained, healthy adoption or unintentionally encourages over-reliance.

Methodologically, a mixed-methods approach will be most informative: automated usage analytics (component selections, regenerate actions, undo frequency) combined with periodic qualitative interviews and diary studies. This combination uncovers not only what changes over time, but why. For instance, rising acceptance may reveal increasing trust due to improved accuracy, or conversely, growing dependency may highlight areas where safeguards or refresher training are needed.

Results from longitudinal work can feed concrete design improvements. For instance adaptive onboarding that fades as users gain expertise, built-in challenge prompts that encourage manual verification, or periodic "competency checks" that ensure critical skills remain in place. These interventions will help prevent skill atrophy while preserving the efficiency gains that AI affords.

6.3.3 Expanding the Prompt Mode Taxonomy

The *"Modes of Prompt"* matrix introduced in Section 3.2.1 offers a useful conceptual axis for categorizing AI interactions by users' goal clarity and input familiarity. A natural next step is to operationalize that matrix into a richer set of interaction templates: predefined prompt patterns, example phrasings, and component combinations tailored to common quadrants (e.g., goal-known/input-unknown or goal-unknown/input-known). These templates should include explicit UI affordances (quick prompts, guided forms, multi-step assistants) and recommended follow-up actions so that designers can compose higher-level flows without inventing prompt logic from scratch.

Furthermore, the taxonomy can be linked to the *Selection Wizard* so that when a designer selects a prompt mode, the wizard automatically proposes component sets, confidence thresholds, and explanation formats suited to that mode. Over time, telemetry about which templates are chosen and how users refine prompts can be collected to iteratively improve template quality and coverage.

Taken together, these future directions form a practical roadmap: institutionalize cross-functional processes, extend to mobile and voice, observe adoption over time, and enrich prompt patterns. Implemented iteratively, these steps will help the framework remain relevant, scalable, and responsibly embedded within operational practice.

6.4 Limitations

The following limitations should be kept in mind when interpreting the applicability and impact of the AI interaction framework. Though the framework was carefully grounded in qualitative research and validated through targeted testing, its current scope is bounded by specific contextual, methodological, and temporal factors. Acknowledging these constraints not only situates the framework's strengths but also highlights areas where further research and adaptation are necessary.

6.4.1 Contextual and cultural specificity

The framework was developed and tested exclusively within Van Lanschot Kempen, a Dutch private bank operating under European financial and data protection regulations. As such, the design solutions reflect cultural norms around data privacy, user interface conventions, and regulatory requirements specific to the Netherlands and the EU's GDPR framework. In markets with different privacy laws or banking practices, such as Asia's varied data jurisdictions, users may have different expectations around consent, data visibility, and compliance workflows.

Moreover, linguistic and cultural factors can influence terminology comprehension and interface metaphors. For example, Dutch bankers might be comfortable with certain phraseologies or iconography, whereas users in other regions could find those same elements unclear or even off-putting. Consequently, the framework's labels, tooltips, and interaction metaphors may require localization and adaptation to maintain clarity and trust across diverse user groups.

Finally, client relationship models differ internationally. In some cultures, personal touch and face-to-face interactions remain paramount, while others lean heavily on digital self-service. Such variations will affect how AI suggestions are received and trusted. Before applying this framework in other banking contexts, a fresh round of user research would be necessary to recalibrate principles, components, and wizard flows to local practices and regulatory environments.

6.4.2 Evaluation sample and environment

The formative evaluation involved a small group of five participants working in a moderated, lab-style setting. While this approach yielded rich qualitative insights, its limited scale may not capture the full spectrum of perspectives across different teams. Larger, more diverse samples could reveal additional usability challenges that did not emerge in the initial tests.

Furthermore, the controlled environment where participants could focus uninterrupted on the wizard differs from real-world conditions characterized by competing deadlines, multitasking, and fluctuating network performance. In an actual production setting, interruptions, or organizational pressures might affect how quickly and confidently users navigate the framework. Field studies or in-situ evaluations would be necessary to validate performance under these more chaotic, day-to-day circumstances.

Finally, the short duration of each session (approximately 30 minutes) limited the ability to observe learning curves over time. Participants' initial impressions and ratings may evolve with repeated exposure, revealing different trust trajectories or workflow integration challenges. Longitudinal studies would help determine whether the framework's benefits persist or require iterative refinement as users become more adept or as organizational processes change.

6.4.3 Evolving AI familiarity and dependency

Artificial intelligence capabilities are advancing at a rapid pace, and users' familiarity and even dependency on AI tools is likely to grow over time. Early in adoption, users may approach AI with caution and curiosity, valuing clear guidance and explicit control. However, as proficiency increases, expectations for automation sophistication, customization, and seamless integration will rise. Features that feel innovative today may be perceived as rudimentary tomorrow, potentially exposing gaps in the framework's ability to support more advanced or speculative AI modalities.

Moreover, long-term reliance on AI introduces new concerns around skill atrophy and over-dependence. In scenarios where AI suggestions become the default decision-maker, users may lose critical domain expertise or fail to question erroneous outputs. The current framework emphasizes human control and transparency as safeguards, but it does not yet address mechanisms for fostering critical reflection or ongoing skill development when AI becomes deeply embedded in workflows. This longitudinal dimension: how user preferences, trust, and competency co-evolve with AI proficiency, lies beyond the scope of the present study but warrants future investigation.

Finally, novel interaction paradigms such as fully autonomous AI agents, adaptive interfaces that learn from user behavior, or AI-mediated collaboration among teams are emerging rapidly. The existing component library and wizard model cover today's most common patterns, but may require significant extension to accommodate these next-generation AI experiences. Continuous monitoring of technological trends and periodic framework updates will be essential to sustain its relevance and effectiveness over time.

7. Conclusion

This project investigated how to design a scalable, human-centered AI interaction framework for internal wealth-management workflows at Van Lanschot Kempen (VLK). The research question centered on how AI features can be integrated into professional tooling so that they reduce operational burden, preserve human judgment, and scale across teams without producing inconsistent or opaque experiences. The outcome is a practical, research-grounded design system intended for the designers and product owners who build AI capabilities, and ultimately for the employees who rely on those capabilities in daily work.

The approach combined qualitative field research with design-led development and formative evaluation. Semi-structured interviews and workflow mapping identified key pain points, *repetitive administrative tasks, fragile trust in black-box outputs, and unclear prompts*, that motivated the design goals. Those empirical findings informed three interlocking deliverables: *a set of Design Principles (Human Control, Transparency, Clarity, Actionable Feedback), an atomic Design Component Library (inputs, chat bubbles, panel formats, templates), and an AI Feature & Component Selection Wizard* that translates research insights into concrete feature specifications. A moderated, task-based evaluation with designers and product owners tested whether the framework supported realistic design work and cross-role collaboration.

Several contributions emerged from the work. Empirically, the study surfaced how trust and usability constraints in financial workflows differ from general consumer contexts: *correctness, provenance, and reversibility are prioritized, and interaction patterns must respect domain-specific compliance and relationship norms.* Design-wise, the framework operationalizes HCI principles, *human-centered AI, explainability, and progressive disclosure*, into reusable components and a decision workflow. The Selection Wizard proved particularly effective at prompting product owners and designers to specify inputs, outputs, and interaction depth before prototyping, addressing a frequent stumbling block where features are otherwise developed with under-specified prompts and assumptions. Formative evaluation indicated broad willingness to adopt the framework: participants found it intuitive, helpful for scoping features, and flexible enough to cover both simple and extensible AI use cases. Role differences were informative, *designers emphasized direct access to components and fidelity in Figma, while product owners valued the wizard's capacity to structure business requirements and spark stakeholder alignment.*

Practically, the framework offers several immediate benefits for Van Lanschot Kempen and similar organizations. It provides consistent ways of implementing AI features, which helps prevent fragmented user experiences across different tools. By embedding transparency and reversible actions at the component level, the framework supports higher interpretability and user control, qualities that are crucial for regulated environments. The Selection Wizard provides a reproducible handoff artifact that can feed design sprints and backlog creation, improving cross-functional communication between product, design, engineering, and compliance teams.

The work also has clear limitations that frame its current applicability. The design and evaluation were situated within a single Dutch private bank; cultural norms, regulatory regimes and domain conventions may differ elsewhere, requiring localization and additional validation. The evaluation cohort was small and composed of internal stakeholders in a lab-style setting; larger scale, in-situ studies are necessary to observe long-term adoption patterns and operational constraints. Finally, AI capabilities and user familiarity are evolving rapidly, changing user expectations and potential dependency on automated suggestions introduce longitudinal risks such as skill atrophy or over-trust that the present study cannot fully resolve.

These limitations suggest a focused agenda for future work. Firstly, cross-sector pilots would help distinguish domain-agnostic principles from industry-specific needs and demonstrate transferability. Secondly, longitudinal adoption studies are needed to track trust, reliance, and proficiency as users incorporate AI into routine tasks. Finally, the Prompt Mode Taxonomy should be expanded into actionable templates and multi-modal pattern packs that the wizard can surface automatically.

In closing, this project demonstrates that design systems can serve as an effective bridge between HCI theory and enterprise practice: design principles grounded in qualitative research can be concretized into components and decision tools that support responsible, scalable AI adoption. For wealth-management organizations seeking to harness generative AI without sacrificing client trust or regulatory compliance, the framework offers a practical starting point. Continued iteration will be necessary to sustain relevance as AI capabilities and workplace norms evolve.

8. References

A

Abedin, B., Meske, C., Junglas, I., Rabhi, F., & Motahari-Nezhad, H. R. (2022). Designing and Managing Human-AI Interactions. Information Systems Frontiers, 24(3), 691–697. https://doi.org/10.1007/s10796-022-10313-1

Addula, S. R., Meduri, K., Nadella, G. S., & Gonaygunta, H. (2024). AI and Blockchain in Finance: Opportunities and Challenges for the Banking Sector. IJARCCE, 13(2). https://doi.org/10.17148/IJARCCE.2024.13231

AI UX Patterns | Filters | ShapeofAI.com. (n.d.). Retrieved July 14, 2025, from https://www.shapeof.ai/patterns/filters

B

Balakrishnan, A. (2024). Leveraging Artificial Intelligence for Enhancing Regulatory Compliance in the Financial Sector (SSRN Scholarly Paper No. 4842699). Social Science Research Network. https://papers.ssrn.com/abstract=4842699

Boustani, N. M. (2021). Artificial intelligence impact on banks clients and employees in an Asian developing country. Journal of Asia Business Studies, 16(2), 267–278. https://doi.org/10.1108/JABS-09-2020-0376

C

Cao, L. (2022). AI in Finance: Challenges, Techniques, and Opportunities. ACM Comput. Surv., 55(3), 64:1-64:38. https://doi.org/10.1145/3502289

Cho, S. H., Jon, S., Jin, Y., Jung, J., & Oh, C. (2024). Understanding the Dynamics in Creating Domain-Specific AI Design Guidelines: A Case Study of a Leading Digital Finance Company in South Korea. Extended Abstracts of the CHI Conference on Human Factors in Computing Systems, 1–7. https://doi.org/10.1145/3613905.3650759

Clarke, V., & Braun, V. (2014). Thematic Analysis. In Encyclopedia of Critical Psychology (pp. 1947–1952). Springer, New York, NY. https://doi.org/10.1007/978-1-4614-5583-7_311

D

Design Systems 101. (n.d.). Nielsen Norman Group. Retrieved May 20, 2025, from https://www.nngroup.com/articles/design-systems-101/

E

Elshan, E., Zierau, N., Engel, C., Janson, A., & Leimeister, J. M. (2022). Understanding the Design Elements Affecting User Acceptance of Intelligent Agents: Past, Present and Future. Information Systems Frontiers, 24(3), 699–730. https://doi.org/10.1007/s10796-021-10230-9

F

Frost, B. (2016). Atomic design. Brad Frost.

G

Gururangan, S., Marasović, A., Swayamdipta, S., Lo, K., Beltagy, I., Downey, D., & Smith, N. A. (2020). Don't Stop Pretraining: Adapt Language Models to Domains and Tasks (No. arXiv:2004.10964). arXiv. https://doi.org/10.48550/arXiv.2004.10964

H

Huang, C., Zhang, Z., Mao, B., & Yao, X. (2023). An Overview of Artificial Intelligence Ethics. IEEE Transactions on Artificial Intelligence, 4(4), 799–819. https://doi.org/10.1109/TAI.2022.3194503

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. ACM Transactions on Information Systems, 43(2), 1–55. https://doi.org/10.1145/3703155

J

Jiang, T., Sun, Z., Fu, S., & Lv, Y. (2024). Human-AI interaction research agenda: A user-centered perspective. Data and Information Management, 8(4), 100078. https://doi.org/10.1016/j.dim.2024.100078

K

Kahn, L. H., Savas, O., Morrison, A., Shaffer, K. A., & Zapata, L. (2020). Modelling Hybrid Human-Artificial Intelligence Cooperation: A Call Center Customer Service Case Study. 2020 IEEE International Conference on Big Data (Big Data), 3072–3075. https://doi.org/10.1109/BigData50022.2020.9377747

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. Journal of Advanced Nursing, 72(12), 2954–2965. https://doi.org/10.1111/jan.13031

L

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks (No. arXiv:2005.11401). arXiv. https://doi.org/10.48550/arXiv.2005.11401

Li, Y., Wang, S., Ding, H., & Chen, H. (2024). Large Language Models in Finance: A Survey (No. arXiv:2311.10723). arXiv. https://doi.org/10.48550/arXiv.2311.10723

M

miglaros. (n.d.). Creating a dynamic UX: Guidance for generative AI applications. Retrieved May 19, 2025, from https://learn.microsoft.com/en-us/microsoft-cloud/dev/copilot/isv/ux-guidance

Moser, C. A. (1952). Quota Sampling. Journal of the Royal Statistical Society. Series A (General), 115(3), 411–423. https://doi.org/10.2307/2980740

P

Park, S., Wang, A., Kawas, B., Liao, Q. V., Piorkowski, D., & Danilevsky, M. (2021). Facilitating Knowledge Sharing from Domain Experts to Data Scientists for Building NLP Models. 26th International Conference on Intelligent User Interfaces, 585–596. https://doi.org/10.1145/3397481.3450637

People + AI Guidebook. (n.d.). Retrieved 19 August 2025, from https://pair.withgoogle.com/guidebook

P, G., S, S., P, S., & M, N. (2023). ANALYSING THE MAJOR DETERMINANTS OF IMPLEMENTING ARTIFICIAL INTELLIGENCE IN ENHANCING CUSTOMER RELATIONSHIP MANAGEMENT IN PRIVATE BANKING COMPANIES. Russian Law Journal, 11(7S), Article 7S.

R

Rukmana, D. (2014). Quota Sampling. In Encyclopedia of Quality of Life and Well-Being Research (pp. 5382–5384). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0753-5_2393

S

Sanders, E., & Stappers, P. (2012). Convivial toolbox: Generative research for the front end of design. Bis.

Sharma, G., Solanki, U., Solanki, V., & Baker, M. R. (2024). Impact of Artificial Intelligence in Personalized Banking Service. In Handbook of Artificial Intelligence Applications for Industrial Sustainability. CRC Press.

Shetty, S. K., Spulbar, C., Birau, R., & Filip, R. D. (2022). Impact of Artificial Intelligence in Banking Sector with Reference to Private Banks in India.

T

The Shape of AI | UX Patterns for Artificial Intelligence Design. (n.d.). Retrieved July 12, 2025, from https://www.shapeof.ai

W

Walker, B. (2024, March 11). Microsoft Unveils AI-Driven Copilot for Finance to Boost Productivity. B2Bdaily.Com. https://b2bdaily.com/fintech/microsoft-unveils-ai-driven-copilot-for-finance-to-boost-productivity/

Y

Yildirim, N., Pushkarna, M., Goyal, N., Wattenberg, M., & Viégas, F. (2023). Investigating How Practitioners Use Human-AI Guidelines: A Case Study on the People + AI Guidebook. Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 1–13. https://doi.org/10.1145/3544548.3580900

TU Delft

VAN LANSCHOT KEMPEN