In Artificial Intelligence (AI) We (Dis)Trust? Navigating Institutional Pressures for Automation and Augmentation in the Implementation of AI in Organizations

Abstract

AI brings competing demands to organizations from pressures towards efficiency and standardization versus contextual responsiveness and ethical judgment. These demands become particularly salient when some areas of organizations push for automation while others for augmentation, as two distinct paradigms of AI implementation. It is therefore important to understand how organizations navigate these competing institutional pressures. Drawing on a nested case study of a European airline, we find that the choice between automation and augmentation is not solely a matter of task–technology fit. It is instead also shaped by how actors configure trust and distrust in AI systems in response to two coexisting institutional logics: instrumental–analytic and contextual–normative. We show how these two logics stimulate different trust–distrust configurations, which in turn guide how AI is implemented and adopted within organizations. We identify two reconciliation practices that help organizational actors manage inherent tensions between these competing institutional pressures: *mindful evaluation* and *proactive safeguarding*. The research reveals how AI implementation and adoption reflects conflicts between dominant institutional logics and contributes with a novel perspective on the role of institutional logics and trust in projects of AI implementation.

Keywords

artificial intelligence, automation, augmentation, trust, distrust, institutional pluralism, case study

1. Introduction

Artificial intelligence (AI) is integrating and gradually reconfiguring organizational decision-making across industries and business functions such as in dynamic pricing, loan approvals and supply-chain optimization (Constantiou et al., 2024; Shollo & Galliers, 2024; Shrestha et al., 2019). While some take a utopian view of this AI driven future others have concerns and are more pessimistic (Carroll et al 2024). Regardless, we see businesses pushing ahead with vast investments in AI with recent reports showing that over 55% of firms already use AI in at least one function, and many organizations reporting substantial gains with up to 20% of EBIT growth due to AI initiatives (Maslej et al., 2025). Despite the exponential investments in AI many implementations fail or fall short of their aims hindered by misaligned expectations and organizational resistance (Benbya et al., 2021; Glikson & Woolley, 2020; Huysman, 2025; Mayer et al., 2024) often owing to poor recognition of the wide and deep effects of AI in reconfiguring boundaries and agency between human and machine decision-making (Baptista et al., 2020).

Previous research has shown that the implementation and adoption of AI within organizations is framed in terms of two paradigms: the paradigm of *automation* and the paradigm of *augmentation* (Langer & Landers, 2021; Raisch & Krakowski, 2021; Teodorescu et al., 2021). *Automation* entails the delegation of tasks, especially routine and data-intensive ones, to AI, whereas *augmentation* emphasizes collaboration between humans and AI for complex and context sensitive decision-making. Implementing AI introduces technologies that are characteristically more autonomous, adaptive and opaque (Berente et al., 2021); however, each paradigm implies different levels of control, accountability and discretion in jointly distributed human–AI decision-making (Hayes, 2008). As a result, AI implementation becomes a site of organizational contradiction, marked by competing imperatives such as speed vs. judgment, standardization vs. discretion and optimization vs. responsibility. As these competing paradigms unfold, organizations face mounting pressure to balance efficiency against contextual judgment in their AI strategies, making selective decisions about which activities to delegate to AI and which to retain as human- or joint human-AI work (Lei & Kim, 2024). However, while the distinction between automation and augmentation has been established (Baer et al., 2025; Raisch & Krakowski, 2021), contemporary research typically frames this distinction as a matter of task–technology fit and work design (Grønsund & Aanestad, 2020; Lebovitz et al., 2021). In doing so, it often treats automation and augmentation as relatively independent design choices, and with little attention to the underlying values, assumptions and belief systems that drive and make one paradigm emerge as stronger than the other in different organizational and social contexts (Raisch & Fomina, 2025; Shao et al., 2024). Consequently, such accounts do not fully capture how broader social and organizational environments, and their associated normative expectations, generate tensions between automation and augmentation and ultimately shape AI implementation and adoption. This leaves us with a limited understanding of why ostensibly similar AI systems are implemented in different ways across organizations, and how organizations experience, navigate and reconcile the competing demands of automation and augmentation in practice.

An institutional logics perspective provides a suitable lens for analyzing tensions between automation and augmentation by linking local and situated choices about AI implementation with the broader organizational, cultural and societal contexts in which those choices are made. Institutional logics are socially constructed belief systems and associated material practices that provide meaning and guide behavior within organizational fields (Friedland & Alford, 1991; Thornton et al., 2012). It is however important to note that AI implementation unfolds within a context of institutional pluralism, where logics emphasizing efficiency and scalability, such as market and corporate logics, coexist and may conflict with those that privilege human judgment, discretion and ethical responsibility, such as profession and community logics (Greenwood et al., 2011). This institutional logics perspective has been used in information systems (IS) research to show how technology can embody and transform institutional arrangements, particularly by examining organizational responses to multiple, coexisting institutional logics (Berente & Yoo, 2012; Faik, et al., 2020; Faik et al., 2026; Mignerat & Rivard, 2009). Yet this work has largely focused on earlier generations of IT. We still know little about how the distinctive properties of contemporary AI—such as autonomy, learning capacity and opacity—interact with this institutional pluralism to generate specific pressures and dynamics, and how these dynamics influence organizational interpretations of AI's legitimacy, risks and benefits (Huysman, 2025).

Institutional logics influence AI implementation indirectly and explain how various actors within organizations develop trust and distrust in AI (Guo et al., 2017; Lumineau, 2017). While *trust* reflects confidence in AI's capabilities (e.g., data processing, pattern recognition), *distrust* signals caution about its limitations (e.g., opacity, lack of contextualization). These orientations often interact, as actors may trust AI for computational tasks yet distrust their capacity for ethical or situational reasoning. From an institutional logics perspective, the trust/distrust orientations are not merely idiosyncratic individual level psychological traits, but instead embedded in patterns and taken-for-granted assumptions about the role of technology, expertise and accountability in organizational life (Slavova & Karanasios, 2018). The configuration of trust and distrust in AI therefore also reflects potential inherent conflicts of institutional logics (Currie & Guah, 2007) which ultimately shape whether organizations lean toward automation or augmentation approaches in AI implementation in organizations.

Despite their fundamental role and a growing interest in how trust and distrust shape AI adoption (Scharowski et al., 2025), we still know little about how they function as mechanisms through which actors respond to competing logics in the institutional environment. Existing studies often treat trust and distrust as psychological or mostly technical assessments (Glikson & Woolley, 2020) overlooking their embeddedness in institutional environments. Our study responds to this limitation by examining how organizations navigate institutional conflicts for automation and augmentation in projects of AI implementation through the configuration of trust and distrust. Accordingly, we pose the following research question:

How do competing institutional pressures for automation and augmentation shape configurations of trust and distrust in AI during AI implementation in organizations?

To address this question we follow a qualitative case study approach of a European airline that has implemented AI across three critical operational areas of the business. Our findings show that institutional pluralism becomes consequential when actors interpret AI through either *an instrumental–analytic logic*, focusing on calculative rationality, internal coherence and replicable performance; or *a contextual–normative logic*, foregrounding practical appropriateness, situational responsiveness and ethical accountability. These interpretations give rise to distinct configurations of trust and distrust, which, in turn, shape different implementation pathways. Furthermore, we identify two practices that reconcile the navigation of competing AI implementation paradigms: (1) *mindful evaluation* reinforces the cognitive capacity for critical engagement, enabling actors to assess when AI aligns with human expertise; and (2) *proactive safeguarding* reinforces the behavioral structures that ensure AI stays aligned with operational realities.

This study contributes to research on AI implementation in organizations and to institutional theory informed IS scholarship in three ways. First, it reveals how trust and distrust in AI are not merely individual attitudes or technical judgments (Omrani et al., 2022), but institutionally grounded mechanisms through which organizations respond to higher level pluralistic demands in the social order. Second, it further theorizes automation and augmentation as institutionally influenced implementation paradigms (Avgerou, 2001), shaped by how actors negotiate alignment between AI capabilities and institutional expectations. Third, it identifies organizational practices that mediate these dynamics, offering insight into how firms can pragmatically manage AI implementation while preserving legitimacy and strategic flexibility (Grønsund & Aanestad, 2020).

The paper proceeds with a section developing the conceptual foundation for this study based on an expanded view of pluralistic institutional dynamics in organizations and of trust and distrust in organizational responses to AI. We then provide a description of the methods and methodological approach used to capture these effects and our process of data collection and analysis. The next section presents our key findings, organized around the emergent configurations of trust and distrust and their corresponding AI implementation pathways, as shaped by two overarching evaluative institutional logics. Finally, we present our theoretical contributions and conclude with implications for research on AI implementation and institutional theory.

2. Conceptual Foundations

To understand how organizations navigate the competing institutional pressures in the implementation of AI in organizations we draw on three interrelated conceptual foundations. We begin by developing the concepts of automation and augmentation as paradigms of AI implementation. We then situate them within the broader context of institutional pluralism. Finally we link and show the connection between institutional pluralism and mid-range concepts of trust and distrust in AI within organizations.

2.1 Automation and Augmentation as Distinct Paradigms of AI Implementation

We define AI in line with Berente et al. (2021) as a fluid and evolving technological frontier that enables machines to perform tasks that previously required human cognition, such as classification, prediction and optimization. In our empirical setting, AI refers to machine learning–based systems implemented to support high-stakes operational decision-making within various areas of organizations.

Automation and augmentation represent two distinct AI implementation paradigms[1] (Raisch & Krakowski, 2021; Faraj et al., 2018). *Automation* involves delegating tasks or decisions entirely to AI systems to improve efficiency, scalability and standardization (Raisch & Krakowski, 2021; Shao et al., 2024). On the other hand, *augmentation* positions AI as a collaborator that enhances rather than replaces human capabilities, emphasizing contextual reasoning, discretion and human–AI co-performance (Faraj et al., 2018; Raisch & Fomina, 2025). In practice, however, organizational deployments of AI seldom conform to either paradigm in a pure form. Some tasks or decision points are tightly automated, while others are deliberately organized as human–AI collaboration, producing hybrid approaches of AI use across processes and domains (Agrawal et al., 2023).

Although prior research has examined automation and augmentation as AI implementation paradigms, the differences between them are typically framed in functional or technical terms (Baer et al., 2025). Specifically, the literature often explains the choice between automation and augmentation through factors such as task routineness, system transparency or user cognition (Agrawal et al., 2023; Berente et al., 2021; Lebovitz et al., 2022). These perspectives offer useful insights into how individuals and units within organizations assess AI implementation, but they tend to overlook the broader organizational environments and normative expectations that shape such assessments.

Further, a closer look at the literature reveals three limitations in how automation and augmentation are conceptualized (we provide a summary of this in Appendix A). First, they are often treated as isolated design choices, with little attention to the underlying values or

[1] We define AI implementation paradigms as organizationally patterned approaches to enacting AI that reflect broader interpretive, strategic and operational commitments.

assumptions about the roles of technology and humans operating and making decisions within social settings (Raisch & Fomina, 2025; Shao et al., 2024). Second, studies are fragmented across levels. Micro-level work emphasizes individual trust or learning (e.g., Glikson & Woolley, 2020), while macro-level research focuses on structural trends like labor substitution or regulation (e.g., Agrawal et al., 2023). This means that the organizational level, where implementation paradigms are negotiated in practice, remains underexplored (Berente et al., 2021; Lei & Kim, 2024). Third, although some studies raise concerns about governance, fairness or ethics (e.g., Berente et al., 2021; Heyder et al., 2023), these are often treated as background issues rather than central to how AI implementation is interpreted and enacted in practice. Therefore, the underlying tensions and legitimacy pathways between automation and augmentation paradigms call for greater attention to the cultural and normative structures that shape such paradigms, and in conceptualizing how they are interpreted and enacted within organizations.

2.2 Institutional Pluralism in projects of AI Implementation

Institutional logics explain how cultural belief systems and normative structures shape organizational cognition, meaning-making and material practices in social settings (Friedland & Alford, 1991; Thornton et al., 2012). They are historically situated, socially constructed frameworks that guide what actors perceive as legitimate goals, appropriate conduct and credible knowledge within organizational life (Thornton & Ocasio, 2008). Although early institutional theory emphasized processes of convergence and isomorphism (DiMaggio & Powell, 1983; Meyer & Rowan, 1977), later work has drawn attention to *institutional pluralism* which recognizes the simultaneous presence of multiple and often conflicting logics within and across organizations (Greenwood et al., 2011; Kraatz & Block, 2008).

A central step in this development was Friedland and Alford's (1991) seminal article, in which they conceived society as a system of multiple institutional orders, each organized around a defining central logic. These institutional orders demarcate different spheres of social life by providing distinct symbolic systems and ways of ordering reality that render actors' experiences of time, space and meaning intelligible (Friedland & Alford, 1991). Examples of such orders include the market, corporation, professions and community, each associated with a central dominant logic. The market logic centers on efficiency, competition and calculability; the corporate logic stresses hierarchical coordination, managerial control, standardization and performance monitoring; the profession logic emphasizes expertise, discretion and context-sensitive judgment; and the community logic foregrounds shared values, mutual obligation and the preservation of collective cohesion and reputation (Dunn & Jones, 2010; Reay & Hinings, 2009; Thornton et al., 2012).

Importantly, while these logics define an institutional order at the societal level they are also reproduced, specified and recombined at field and organizational levels, either as narrowly focused variants, for instance, a service-oriented commercial logic elaborating the broader market logic) or as hybrids, for example, a social-entrepreneurship logic that weaves together market and community logics (Faik et al., 2020; Smets et al., 2015). This multilayered view of institutional orders and their local enactments opens analytical space for examining how organizations develop composite, situated logics that are grounded in broader institutional orders yet adapted to local conditions - a perspective that we take in our analysis.

Against this backdrop, organizations situated within such pluralistic environments must navigate not only structural contradictions but also competing cultural expectations about what constitutes legitimate goals, appropriate practices and acceptable trade-offs. Institutional pluralism becomes particularly salient in periods of technological change (Seidel et al. 2025), such as the integration of AI in various areas of organizations. AI introduces capabilities marked by autonomy, malleability and opacity (Berente et al., 2021) which destabilize existing normative and epistemic anchors. For example, the capacity of AI to generate probabilistic, non-transparent outputs challenges professional expectations for reasoned judgment and accountability (Faraj et al., 2018). In such contexts, plural logics become enacted and amplified. For example, while a corporate logic will view AI as a means of control and standardization (Marett et al., 2013), a profession logic will resist opaque automation that displaces human expertise (Hultin & Mähring, 2014). Further, while a market logic may valorize revenue, optimization and speed (Martinsons, 2008), a community logic will foreground concerns over fairness, inclusion and responsibility (Miranda et al., 2015).

Importantly, institutional pluralism is not uniformly experienced across the organization (Berente & Yoo, 2012). It is interpreted and enacted in situated ways by organizational members engaging with specific technologies and decisions (Faik et al., 2020). In the case of AI, the implementation paradigms of automation and augmentation may be received differently across organizational units depending on how institutional logics are prioritized and interpreted. For instance, the same AI system may be embraced in one setting as a cost-saving innovation aligned with corporate or market expectations, while in another, it may raise concerns about diminished human discretion or professional accountability. These differences will deepen conflicts over what constitutes legitimate and responsible use of AI. Such conflicts often manifest in the form of coexisting attitudes which shape the way trust and distrust in AI is formed by different groups. This, in turn, shapes how individuals and groups evaluate and respond to AI within these pluralistic institutional environments.

2.3 Trust and Distrust Toward AI

Trust and distrust are distinct and coexisting attitudes that shape how organizations respond to technological change (Kostis et al., 2022). *Trust* involves a sense of hope, encompassing a confident expectation that the other party (or system) will act in beneficial and predictable ways (Mayer et al., 1995). Trust is often accompanied by a willingness to rely on that party despite uncertainty, based on the belief that they will support the trusting party’s objectives or responsibilities (Mayer et al., 1995). On the other hand, *distrust* reflects a sense of fear - a belief that the other party may act in harmful, unpredictable or misaligned ways (Lumineau, 2017). Distrust encourages one party to guard against the risks posed by the other's potential for undesirable actions (Luhmann, 2018).

However, actors can simultaneously trust and distrust the same entity, particularly under conditions of uncertainty or competing demands (Saunders et al., 2014). For example, organizational actors may trust AI to efficiently process large datasets, yet simultaneously, distrust it in ethically sensitive contexts that require human discretion (Glikson & Woolley, 2020; Tang et al., 2023). Such ambivalence reflects broader patterns of attitudinal coexistence, especially when new technologies disrupt established roles, norms, or expectations (van Harreveld et al., 2009; Katz & Hass, 1988). Importantly, as prior research has shown, increased distrust does not necessarily reduce trust, nor does the presence of trust eliminate distrust - rather, both can emerge and evolve independently (Dimoka, 2010; Lewicki et al., 1998). This conceptual distinction allows for more nuanced interpretations of organizational responses to AI under conditions of uncertainty and normative complexity.

In the context of AI, trust and distrust manifest as domain-specific judgments, shaped by perceptions of system functionality, transparency and goal alignment (Kostis et al., 2022; Lankton et al., 2015). *Trust in AI* entails confident positive expectations that the system will perform tasks reliably, accurately and in accordance with intended organizational objectives (Glikson & Woolley, 2020). *Distrust in AI* arises from confident negative expectations that AI may malfunction, produce harmful outcomes, misalign with organizational values and goals (Scharowski et al., 2025). These attitudes become especially salient in contexts where AI's malleability and opacity amplify concerns around accountability, discretion and ethical integrity (Berente et al., 2021; Raisch & Krakowski, 2021).

Although traditional perspectives (e.g., Mayer et al., 1995; Schoorman et al., 2007) emphasize the link between traits such as benevolence and integrity to interpersonal trust, AI-specific trust and distrust are more closely associated with system-level attributes such as explainability, controllability and context-awareness (Berente et al., 2021; Tams et al., 2018). However, as previously mentioned, even recent perspectives often frame trust and distrust as individual or technical assessments (Saunders et al., 2014) and therefore overlook the embeddedness of these constructs in the social context and in the institutional environment of organizations.

This is however an important oversight because what is considered trustworthy or untrustworthy in organizational practice is not just a matter of interface design or technical robustness, it is also shaped by institutional logics (Fuglsang & Jagd, 2015). These historical belief systems, such as professional values, corporate priorities or community expectations, influence how actors interpret the legitimacy, risk and appropriateness of AI (Greenwood et al., 2011; Karunakaran et al., 2022). Furthermore, while trust is often assumed to enable automation, and distrust to obstruct it (Starke & Lenca, 2024) such binary framings obscure how trust and distrust interplay and coexist in guiding AI implementation and adoption choices, sometimes reinforcing one another, and sometimes acting in tension, especially in environments marked by institutional pluralism.

3. Methods

We employed an exploratory qualitative research design to examine how organizations navigate AI implementation amid institutional plurality and dynamics of trust and distrust. This approach is well suited to capturing complex, context-dependent phenomena and uncovering emergent patterns that contribute to theory development (Yin, 2014). More specifically, we conducted an in-depth nested case study[2], which allowed to investigate intra-organizational variation in how AI systems are interpreted, trusted, distrusted, and enacted. This approach is particularly appropriate for studying the complexity of AI implementation, given its cross-functional embeddedness and its implications for diverse user groups and institutional expectations (Berente et al., 2021; Glikson & Woolley, 2020). By tracing how institutional logics are instantiated in different business units, we identify patterned similarities and differences in how trust and distrust shape AI implementation. These nested insights and cross case analysis allowed us to theorize the organizational dynamics through which broader institutional pluralism is enacted, mitigated or reinforced in practice.

3.1. Case Selection and Data Collection

Our study is based on extensive engagement with a European airline, referred to here as *BlueSky*. In response to a strategic mandate to boost operational efficiency and resilience, BlueSky launched a multi-year AI transformation initiative centered on adopting machine-learning-based tools in various back-office and non-client-facing functions. These AI systems were implemented in areas such as route planning, crew scheduling, cargo management, dynamic pricing and procurement. Our empirical study focused on three departments where AI was adopted: Customer Retention Management (CRM), Revenue Management (RM) and the Data Science Lab (DSL). Across these units, AI systems were deployed to analyze behavioral data, optimize pricing strategies and to handle large-scale operational data.

We selected BlueSky for two primary reasons. First, the airline industry is among the early adopters of AI and a leader in digital transformation (Kell, 2023). Airlines are knowledge-intensive organizations that rely heavily on timely and accurate decision-making based on

---

[2] A nested case study in qualitative research is a methodological approach that involves examining multiple sub-units or cases within a larger, overarching case (Yin, 2014). This approach allows researchers to analyze complex phenomena at different levels within a single organization or context.

complex and high-volume data. AI is therefore viewed as a critical enabler of performance optimization in such data-rich environments. Second, airlines, like other knowledge-intensive organizations, face persistent dilemmas around when and how to intervene in AI-supported decisions (Cheatham et al., 2019; McKendrick & Thurai, 2022). Too much human intervention may undercut efficiency gains, whereas too little may raise issues of accountability and risk.

We adopted a multi-method data collection strategy combining semi-structured interviews, document analysis and field observations to ensure triangulation (Jick, 1979). One author previously worked at BlueSky and possessed valuable insider knowledge about the company and its strategy, culture and high-level AI implementation efforts. This background proved beneficial in preparing for interviews and conducting direct field observations. We conducted 21 semi-structured interviews across CRM, RM and DSL. Interviews lasted between 35 and 78 minutes (average: 46 minutes) and covered a range of themes, including participants' perceptions of AI within their work context, the degree and nature of trust and distrust they associated with AI systems, and how these attitudes shaped their adoption experiences and decision-making. While structured around key focal topics, the interviews remained open-ended and adaptive to emergent themes and unanticipated insights. Appendix B provides further detail on the interview protocol and participant backgrounds.

To enrich and triangulate the interview data, we collected documents such as strategy reports, project memos, presentation slides and meeting notes. These materials offered insight into the institutional and strategic framing of AI at BlueSky. Additionally, one author conducted a week-long field visit to BlueSky's headquarters, during which they attended presentations, observed AI demonstrations and joined brainstorming sessions. These interactions provided additional opportunities to contextualize interview responses and deepen our understanding of how trust, distrust and institutional logics were enacted in practice. Table 1 summarizes our data collection process.

Table 1. Overview of Data Collection

Source	Data Collected	Key Items Investigated
Semi-structured interviews	21 interviews with practitioners (e.g., heads of department, data analysts, and data scientists) across different business units (e.g., revenue management, customer retention management, and data science lab)	Perceptions of AI's organizational role; expectations of AI-driven business operations; levels of trust and distrust in AI; impact of AI on business operations; approaches to AI utilization; mechanisms supporting AI implementation
Documents	Five BlueSky annual reports (2018–2022; avg. 204 pages), five BlueSky sustainable development reports (2018–2022; avg. 97 pages), two third-party case studies of BlueSky (avg 38 pages), and 20 regional aviation (Eurocontrol) think papers (avg 12 pages)	Necessity of AI in the organization; institutional and environmental factors enabling or inhibiting AI implementation; organizational culture and values surrounding AI; anticipated effectiveness of AI in operational processes
Field observations	48 hours of direct participation (one author) in brainstorming sessions, project presentations, operational demonstrations, and unstructured conversations	Practitioners' interactions with AI in daily operations; attitudes of (dis)trust toward AI as reflected in everyday tasks; challenges to trusting AI in decision-making

3.2. Data Analysis

Our data analysis followed an interactive process in line with established qualitative data analysis practices. First, we engaged in open coding to thoroughly examine the data, identifying categories and distilling them into codes (Miles et al., 2014). We labeled these codes with descriptive phrases that captured recurring meanings across the data. For instance, we coded situations where BlueSky favored AI-generated pricing insights over traditional human decision-making as "Valuing AI-generated pricing insights over human intuition." This iterative process continued until all authors reached a consensus on the emerging data structure.

Second, we conducted axial and thematic coding (Corbin & Strauss, 2008), reorganizing the open codes into more abstract themes based on their similarities and differences. For example, we grouped open codes related to BlueSky's institutional culture of optimizing revenues through AI under the axial code "Revenue Maximization". The axial codes were then clustered into broader themes, uncovering underlying conceptual structures. At this stage, we began to observe cross-case regularities in how actors evaluated AI that pointed to deeper, shared assumptions about appropriate goals and responsibilities - for instance, recurring tensions between accuracy and fairness, efficiency and judgment, and standardization and contextual fit, as well as patterned linkages between these tensions and actors' role understandings and accountability expectations. Rather than imposing institutional logics as an a priori lens, we treated these inductively derived patterns as prompts for abductive engagement with the institutional logics literature, using it as a framework to refine and label our emerging categories (Thornton et al., 2012; Timmermans & Tavory, 2012). This process enabled us to identify nuanced conceptual dimensions spanning institutional plurality, as expressed in the instrumental–analytic and contextual–normative logics we identified, the configurations of trust and distrust in AI and the two established AI implementation paradigms of automation and augmentation. In this sense, the empirical patterns in principles, goals and identities emerged from the data, while our articulation of them as organizational instantiations of broader institutional logics reflects the theoretical lens we adopted at a later stage of analysis. Appendix C provides an illustration of our coding structure.

In addition to the structured coding process, we adopted an iterative approach, cycling between the data, emerging themes and relevant literature. This approach, grounded in the constant comparative method (Strauss & Corbin, 1998), enabled us to refine our understanding of the relationship among institutional plurality, trust and distrust in AI, and automation and augmentation paradigms. We continually revisited the data to ensure that our interpretations were robust and reflective of the complexities inherent in AI implementation. To further enhance the rigor of our analysis, we conducted multiple rounds of peer debriefing, where all authors critically evaluated the emerging themes and their interrelationships. This collaborative process provided a system of checks and balances, minimizing individual biases and strengthening the credibility of our findings (Morse et al., 2002).

Furthermore, we drew on Langley's (1999) narrative strategy to select and present evidence that illustrates the theoretical relationships uncovered during our analysis. This narrative approach was essential for constructing a coherent theorization that traces how multiple institutional logics were instantiated in practice, how configurations of trust and distrust in AI formed over time, and how these were linked to AI implementation choices across departments, while also highlighting the dynamics of these relationships. Throughout the data analysis, we remained attentive to key concepts from our literature review and conceptual framing, ensuring that our findings accurately reflected the embedded nature of AI implementation within the broader discourse of institutional logics. Our methodological approach is summarized in Appendix D.

4. Findings

We now present findings from the three nested cases, organizing the empirical material into three subsections for each of the three case studies. For each case, we first outline the organizational purpose of the AI project implementation. We then examine the institutional pressures that shaped varying emergent configurations of trust and distrust in AI. Finally, we discuss the emergent practices that helped navigate inherent conflicts in the institutional environment. In the concluding subsection, we synthesize insights across the three cases to identify cross-cutting patterns that inform our theorization, which we then further develop in the section 5 (Theorization).

4.1. Nested Case One: AI Implementation in the CRM Department

The CRM (Customer Relationship Management) department was responsible for leveraging customer data to support commercial decisions, particularly in designing promotional offers and refining marketing campaigns. The CRM analysts, who were proficient in data analytics and statistical tools, were tasked with supporting telemarketing campaigns launched by the marketing department. One of these campaigns was aimed to promote a co-branded credit card, developed with a major bank, offering exclusive perks to loyal customers. Traditionally, such campaigns relied on simplistic rules or subjective judgment, such as targeting frequent flyers who had completed more than 20 trips. As the head of CRM analysts noted:

In the past, we would call customers considered frequent flyers. Then, who is a frequent flyer? Someone with over 20 trips. And how many customers have more than 20 trips? 10,000. Okay, let's call them [frequent flyers].

However, this approach overlooked the complexity and diversity of the airline’s customer base, often resulting in poor conversion rates and unnecessary costs. As the head analyst further explained:

Most of the rules used in previous campaigns were simple heuristics, and some, even worse, speculations. Heuristics can be useful, but you cannot rely on such simplistic views for designing modern campaigns.

To overcome these limitations, the CRM analysts applied their expertise in data analytics and machine learning to develop a predictive AI model that estimated the likelihood of a positive response for each customer, enabling more targeted and personalized messaging.

Most CRM analysts approached their work through a scientific lens, reflecting their academic training in computer science or operations research. Several held PhDs and were drawn to the airline industry because of its strong reputation for scientific rigor and data-driven innovation. In addressing problems, they consistently turned to academic literature, placing a premium on methodological precision and controlled experimentation. As a result, they prioritized analytical rigor, often emphasizing empirical testing over practical input. Notably, many deliberately avoided collaboration with other business units, believing that external input might compromise the integrity of their scientific process. As the head of CRM analysts explained:

We are somewhat isolated from the business side, and to be honest, most of us prefer not to be involved in [collaborating with other units]. It's more a matter of character or a philosophical view of science. We often distance ourselves from [collaboration] to maintain our academic perspective and scientific integrity.

Against this backdrop, the CRM analysts expressed strong confidence in AI's ability to process complex multidimensional data - capabilities they believed far exceeded human cognitive limits. They saw AI as particularly adept at uncovering intricate patterns, correlations and subtle relationships that would likely escape human analysis. As one senior analyst noted:

For any task involving big data, AI is essential. The human brain simply cannot handle it. We can't make predictions as accurate and complex as those generated by an AI system. So, when it comes to the volume and complexity of data, we rely on AI for support.

Similarly, they believed that AI not only offered clear advantages over prior manual processes but also entailed minimal risk of harm. As the head of CRM analytics noted:

For very long time we have done it ourselves and so we knew the data, we knew the errors, we knew the anomalies. We can anticipate, we know what to expect.

In this case, the CRM analysts worked within an environment that deeply valued scientific objectivity, which shaped how they approached decision-making. They believed that less automated processes were prone to subjective judgments and heuristics. Confident that AI could deliver more reliable outcomes with minimal risk, they deliberately bypassed collaboration with the marketing department, preferring to rely on the mathematical precision of AI. They deliberately minimized human input to produce decisions rooted in rigorous, data-driven analysis rather than intuition or experiential bias. As the head analyst explained, "*We avoid interventions to mitigate human bias in the process. If I intervene, I'm just finding a more 'scientific' way to apply the business rules I've already assumed.*"

However, when the CRM analysts presented the AI-generated customer clusters to the marketing team, they faced difficulties in conveying the practical relevance of the outputs. The clusters which were produced through an opaque process lacked clear connections to real-world factors, making them difficult to interpret. As one analyst admitted, "*In some cases, we created clusters, and we had no idea what they represented. I couldn't relate them to the real world.*" As a result, the marketing team struggled to make use of these algorithmic outputs, largely because they couldn't understand how the AI had arrived at them. The same analyst elaborated:

> *When we told marketing that the elbow test indicated we needed 11 clusters, they didn't understand and didn't care to understand because it wasn't part of their job... They just needed a result they could make sense of.*

The difficulty in translating AI outputs into actionable strategies revealed a key limitation: although AI could process large datasets and identify robust patterns, it often lacked contextual information needed for practical use. Although CRM analysts remained confident in AI's technical strengths, they acknowledged the risks of misalignment when outputs failed to meet marketing's real-world needs. Their commitment to scientific objectivity paired with limited cross-unit collaboration amplified this disconnect. The marketing team's dissatisfaction stemmed not from opposition to AI, but from the lack of practical relevance of the outputs of the AI system. This highlighted the need to complement AI's technical precision with contextual understanding to ensure that the outputs are meaningful and actionable in real-world settings. In response, the analysts re-ran the model with marketing input and initiated a pilot study to ensure alignment with operational goals. As one analyst explained:

> *We created a sample, and we immediately incorporated business ideas to improve it. It's not a random sample of all our customers like what the AI will do. It's a sample of customers that our business thinking suggests have a non-zero probability [of responding]. For instance, we're not going to include someone who lives in*

Canada and doesn't have a banking relationship with us. This is where business thinking comes in.

Furthermore, the CRM analysts invested considerable effort in the post-analysis phase, conducting descriptive analyses to contextualize the AI-generated clusters. This effort was to ensure that the outputs, while grounded in rigorous scientific methods, were also translated into actionable strategies relevant to the marketing domain, as expressed by one analyst:

I took the clusters and conducted descriptive analysis to understand what these variables [used by AI] mean and how they emerged. I'll ask the marketing team how many personality types they want to have. I won't adhere rigidly to statistical metrics as if they're the Holy Book, or as if I were a robot. I should be able to understand how the phenomenon I'm studying relates to the real world. This requires commercial perception.

The misalignment between AI's outputs and BlueSky's operational needs revealed a deeper conflict between two coexisting deeply rooted orientations. While scientific objectivity remained the dominant lens among CRM analysts driving confidence in AI's analytical precision, this orientation clashed with the marketing team's emphasis for contextual applicability and strategic alignment. As a result, business pragmatism emerged not just as a practical adjustment, but as a countervailing force challenging the sufficiency of data-driven logic. The sustained belief in AI's technical rigor allowed the initiative to progress, yet the rise of concerns about its practical fit created pressure to reintroduce human oversight and contextual judgment.

Importantly, the CRM analysts did not view AI implementation as a linear process in which human involvement simply fills the gaps left by technology. Instead, they approached it as a dynamic and iterative endeavor: one that required ongoing judgment to align algorithmic capabilities with practical business needs. This approach became evident in two key shifts.

First, analysts developed a sustained attentiveness to commercial applicability. While confident in AI's technical sophistication, they remained alert to the possibility that even robust analytical outputs might fail to generate usable customer insights. As one CRM analyst emphasized:

> *This work is not done to satisfy theoretical satisfaction… It is done to increase revenues by enabling better, more targeted services. And that happens only when those designing and using these tools truly understand our business needs, who our customers are.*

Second, the CRM team formalized routine coordination with the marketing department. Moving away from earlier patterns of isolated decision-making, analysts proactively engaged stakeholders to ensure AI-generated segments could be interpreted and applied effectively. This shift reversed the traditional flow of analytics from data-first to use-case-first; and introduced a layer of collaborative filtering to ensure contextual fit. As one analyst reflected:

> *Before anything, we go to marketing to understand their needs because in the end that's what really counts. The data scientist doesn't just present results; the process has to go in the opposite direction. We ask, 'What are you comfortable working with? How many clusters can you actually use?' Otherwise, the whole project ends up like a paper no one reads.*

Together, these shifts signaled a growing awareness that effective AI use required more than technical proficiency. CRM analysts increasingly recognized the importance of aligning analytical insights with business realities, reflecting a cognitive recalibration toward commercial relevance. This attentiveness was accompanied by a growing willingness to engage end users early in the development process, ensuring that AI outputs would be interpretable and integrated into marketing practice. These adjustments marked a subtle but important inflection point: a movement toward more deliberate evaluation and structured coordination.

4.2. Nested Case Two: AI Implementation in the RM Department

The RM (Revenue Management) department was responsible for analyzing and optimizing pricing strategies across flight routes and fare classes to maximize revenue. To support this function, the department integrated a third-party AI-driven revenue optimization platform that used machine learning algorithms to project demand and recommend pricing adjustments. The system drew on both internal historical data from other departments and external inputs such as travel agency bookings. Revenue analysts, depending on their level of authorization, were responsible for segmenting historical data, monitoring demand curves generated by the AI, and adjusting input parameters when forecasts diverged from actual trends.

The introduction of the AI optimizer changed established work routines within the RM team. Previously, analysts compiled data from multiple sources - such as passenger loads, flow patterns, and average ticket prices - and used this information to manually construct pricing strategies. With AI integration, much of this process became automated, shifting the analysts' role from hands-on data preparation to oversight and selective intervention. As one analyst put it, their job had moved from "*manually processing and analyzing data*" to "*intervening in an automated procedure.*"

From the outset, RM analysts lacked confidence in surrendering pricing decisions to an external system. Their initial deployment of the AI optimizer was intentionally limited to a recommendation-and-alert mode. While analysts recognized the system's technical sophistication, many viewed AI as lacking the route-specific intuition and contextual insight that came from years of professional experience. As one analyst explained, "*Analysts cannot be substituted by an AI… Even with the perfect AI, it cannot take the experience of someone who has had the route for a long time*." This guarded approach reflected a professional commitment to domain mastery, rather than fear of harm. Indeed, analysts often described pricing as “more straightforward” and “AI-friendly,” suggesting that their hesitation was grounded in doubts about added value rather than concern about algorithmic failure.

Soon after deploying the AI platform as a recommendation-and-alert tool, RM analysts noticed that its pricing suggestions frequently conflicted with their own assumptions and experience. A telling example emerged on a route where BlueSky competed directly with a low-cost carrier. Historically, analysts kept prices closely aligned with the competitor, believing that significant price premiums, even with BlueSky’s superior service, would risk losing customers and reducing revenue. However, in contrast to this established logic, the AI system repeatedly recommended raising ticket prices.

The RM team, after following the AI’s recommendations, was surprised to discover that ticket sales actually increased despite the higher prices. Flights were consistently departing at full capacity. As one RM analyst noted:

> *The algorithm recommended selling at even higher prices, and I was thinking, 'Even more expensive? But there's a low-cost airline operating too.' We kept increasing the price despite the low-cost airline selling cheaper tickets, and we were still filling up. We couldn't understand why passengers were preferring us over a much cheaper alternative.*

These early conflicts between analyst intuition and algorithmic output gradually softened. Through controlled trials on selected routes, RM analysts came to recognize that AI’s superior data processing capabilities could reveal pricing opportunities they had previously missed.

Encouraged by the unexpected success on that route, the RM analysts conducted additional tests on routes with similar characteristics, such as competitive dynamics, demand profiles, flight schedules, and market trends. To their surprise, the results consistently echoed the initial case. The analysts attributed this performance to the AI’s superior computational power, which allowed it to objectively process vast amounts of real-time data from both internal and external sources. They viewed this analytical capacity as exceeding the limits of human cognition.

The AI's recommendations revealed opportunities that human analysts had overlooked, often due to ingrained assumptions, heuristics and constrained data interpretation. A pivotal insight from the AI's analysis was that BlueSky and its low-cost competitor were serving distinct customer segments. This realization enabled the RM department to better understand consumer behavior and adopt more revenue-oriented pricing strategies. As the head of the RM department explained:

> *It [AI] uses available data from various sources. For example, it conducts shadow shopping across all GDS [a ticket sales platforms used by airlines and travel agencies] and websites, calculating opportunities for upselling and profit maximization. This, I'd say, is a true evolution, a turning point for us in terms of revenue growth. Previously, analysts were often more conservative in their pricing, potentially leaving money on the table. But this tendency wasn't necessarily based on data; it was more like an unwritten rule for us humans, a collective reaction.*

Encouraged by the platform's success, the RM department shifted from using the AI tool merely for recommendations to fully automating revenue optimization on comparable routes. Confident in the system's analytical strength and perceiving little risk in its autonomous decisions, they further reduced human oversight to prevent the reintroduction of subjective bias. As the Head of the RM department noted: "*At the moment, what we are trying to do and say is, as little intervention in the system as possible*".

During the automated pricing optimization phase, an unexpected incident compelled analysts to reintroduce human oversight, highlighting the limitations of AI autonomy in ethically sensitive situations. A regional train accident disrupted services, leaving air travel as the sole transportation option for affected residents. In response to the sudden spike in demand for flights, the AI platform began raising ticket prices in line with its revenue-maximization objective.

However, RM analysts recognized that pricing decisions carry reputational consequences. Automatically increasing fares in such circumstances risked appearing exploitative, potentially undermining customer trust and harming the airline's public image. The incident underscored the AI's inability to account for social responsibility considerations. In response, the RM team became more cautious about fully delegating pricing authority to the system. As one RM data analyst explained:

> *We don't know how passengers will react to this [exploitative and unethical pricing practices during a crisis]. If they respond negatively, which is likely, we'll have to invest lots of time and resources to rebuild trust and regain their business on this route.*

It is important to recognize that the operational ethos of the BlueSky extended beyond purely economic objectives. Customer-centric values were central to the airline's identity. This commitment was evidenced by numerous awards recognizing its customer-friendly practices and high levels of customer satisfaction. BlueSky also actively engaged in broader societal initiatives, including the adoption of eco-friendly aircraft, reinforcing its dedication to sustainability and environmental stewardship.

To uphold the airline's social values and avoid perceptions of exploitation, RM analysts intervened to prevent reputational damage among customers, the public and employees. The incident raised concerns about AI's ability to navigate the contextual nuances of socially responsible action. In response, the team paused the AI platform to revise its rules and parameters. Importantly, this decision to override the AI platform was not triggered by model failure or clear evidence of long-term financial harm - indeed, the AI platform correctly detected and responded to the surge in demand - but by a perceived violation of the airline's commitments to social responsibility. As one RM team leader explained:

> *We didn't want to charge high prices when people had no other means of transportation [due to the accident]. So, we immediately halted the system because we didn't want to exploit the demand in an irresponsible way. The system knows how to maximize revenues but can't handle these extreme situations ethically... We have a social responsibility. If the analyst isn't alert to readjust rules and thresholds, the system can't understand all these nuances by itself.*

The RM department's experience with the AI revenue optimizer during the crisis highlighted a dual evaluative stance: while analysts expressed strong confidence in the system's ability to optimize for revenue, they concurrently voiced concerns about its limitations in ethically sensitive contexts, particularly its capacity to account for broader societal impacts or uphold the airline's social values.

As the AI revenue optimizer became embedded in BlueSky's pricing operations, it revealed a core challenge: aligning algorithmic logic with broader organizational values. Rather than treating the system as an infallible authority or a one-time deployment, RM analysts came to see AI implementation as a dynamic process - one requiring deliberate reflection, continuous oversight and organizational learning.

To support this, the department introduced a range of measures aimed at fostering more thoughtful engagement with the system. Analysts were trained not just to operate the platform, but to critically evaluate its suggestions. Senior team members provided informal mentorship, helping newer analysts understand when and how to question AI decisions, particularly when recommendations risked misaligning with BlueSky's public image or social commitments. As one team leader emphasized:

I tell them [junior analysts], it's their responsibility to monitor it [the AI]. Through daily analysis, I try to intrigue them on how to become better... You may not be able to monitor the entire process, but you must monitor the input and output. Do you simply accept its output and click 'approve'? Or do you challenge it to understand what happened? It's crucial to critically engage with the system's decisions.

This emphasis on discretion was further reinforced by structured comparison exercises. The RM team conducted regular tests comparing human and AI performance across different pricing scenarios, enabling analysts to detect when automated outputs might miss important contextual cues. As the department head explained:

We conduct experiments, human versus machine, to see who performs better under what conditions. If the machine does better, it should guide us, and we shouldn't intervene. Conversely, when humans outperform machines, we work to train the machine to reach the human level.

Over time, the RM department cultivated a measured approach that integrated AI's strengths with human intuition and value-based reasoning. Analysts became increasingly adept at identifying decision points where automation was appropriate, and where human judgment was indispensable not only to safeguard reputation, but to uphold the airline's ethos of responsible service.

4.3. Nested Case Three: AI Implementation in the Data Science Lab

The primary role of the DSL (Data Science Lab) was to support other departments by optimizing operations through data-driven analysis. One of its critical projects involved enhancing air freight operations for the cargo department. This department handled commercial merchandise transported in the aircraft's cargo holds - separate from passenger luggage. A persistent operational challenge was accurately estimating available cargo capacity, as passenger baggage was prioritized and total load was only finalized at the last-minute during airport check-in. As one data scientist explained: "*Some customers buy tickets at the counter, and some carry-ons get moved to the hold at the last minute.*" This last-minute variability made it difficult for cargo officers to confidently accept or reject freight requests for specific flights.

Historically, these decisions were made manually, relying heavily on the intuition of cargo officers. This approach often resulted in missed revenue opportunities due to underutilized space or operational issues when cargo was overcommitted. As one data analyst reflected:

> *Cargo officers were making decisions based solely on experience, sometimes blindly. This resulted in situations where more cargo was sent to the airport than the aircraft could accommodate, especially during busy periods like summer. These miscalculations caused significant disruptions and confusion in airport operations.*

To address these challenges, the cargo department collaborated with the DSL to develop an AI-driven tool designed to support more accurate, data-informed decisions about accepting or rejecting freight requests for specific flights.

A core challenge for the cargo department was fragmented information. Cargo analysts were required to manually consult multiple business units to assemble the necessary data, creating inefficiencies and increasing the likelihood of error. To overcome this, the DSL aggregated these disparate data sources into a unified dataset and developed an AI tool to generate daily predictions of available cargo space across all BlueSky flights. The aim was to replace the previously inconsistent and ad hoc decision-making with an automated, standardized process. As a data scientist explained:

> *The cargo department had to contact multiple departments, such as pet bookings, special equipment, passenger counts, and so on…. It was time-consuming…. By organizing historical data and building a prediction tool, we significantly improved decision-making.*

The development and implementation of the AI-based prediction tool reflected the DSL team's strong confidence in AI's ability to enhance operational efficiency. They believed the tool could seamlessly integrate data from multiple sources, deliver consistent, data-driven forecasts, and significantly reduce time inefficiencies. This confidence led the DSL to design a machine learning–driven tool capable of processing large volumes of cross-departmental data and generating more accurate predictions than manual methods. Framing cargo optimization as an engineering problem, the data scientists saw little risk of adverse outcomes. As one explained: "*AI is good for such tasks requiring engineering, coding, and so on… I trust it and if it shows an alert for a bug, I fix it, I trust it.*"

The DSL team's confidence in AI's ability to impose order on a fragmented system shaped the tool's design: it was built to handle complex, cross-departmental inputs and generate standardized actionable outputs. As one data scientist noted:

> *By organizing historical data from different sources and then building a prediction tool, we were able to improve the decisions-making in the cargo department... The cargo department manager even made a presentation, showing that by doing that they can have an extra revenue of almost a million.*

In this case, the inefficiency and fragmentation of cargo operations created conditions well aligned with AI's strengths. The opportunity to streamline workflows, paired with a low perceived risk, ultimately led the DSL to delegate core aspects of cargo management to the AI-powered tool.

Despite initial confidence in the AI tool, the efforts to automate cargo management soon revealed significant limitations. Although the AI often delivered promising results, it also produced inaccurate forecasts that caused operational disruptions. A data scientist recalled one such incident:

The cargo manager called me, clearly frustrated. He explained that the AI had predicted 450 kilograms of available cargo space for a flight. Being cautious, he had only sent 350 kilograms to the airport, yet even that couldn't be loaded.

This incident reflected AI’s limitations in handling the complex, cross-functional dynamics of cargo operations.

The goal to establish a fully automated and standardized process was impeded by the inherent complexities of airline logistics. Coordination among multiple departments, each with its own specialized knowledge and operational routines, introduced layers of complexity that the AI struggled to handle. For instance, the load control office, which managed weight distribution, relied on practices that were often opaque to both data scientists and cargo officers. External stakeholders, such as ground operations teams at airports, added further unpredictability that the AI tool found difficult to incorporate. As one data scientist explained:

Sometimes identical flights would carry the same cargo, but load control would approve one and deny the other. How is that possible? Or how can the algorithm account for who’s loading the aircraft? Some loaders are highly skilled, stacking cargo like Lego, while others aren’t as proficient. These multi-dimensional factors are too difficult to quantify and predict… There you want the human to intervene and feed this back to the algorithm”

In this case, concerns emerged regarding AI’s capacity to navigate the nuanced, cross-domain demands of cargo management. Its inability to account for context-specific variability and adapt to the dynamic needs of multiple stakeholders resulted in significant operational disruptions, raising doubts about its autonomous use. At the same time, the DSL team maintained strong confidence in AI’s ability to integrate large-scale, multi-source data and produce consistent predictive outputs. To leverage these strengths while addressing its limitations, the team supplemented the AI tool with human judgment and coordination, offering the situational awareness and cross-functional responsiveness that the system could not replicate.

As AI became embedded in cargo operations, the DSL team came to recognize that strict process standardization was not always viable in highly cross-domain contexts. While they continued to view AI as essential for streamlining fragmented workflows, they also acknowledged its limitations in handling the contextual variability and operational uncertainty characteristic of cargo logistics. In response, the DSL and cargo managers implemented a cross-functional coordination structure to surface and resolve issues that the AI tool could not address on its own. This adaptive approach helped clarify which decisions could be reliably delegated to the AI system and which required human discretion. As one data scientist reflected:

> *We are not a tech company; we are an airline…. Each field has its own peculiarities. Automation requires generalization, but becoming completely domain-agnostic is impossible. You need surrounding knowledge and steering of the process.*

To support this boundary work, the DSL and cargo team introduced mechanisms that paired automation with structured attentiveness. They regularly compared the AI system's predictions against actual cargo outcomes, using mismatches as signals to reassess inputs or update the model's logic. This encouraged a reflexive posture - one in which technical performance was not taken for granted but monitored for breakdowns that might reveal hidden constraints or overlooked factors. Disruptions became opportunities to recalibrate and inform future iterations. As one data scientist recounted:

> *A few months later, we acquired a new aircraft, and the algorithm struggled. We didn't know why, since decisions from other departments, like weight loading, were unknown to us. So, we coordinated across the organization to identify missing parameters and incorporate human feedback… If something went wrong, we'd pause, reflect, and adjust the tool to include buffers for future cases.*

In summary, these practices allowed the DSL team to draw a flexible boundary between what AI could standardize and where human interpretation remained essential. They learned to combine predictive automation with attentive monitoring to retain the benefits of algorithmic efficiency while adapting to the complexities of real-world operations.

4.4. Cross-Case Synthesis

Although the three departments at BlueSky implemented AI for different functional purposes, our analysis reveals important commonalities and contrasts in how AI was interpreted, adapted and stabilized in the organization. This subsection compares these implementation trajectories to distill the underlying logics and sensemaking processes shaping their evolution.

In CRM, the AI model was introduced to replace intuition-driven segmentation with analytically grounded models. Confidence in AI stemmed from its perceived methodological rigor, consistent with the team's broader commitment to systematic and replicable practices. Analysts' deep familiarity with the data, along with the system's ability to flag anomalies, reinforced a sense of control over potential errors - further minimizing perceived risks. This combination supported initial efforts to automate the identification of non-obvious customer segments. However, concerns soon emerged regarding the practical relevance of the outputs: while technically robust, they often failed to resonate with the marketing department's goals and priorities. These tensions reflected a broader divide between the analytical culture of the CRM team and the more pragmatic, action-oriented mindset of marketing. In response, analysts shifted their approach towards augmentation by reintroducing human oversight to contextualize insights for marketing use, while retaining an approach of automation for analytical clustering. To support this, they adopted *mindful evaluation practices*, regularly scrutinizing the outputs' marketing fit, and engaged in *proactive safeguarding* by working with marketers to iteratively refine segmentation. These adaptations helped preserve implementation momentum while ensuring insights remained actionable and aligned with commercial needs.

In RM, initial responses to the AI platform were reserved. While analysts acknowledged its potential to improve traditional pricing models, they approached its adoption with measured expectations. Confidence grew as the system consistently delivered measurable financial gains, particularly in routine pricing decisions. These early successes reinforced perceptions of AI's analytical reliability and led to its adoption as an autonomous pricing optimizer under stable market conditions. At this stage, the team perceived minimal risk, resulting in a relatively low degree of caution. However, concerns emerged when the system produced ethically questionable outputs, such as sharp price increases during periods of crisis, that risked damaging customer trust and public perception. Despite these concerns, confidence in the system's computational strengths remained intact. Analysts continued to value AI's pricing precision but recognized the need for human judgment in contexts requiring social sensitivity and reputational care. Situated between commercial performance demands and public accountability, the team adopted new practices to navigate competing pressures by automating routine decisions while actively intervening in value-laden scenarios. The practice of mindful evaluation enabled anticipatory judgment about when AI might misalign with broader expectations, while proactive safeguarding embedded routines like AI–human comparisons and escalation pathways.

In DSL, confidence in AI was strong from the outset. Data scientists viewed the system as a way to bring structure to a fragmented and variable operational environment. Its capacity to integrate data across time and space, and align the departments with their systems-oriented approach and engineering mindset. Initial confidence was further reinforced by the belief that cargo processes, though complex, followed recognizable patterns, minimizing early concerns about harmful consequences. This combination of methodological conviction and low perceived risk led the team to automate core elements of cargo coordination. As implementation progressed, however, cargo managers increasingly encountered exceptions, such as weather disruptions or last-minute changes, that the model could not adequately anticipate. These limitations were particularly salient given the nature of airline cargo management as a professional domain. The role demands coordination across a diverse set of internal and external stakeholders under conditions of high temporal pressure and frequent unpredictability. Within this context, the ability to respond flexibly to emergent conditions is not just a technical necessity but a core professional expectation. Recognizing the model's constraints in managing such situational complexity, the DSL maintained confidence in its analytical strengths while becoming more attuned to its contextual limitations. This evolving sensitivity led to a differentiated approach: stable, pattern-based components were automated, while segments requiring real-time human judgment and adaptation were augmented through human–AI collaboration. Alongside this, they also established interpretive routines that framed AI outputs as hypotheses to be tested and implemented safeguarding mechanisms linking modeling outputs with frontline feedback. These practices preserved the structured benefits of automation while also ensuring the adaptability required for dynamic cargo operations. Table 2 synthesizes these dynamics. A more detailed chain of evidence is provided in Appendix E.

Table 2. Cross-Case Comparison of AI Implementation Across BlueSky Departments

Dimension	CRM	RM	DSL
AI implementation goal	To systematize customer segmentation and value prediction by replacing heuristic marketing judgments with analytically driven models.	To optimize ticket pricing in real-time to increase revenue yield per route, using performance-based learning from historical trends	To forecast and allocate cargo space by integrating heterogeneous operational variables into standardized, scalable prediction models

Underlying rationality	Emphasis on analytical rigor and marketing usability: model validity mattered (*scientific objectivity*), but outputs needed to fit evolving campaign needs *(business pragmatism*)	Emphasis on financial returns and public defensibility: pricing strategies had to drive revenue (*revenue maximization*) while avoiding fairness breaches (*social responsibility*)	Emphasis on workflow stability and frontline adaptability (a*daptive flexibilit*y): forecasts had to support cross-department coordination amid changing constraints (*process standardization*)
Role positioning	Dual-positioned as analytical stewards: upholding methodological rigor while adapting outputs to evolving campaign relevance	Dual-positioned as strategic custodians: balancing revenue-driven analysis with ethical sensitivity in pricing	Dual-positioned as systemic integrators: building structured AI processes while adjusting for live operational variation
Trust in AI	Affirmed: Belief in AI's objectivity and methodological integrity as a superior alternative to intuition	Guarded: Initial limited expectations of performance gains from AI implementation Affirmed: Trust increased after AI proved superior in identifying pricing efficiencies	Affirmed: Perception of AI as a dependable engine for integrating diverse datasets and structuring decisions
Distrust in AI	Latent: Limited concern over algorithmic bias in AI outputs Pronounced: Salient concern over output relevance to campaign design.	Latent: Minimal unease about risk in routine pricing scenarios Pronounced: Heightened wariness when AI conflicted with fairness norms;	Latent: Absence of worry over AI errors going undetected and causing operational issues Pronounced: Acute concern about model inflexibility in unanticipated disruptions.
Mindful evaluation	Critically interpreted AI segments, assessing marketing fit and revalidating outputs through collaborative reviews	Reflected on AI pricing through fairness and appropriateness, especially during internal reviews and crises	Treated AI forecasts as hypotheses, comparing model outputs with frontline realities and adjusting judgment
Proactive safeguarding	Co-refined segments with marketers and incorporated interpretive feedback in iterative cycles	Conducted experimental AI-human output comparisons and instituted regular mentorship between senior and junior analysts	Created team-based alert systems and embedded real-time feedback channels with ground operations and load teams

Our cross-case synthesis reveals that AI implementation trajectories are shaped not only by a system's technical features but more fundamentally by how actors interpret and evaluate those features through shared assumptions about what counts as effective, appropriate and meaningful decision support. Across the three departments, AI was engaged with not simply as a tool, but as a system whose value depended on how it fit with professional practices, domain priorities and operational contingencies. These interpretations gave rise to distinct configurations of confidence and concern, which in turn influenced how AI was embedded, challenged or modified in practice, which we theorize in the next section.

5. Theorization

As discussed above, AI implementation trajectories reflect the institutional pluralism in which contemporary organizations are embedded (Thornton et al., 2012; Greenwood et al., 2011). Yet, institutional pluralism rarely takes the form of overt clashes between fully articulated societal-level institutional orders and their associated logics (Besharov & Smith, 2014; Pache & Santos, 2013). Rather, it is enacted in situated and embodied ways - through selective, condensed and often hybridized instantiations of these broader orders that guide actors' interpretations and responses to novel technologies (Berente et al., 2019; Dunn & Jones, 2010). Our study contributes to this perspective by showing how institutional pluralism becomes actionable in the field through specific evaluative orientations that structure how actors assess AI systems as more or less appropriate, legitimate or valuable.

Across our cases, we found that these evaluative orientations, while varied in expression, consistently clustered around two dominant and overarching logics: *an instrumental–analytic logic*, focusing on calculative rationality, internal coherence and replicable performance; and *a contextual–normative logic*, foregrounding ethical accountability, practical appropriateness, and situational responsiveness. Importantly, we conceptualize the two logics we identify not as idiosyncratic products of our cases, but as organizational instantiations of wider societal-level institutional orders that actors draw upon and hybridize to make sense of their work and to evaluate the legitimacy of practices and technologies (Faik et al., 2020). These logics are therefore grounded in extra-organizational institutions rather than generated anew within a single case. In line with the ideal-type orders in the institutional logics literature (Thornton et al., 2012), we interpret the *instrumental–analytic logic* as a hybrid of corporate and market logics, given its emphasis on performance management, control and efficiency; and the *contextual–normative logic* as a hybrid of profession and community logics, given its focus on situational understanding, ethical responsibility, and shared norms. In our data, what emerges empirically are not entirely new, purely local logics, but rather specific enactments and combinations of these institutionalized logics that guide actors' interpretation of legitimacy, value and appropriate conduct in the context of AI implementation.

To examine how these logics materialized in practice, we focused on three elemental categories that have proven central to understanding how logics shape organizational action: principles, goals and identities (Greenwood et al., 2011; Thornton et al., 2012). We summarize and capture these categories in Appendix F. These categories were not chosen arbitrarily but because they capture the cognitive (principles), teleological (goals) and normative/affective (identities) aspects of logics, offering a composite yet parsimonious lens for interpreting how actors navigate competing expectations in AI implementation and adoption in organizations (Berente & Yoo, 2012; Glynn, 2008; Lok, 2010; Reay & Hinings, 2009).

Our analysis shows that the instrumental–analytic logic embodies a deep belief in the authority of systematic reasoning and structured control. This logic rests on the premise that legitimacy stems from demonstrable rigor, consistency, and replicability. Within this logic, AI is valued for its ability to stabilize decision-making through codified procedures and evidentiary standards (Marett et al., 2013). This orientation foregrounds goals such as predictive precision, operational efficiency and the disciplined orchestration of complexity. It also shapes a corresponding sense of professional identity, one that casts actors as stewards of analytic integrity or architects of system coherence (Berente and Yoo, 2012). From an institutional logics perspective, we interpret the instrumental–analytic logic as an organizational instantiation of a hybrid between the corporate and market institutional orders (Thornton et al., 2012). It translates corporate concerns with hierarchical coordination, standardization and internal control, together with market emphases on efficiency, calculability and comparative evaluation, into local expectations about what counts as “proper” AI use. Across our cases, this hybrid logic guided how actors interpreted the value of AI, affirmed when it reinforced institutionalized standards of calculative order, and questioned when it failed to accommodate the situated demands of application.

The analysis then also shows that the contextual–normative logic is grounded on a moral and interpretive sensibility that views technological legitimacy as contingent on responsiveness to social expectations and situated accountability. At the organizational level, this logic privileges ethical attunement, reputational stewardship and the practical intelligibility of system outputs (Miranda et al., 2015). Here, AI is not only judged by its formal accuracy, but by its capacity to accommodate the situational ambiguities and value-laden demands of real-world contexts. Goals tied to this logic emphasize fairness, stakeholder awareness and contextual appropriateness (Hultin & Mähring, 2014). Actors positioned within this logic see themselves as moral custodians or responsive coordinators, those responsible for translating algorithmic abstraction into practical, socially meaningful and justifiable action. We therefore read the contextual–normative logic as an organizational enactment of a hybrid between the professional and community institutional orders (Thornton et al., 2012) because it joins a professional emphasis on expert, context-sensitive judgment and accountable discretion with a community orientation toward shared values, ethical standards, mutual care and the maintenance of collective reputation (Thornton et al., 2012). In our data, this logic surfaced when the perceived neutrality of AI clashed with expectations of fairness, sensitivity or applicability, prompting critical engagement with the system’s normative implications.

We then noticed that these two dominant logics of *instrumental-analytic* and *contextual–normative* influenced the development of trust and distrust in AI, and that trust and distrust emerged as institutionally embedded constructs rather than just individual psychological traits as often portrayed in the literature (Glikson & Woolley, 2020). This is consistent with Giddens' (1990) views that modern trust is often vested in abstract systems of expertise rather than in individuals, suggesting that confidence in AI is also grounded in these institutionalized expectations in social setting and within organizations. Specifically, trust in AI reflects a perceived alignment between the AI system's capabilities and the dominant institutional logic, whereas distrust reflects a misalignment between the two. For example, in our case, *trust* tends to increase when AI is seen to fulfill the instrumental–analytic logic, offering perceived benefits such as efficiency, consistency or analytic rigor; and *distrust* intensifies when AI appears to violate the contextual–normative logic through opacity, rigidity or heightened wariness of ethical and situational blind spots. Therefore, it is not just the presence of trust or distrust that matters, but the way they emerge to reflect inherent dominant logics within complex institutional environments because this shapes organizational responses to AI implementation.

We suggest that the configurations of trust and distrust operate by translating and reflecting institutional logics as different AI implementation choices, emphasizing automation or augmentation as two different but coexisting pathways (Agrawal et al., 2023). The analysis shows that the configurations of trust and distrust reflect institutionalized affirmation and wariness: trust entails affirmation of positive outcomes (e.g., accuracy, efficiency), while distrust reflects pronounced wariness of negative consequences (e.g., ethical lapses, contextual failure) (Lewicki et al., 1998).

More specifically, when AI systems are perceived to align with the instrumental–analytic logic, that is in fulfilling goals of objectivity, efficiency and scalable decision-making, they foster affirming trust. At the same time, as long as AI systems do not overtly conflict with contextual–normative expectations, distrust remains latent. This configuration, affirming trust coupled with latent distrust, leads organizational actors to view the AI as a reliable and legitimate instrument with minimal perceived risk, thereby legitimizing an *automation* oriented implementation pathway.

However, affirming trust does not always produce automation. When both trust and distrust are strong, organizations recognize AI's instrumental value but also perceive it as misaligned with contextual–normative principles, such as fairness, explainability or situational judgment. In this case, the configuration emphasizes *augmentation* as a human–AI collaboration approach becomes a way to benefit from the AI's strengths while safeguarding against its perceived limitations.

We also observed an indirect configuration of guarded trust and latent distrust in nested case two, where actors prioritized a restricted experimentation with the AI. Because this pattern did not recur across our three nested cases - and focusing on a single instance risks capturing a more idiosyncratic, situation-specific response rather than a shared, institutionally grounded pattern - we retain this configuration in the empirical narrative but do not foreground it in our core theoretical framework, which concentrates on cross-case regularities in how institutional logics shape configurations of trust and distrust in AI implementation.

These distinct institutional pressures lead to inherent tensions that push AI adoption in different directions. We found that organizations actively engage in reconciliation practices to manage the inherent institutional contradictions arising from AI implementation. We identified two key practices: *mindful evaluation* and *proactive safeguarding*. These practices function as practical coping mechanisms that address the collision between the automation and augmentation contradiction. *Mindful evaluation* involves deliberate reflection on AI outputs as propositions subject to institutional scrutiny. Rather than fully accepting or rejecting AI recommendations, actors adopt a questioning stance when following this practice by continuously assessing system performance across multiple criteria, such as technical accuracy, fairness and contextual fit. This practice echoes Weick's (1995) concept of mindfulness in organizing and resonates with Orlikowski's (1996) notion of improvisational use of technology, where routines remain open to adjustment.

On the other hand, the practice of *proactive safeguarding* embeds structural protections into AI use. This practice involves configuring socio-technical routines that define and constrain the domains and conditions under which AI operates. For example, by benchmarking AI against human decisions, conducting scenario testing or instituting oversight checkpoints. These practices function as buffers, mitigating risks of normative misalignment by selectively coupling AI outputs to organizational decision-making processes (Pache & Santos, 2013).

Figure 1 summarizes and provides a visual representation of our theorization. It illustrates how the two dominant organizational level logics identified in our analysis configure trust and distrust in AI. These trust and distrust configurations, in turn, shape how AI is implemented within organizations. The figure also highlights the two practices of mindful evaluation and proactive safeguarding, that operate to manage the inherent tensions between these contradictory guiding logics in everyday organizational practice.

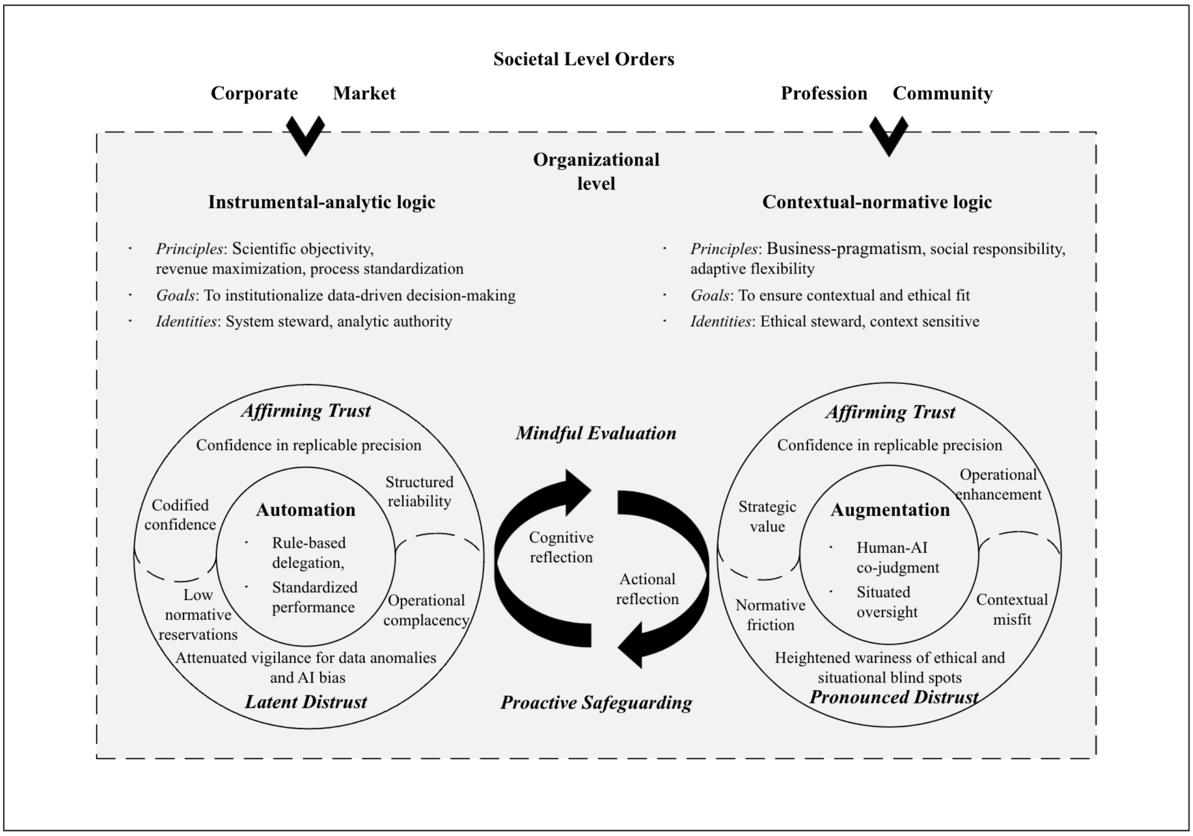

Figure 1. Theoretical Framework

In summary, our theorization explains AI implementation as an institutionally embedded process shaped by how organizations navigate value conflicts under conditions of institutional pluralism. We identify two evaluative institutional logics at the organizational level: *instrumental–analytic* and *contextual–normative*, that guide how organizational actors interpret and evaluate AI. These logics translate abstract institutional pluralism into situated expectations, shaping the emergence of trust and distrust as configurations of intensity that reflect alignment or misalignment with institutional priorities. These configurations act as mechanisms that steer implementation pathways toward automation and augmentation. Importantly, organizations do not passively absorb these conflicts but actively manage them through reconciliation practices: *mindful evaluation* and *proactive safeguarding*. Together, these practices help organizations dynamically balance efficiency with legitimacy.

6. Discussion and Conclusion

This study examined how institutional pressures and pluralism shape AI implementation and adoption within organizations. We find that AI is not adopted solely based on technical performance or user attitudes, but rather through how its capabilities are interpreted in relation to coexisting principles, goals, and identities embedded in different institutional logics. These logics shape distinct configurations of trust and distrust, which in turn orient organizations towards automation or augmentation. Organizations manage the inherent tensions between the two paradigms through two mechanisms of *mindful evaluation* and *proactive safeguarding*, by continually reconciling efficiency with ethical and contextual demands.

6.1. Theoretical Contributions

Our study makes several theoretical contributions. First, the study positions trust and distrust as institutionally grounded mechanisms rather than as psychological states or technical judgments. The literature has treated trust in AI as a function of system performance (e.g., reliability, accuracy) or user disposition (e.g., familiarity, risk tolerance) (Glikson & Woolley, 2020; Yang & Wibowo, 2022). We depart from this literature by showing that trust and distrust in AI are shaped by institutional logics regarding what constitutes legitimate, responsible or desirable technology use. Building on Giddens' (1990) notion of trust in abstract systems and the institutional logics perspective (Thornton et al., 2012), we demonstrate how organizational actors assess AI not in a vacuum, but through the lens of coexisting normative frameworks. Trust arises when AI aligns with valued institutional logics such as analytic rigor or business pragmatism, while distrust emerges when the system is perceived to violate expectations of fairness, transparency or contextual appropriateness. These interpretations are not mutually exclusive and often coexist in dynamic configurations. By conceptualizing trust and distrust as socially embedded, value-laden constructs, our study offers a more sociologically attuned account of how organizations make sense of emerging technologies under normative complexity.

Second, we contribute to AI implementation and adoption research by theorizing automation and augmentation not as neutral design choices, but as institutionally informed, interrelated implementation trajectories. Prior studies have examined automation and augmentation as distinct technological orientations (Jarrahi, 2018; Raisch & Krakowski, 2021; van den Broek et al., 2021), often treating the choice between them as a function of task characteristics or epistemic dynamics during model development. Our findings suggest that the two implementation paths are also shaped by institutional pluralism and are rarely realized in pure form. Instead, organizations combine automated and augmentation-oriented uses of AI across tasks and decision components. We show that these combinations are patterned by how actors configure trust and distrust in response to competing institutional logics. For elements of work where AI is perceived to fulfill instrumental-analytic priorities (e.g., efficiency, analytic objectivity) with minimal normative concerns, automation becomes a legitimate trajectory. However, for elements where confidence in AI's technical capability coexists with pronounced wariness about its contextual, normative or organizational fit, actors favor augmentation to retain discretion and safeguard against misalignment. This insight extends recent work in institutional theory that emphasizes how plural logics shape organizational responses to innovation (Greenwood et al., 2011; Hinings et al., 2018), by showing how these logics become enacted through the trust-distrust configurations that, in turn, drive hybrid mixes of automation and augmentation in concrete forms of AI use.

Third, we advance a processual understanding of how organizations navigate AI implementation under institutional complexity by identifying two reconciliation practices: *mindful evaluation* and *proactive safeguarding*. While prior work has recognized the importance of institutional work (Lawrence & Suddaby, 2006) and loose coupling strategies

(Pache & Santos, 2013) in managing conflicting demands, our study shows how these concepts play out in the context of intelligent technologies. Mindful evaluation enables actors to treat AI outputs not as conclusive judgments but as provisional prompts for reflection. Proactive safeguarding, in turn, builds organizational guardrails - such as human-AI benchmarking, ethical scenario testing, or role-based oversight - that buffer institutional values from being overridden by automation. Together, these practices illustrate how organizations can respond to institutional contradictions not by choosing one logic over another, but by maintaining a productive tension between them. In doing so, we contribute to the literature on institutional reflexivity (Smets et al., 2015), offering a pragmatic account of how AI implementation is continually recalibrated to reconcile legitimacy and performance goals over time.

6.2. Practical Contributions

Our study offers several valuable insights for practitioners navigating the complexities of AI implementation in organizations. First, we highlight the importance of explicitly acknowledging AI as both a technical system and an institutional object. Rather than viewing automation and augmentation as purely functional choices, practitioners should recognize them as value-laden responses to competing organizational priorities. Framing AI implementation in this way enables more deliberate decision-making, matching automation to contexts where confidence in AI is institutionally supported, and reserving augmentation for areas that demand human judgment, accountability or ethical oversight.

Second, we offer a structured lens for embedding responsibility into AI use through two core practices: mindful evaluation and proactive safeguarding. While these practices emerged from our empirical context, they generalize as designable organizational routines. Mindful evaluation refers to a stance of reflective skepticism, where actors treat AI outputs not as definitive answers but as propositions to be interpreted and contextualized. This mindset can be fostered through lightweight interventions such as structured deliberations, pre-deployment scenario analyses, or post-decision reviews that encourage users to critically appraise AI recommendations across multiple criteria (e.g., technical validity, fairness, contextual relevance). Proactive safeguarding represents an action-oriented practice that builds structural protections around AI workflows, such as role-based overrides, exception-handling protocols, or comparative evaluation exercises, that constrain how and where AI can operate. Importantly, by surfacing where AI reliably succeeds versus where it triggers overrides or edge-case flags, these practices furnish the very feedback needed to distinguish tasks for full automation from those better suited to augmentation. The practices of mindful evaluation and proactive safeguarding are not peripheral add-ons; rather, they should be embedded into the organizational infrastructure to ensure that AI use remains aligned with both institutional values and operational goals.

Third, our findings highlight that effective AI implementation is not just a matter of technological fit, but of organizational alignment. Cross-functional coordination is essential to anticipate and resolve misalignments between AI outputs and domain-specific expectations. This requires creating institutional space for deliberation, not only to surface local concerns, but also to ensure that evolving AI capabilities remain anchored to organizational values. Lastly, we encourage organizations to view AI governance as an ongoing process rather than a one-time design. Institutional expectations shift, technologies evolve, and so must the interpretive frameworks that mediate their use. A reflexive, principle-aware approach enables organizations to sustain both legitimacy and performance as they adapt to new sociotechnical realities.

6.3. Limitations and Future Research

We recognize some limitations in our study. First, our findings capture AI implementation at a particular point in time. Given the fast-evolving nature of AI technologies and their organizational uses, longitudinal research is needed to examine how trust and distrust evolve as AI systems become more advanced, autonomous and embedded in core business processes.

Such work could illuminate how implementation trajectories shift over time and how new actors, organizational routines, or external triggers reshape these dynamics.

Second, although we highlight how institutional logics shape AI use within organizations, our analysis does not fully explore how broader environmental factors, such as regulation, media narratives, or shifting societal expectations, interact with internal practices within organizations. Future research could investigate how changes in public discourse or policy frameworks shape organizations' trust-distrust configurations and governance mechanisms for AI. This would enrich understanding of the feedback loops between societal-level institutional logics and organizational implementation strategies. Third, while we gesture toward societal implications, our empirical focus remains of three nested case studies within a single organization. Future studies could expand the unit of analysis to examine cross-organizational or industry-wide patterns, particularly how collective responses to AI emerge and institutionalize over time through professional associations, industry standards or regulatory coalitions. Fourth, our study does not engage with related constructs such as algorithmic aversion or appreciation (Castelo et al., 2019; Logg et al., 2019). These concepts primarily reflect evaluative preferences, that is, whether individuals favor human or algorithmic recommendations in decision-making contexts. In contrast, trust and distrust involve a broader orientation of willingness to rely on a system under conditions of uncertainty, grounded in perceptions of legitimacy, appropriateness and alignment with institutional expectations. Future research could extend this distinction by examining how algorithmic aversion and appreciation themselves may be shaped by institutional environments, roles, and value commitments.

6.4 Conclusion

This study advances understanding of AI implementation and adoption within organizations by revealing how different actors navigate the inherent tensions between the two dominant paradigms of AI implementation of automation and augmentation. Our analysis identifies two distinct normative pressures in AI implementation stemming from *instrumental–analytic* and *contextual–normative logics*. These two organizational level logics shape configurations of trust and distrust which then also guide different AI implementation pathways. These trajectories are stabilized through reconciliation practices of *mindful evaluation* and *proactive safeguarding*. By linking institutional logics, trust dynamics and organizational practices, our study offers a sociotechnical account of responsible AI implementation in complex institutional environments.

References

Agrawal, A., Gans, J. S., & Goldfarb, A. (2023). Do we want less automation? Science, 381(6654), 155-158.

Avgerou, C. (2001). The significance of context in information systems and organizational change. *Information Systems Journal*, 11(1), 43-63.

Baer, I., Waardenburg, L., & Huysman, M. (2025). What Is Augmented? A Metanarrative Review of AI-Based Augmentation. *Journal of the Association for Information Systems*, 26(3), 760-798.

Baptista, J., Stein, M. K., Klein, S., Watson-Manheim, M. B., & Lee, J. (2020). Digital work and organisational transformation: Emergent Digital/Human work configurations in modern organisations. *The Journal of Strategic Information Systems*, *29*(2), 101618.

Benbya, H., Pachidi, S., & Jarvenpaa, S. L. (2021). Special Issue Editorial: Artificial Intelligence in Organizations: Implications for Information Systems Research. *Journal of the Association for Information Systems*, *22*(2), 281–303.

Berente, N., & Yoo, Y. (2012). Institutional Contradictions and Loose Coupling: Postimplementation of NASA's Enterprise Information System. *Information Systems Research*, *23*(2), 376–396.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing Artificial Intelligence. *MIS Quarterly*, *45*(3), 1433-1450.

Besharov, M. L., & Smith, W. K. (2014). Multiple institutional logics in organizations: Explaining their varied nature and implications. *Academy of Management Review*, *39*(3), 364-381.

Binns, R., Van Kleek, M., Veale, M., Lyngs, U., Zhao, J., & Shadbolt, N. (2018, April). 'It's Reducing a Human Being to a Percentage' Perceptions of Justice in Algorithmic Decisions. In *Proceedings of the 2018 Chi conference on human factors in computing systems* (pp. 1-14).

Carroll, N., Holmström, J., Stahl, B. C., & Fabian, N. E. (2024). Navigating the utopia and dystopia perspectives of artificial intelligence. Communications of the Association for Information Systems, 55(1), 854-874.

Castelo, N., Bos, M. W., & Lehmann, D. R. (2019). Task-dependent algorithm aversion. *Journal of Marketing Research*, 56(5), 809–825.

Cheatham, B., Javanmardian, K., & Samandari, H. (2019). Confronting the risks of artificial intelligence. *McKinsey Quarterly*.

Corbin, J., & Strauss, A. (2008). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (3rd ed.). Thousand Oaks, CA: Sage.

Constantiou, I., Joshi, M. P., & Stelmaszak, M. (2024). Research Handbook on Artificial Intelligence and Decision Making in Organizations. Edward Elgar Publishing. https://doi.org/10.4337/9781803926216

Currie, W. L., & Guah, M. W. (2007). Conflicting institutional logics: A national programme for IT in the organisational field of healthcare. *Journal of Information Technology*, 22(3), 235-247.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, *48*(2), 147-160.

Dimoka, A. (2010). What does the brain tell us about trust and distrust? Evidence from a functional neuroimaging study. *MIS Quarterly*, 34(2), 373-396.

Dunn, M. B., & Jones, C. (2010). Institutional Logics and Institutional Pluralism: The Contestation of Care and Science Logics in Medical Education, 1967–2005. *Administrative Science Quarterly*, *55*(1), 114–149.

Faik, I., Barrett, M., & Oborn, E. (2020). How Information Technology Matters in Societal Change: An Affordance-Based Institutional Perspective. *MIS Quarterly*, *44*(3), 1359–1390.

Faik, I., Gwee, M., Tan, F. T. C., Leong, C., & Hastiadi, F. F. (2026). When digital platforms enter informal sectors: Work formalization and institutional change. *The Journal of Strategic Information Systems*, *35*(1), 101941.

Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, *28*(1), 62–70.

Friedland, R., & Alford, R. R. (1991). Bringing society back in: Symbols, practices, and institutional contradictions in the new institutionalism in organizational analysis. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 232–263). University of Chicago Press.

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, *33*(2), 678–696.

Fuglsang, L., & Jagd, S. (2015). Making sense of institutional trust in organizations: Bridging institutional context and trust. *Organization*, 22(1), 23-39.

Giddens, A. (1990). *The consequences of modernity*. Polity Press in association with Basil Blackwell, Oxford, UK.

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, *14*(2), 627–660.

Glynn, M. A. (2008). Beyond Constraint: How Institutions Enable Identities. In R. Suddaby, R. Greenwood, K. Sahlin, & C. Oliver (Eds.), *The SAGE Handbook of Organizational Institutionalism* (pp. 413–430). SAGE Publications Ltd.

Greenwood, R., Raynard, M., Kodeih, F., Micelotta, E. R., & Lounsbury, M. (2011). Institutional Complexity and Organizational Responses. *The Academy of Management Annals*, *5*(1), 317–371.

Gregor, S. (2006). The Nature of Theory in Information Systems. *MIS Quarterly*, *30*(3), 611-642.

Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. *The Journal of Strategic Information Systems,* 29(2), 101614.

Guo, S. L., Lumineau, F., & Lewicki, R. J. (2017). Revisiting the foundations of organizational distrust. *Foundations and Trends in Management*, 1(1), 1-88.

Hayes, N. (2008). Institutionalizing change in a high-technology optronics company: The role of information and communication technologies. *Human Relations*, 61(2), 243-269.

Heyder, T., Passlack, N., & Posegga, O. (2023). Ethical management of human-AI interaction: Theory development review. *The Journal of Strategic Information Systems*, *32*(3), 101772.

Hinings, B., Gegenhuber, T., & Greenwood, R. (2018). Digital innovation and transformation: An institutional perspective. *Information and Organization*, *28*(1), 52-61.

Hultin, L., & Mähring, M. (2014). Visualizing institutional logics in sociomaterial practices. *Information and Organization*, 24(3), 129-155.

Huysman, M. (2025). Studying AI in the Wild: Reflections from the AI@ Work Research Group. *Journal of Management Studies*.

Iansiti, M.., & Lakhani, K. R. (2020). *Competing in the age of AI: strategy and leadership when algorithms and networks run the world*. Harvard Business Review Press.

Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, *61*(4), 577-586.

Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, *24*(4), 602-611.

Karunakaran, A., Orlikowski, W. J., & Scott, S. V. (2022). Crowd-based accountability: Examining how social media commentary reconfigures organizational accountability. *Organization Science*, *33*(1), 170-193.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*(6), 893-905.

Kell, J. (2023, January 31). How the airline industry is using A.I. to improve the entire experience of flying. *Fortune*. https://fortune.com/2023/01/31/tech-forward-everyday-ai-airline-industry-fuelconsumption-food-waste/

Kostis, A., Bengtsson, M., & Näsholm, M. H. (2022). Mechanisms and dynamics in the interplay of trust and distrust: Insights from project-based collaboration. *Organization Studies*, 43(8), 1173-1196.

Kraatz, M. S., & Block, E. S. (2008). Organizational Implications of Institutional Pluralism. In R. Suddaby, R. Greenwood, K. Sahlin, & C. Oliver (Eds.), *The SAGE Handbook of Organizational Institutionalism* (pp. 243–275). SAGE Publications Ltd.

Krakowski, S., Luger, J., & Raisch, S. (2023). Artificial intelligence and the changing sources of competitive advantage. *Strategic Management Journal*, *44*(6), 1425–1452.

Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior*, 123, 106878.

Langley, A. (1999). Strategies for Theorizing from Process Data. *The Academy of Management Review*, *24*(4), 691-710.

Lankton, N., McKnight, D. H., & Tripp, J. (2015). Technology, Humanness, and Trust: Rethinking Trust in Technology. *Journal of the Association for Information Systems*, *16*(10), 880–918.

Lawrence, T. B., & Suddaby, R. (2006). Institutions and Institutional Work. In T. B. Lawrence, C. Hardy, S. R. Clegg, & W. R. Nord (Eds.), *The SAGE Handbook of Organization Studies* (pp. 215–254). SAGE Publications, Limited.

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). Is AI ground truth really true? The dangers of training and evaluating AI tools based on experts' know-what. *MIS Quarterly*, *45*(3), 1501-1526.

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, 33(1), 126-148.

Lei, Y. W., & Kim, R. (2024). Automation and augmentation: Artificial intelligence, robots, and work. *Annual Review of Sociology*, *50,* 251-272.

Lewicki, R. J., McAllister, D. J., & Bies, R. J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, *23*(3), 438-458.

Lifshitz-Assaf, H. (2018). Dismantling Knowledge Boundaries at NASA: The Critical Role of Professional Identity in Open Innovation. *Administrative Science Quarterly*, *63*(4), 746–782.

Lindebaum, D., Vesa, M., & den Hond, F. (2020). Insights from *The Machine Stops* to Better Understand Rational Assumptions in Algorithmic Decision-Making and Its Implications for Organizations. *Academy of Management Review, 45*(1), 247-263.

Lok, J. (2010). Institutional logics as identity projects. *Academy of Management Journal*, *53*(6), 1305-1335.

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103.

Lounsbury, M. (2023). The Problem of Institutional Trust. *Organization Studies*, *44*(2), 308–310.

Luhmann, N. (2018). *Trust and power*. John Wiley & Sons.

Lumineau, F. (2017). How contracts influence trust and distrust. *Journal of Management*, 43(5), 1553-1577.

Lyytinen, K., Nickerson, J. V, & King, J. L. (2021). Metahuman systems = humans + machines that learn. *Journal of Information Technology*, *36*(4), 427–445.

Marett, K., Otondo, R. F., & Taylor, G. S. (2013). Assessing the effects of benefits and institutional influences on the continued use of environmentally munificent bypass systems in long-haul trucking. *MIS Quarterly*, 37(4), 1301-1312.

Markus, M. L., & Silver, M. (2008). A Foundation for the Study of IT Effects: A New Look at DeSanctis and Poole's Concepts of Structural Features and Spirit. *Journal of the Association for Information Systems*, *9*(10), 609–632.

Martinsons, M. G. (2008). Relationship-based e-commerce: Theory and evidence from China. *Information Systems Journal*, 18(4), 331-356.

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., Capstick, E., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., Walsh, T., Hamrah, A., Santarlasci, L., Betts Lotufo, J., Rome, A., Shi, A., & Oak, S. (2025, April). *The AI Index 2025 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734.

McKendrick, J., & Thurai, A. (2022). AI Isn't Ready to Make Unsupervised Decisions. Harvard Business Review. https://hbr.org/2022/09/ai-isnt-ready-to-make-unsupervised-decisions

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, *83*(2), 340-363.

Mignerat, M., & Rivard, S. (2009). Positioning the institutional perspective in information systems research. *Journal of Information Technology*, *24*(4), 369-391.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE.

Miranda, S. M., Kim, I., & Summers, J. D. (2015). Jamming with social media. *MIS Quarterly*, 39(3), 591-614.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, *1*(11), 501–507.

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. *International Journal of Qualitative Methods*, *1*(2), 13–22.

Omrani, N., Rivieccio, G., Fiore, U., Schiavone, F., & Agreda, S. G. (2022). To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts. *Technological Forecasting and Social Change*, 181, 121763.

Orlikowski, W. J. (1996). Improvising organizational transformation over time: A situated change perspective. *Information Systems Research*, *7*(1), 63-92.

Orlikowski, W. J., & Scott, S. V. (2021). Liminal innovation in practice: Understanding the reconfiguration of digital work in crisis. *Information and Organization*, *31*(1), 100336.

Osborne, M. (2023, May 20). Artificial intelligence holds huge promise – and peril. Let's choose the right path. *The Guardian*. https://www.theguardian.com/commentisfree/2023/may/20/artificial-intelligence-holds-huge-promise-and-peril-lets-choose-the-right-path

Pache, A. C., & Santos, F. (2013). Inside the hybrid organization: Selective coupling as a response to competing institutional logics. *Academy of Management Journal*, *56*(4), 972-1001.

Raisch, S., & Krakowski, S. (2021). Artificial Intelligence and Management: The Automation–Augmentation Paradox. *Academy of Management Review*, *46*(1), 192–210.

Raisch, S., & Fomina, K. (2025). Combining human and artificial intelligence: Hybrid problem-solving in organizations. *Academy of Management Review*, *50*(2), 441-464.

Reay, T., & Hinings, C. R. (2009). Managing the rivalry of competing institutional logics. *Organization Studies*, 30(6), 629-652.

Saunders, M. N., Dietz, G., & Thornhill, A. (2014). Trust and distrust: Polar opposites, or independent but co-existing? *Human Relations*, *67*(6), 639–665.

Schad, J., Lewis, M. W., Raisch, S., & Smith, W. K. (2016). Paradox research in management science: Looking Back to Move Forward. *The Academy of Management Annals*, *10*(1), 1-60.

Scharowski, N., Perrig, S. A., von Felten, N., Aeschbach, L. F., Opwis, K., Wintersberger, P., & Brühlmann, F. (2025). To Trust or Distrust AI: A Questionnaire Validation Study. *In Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*, (pp. 361-374).

Schoorman, F. D., Mayer, R. C., & Davis, J. H. (2007). An Integrative Model of Organizational Trust: Past, Present, and Future. *Academy of Management Review*, *32*(2), 344–354.

Seidel, S., Frick, C. J., & vom Brocke, J. (2025). Regulating emerging technologies: Prospective sensemaking through abstraction and elaboration. *MIS Quarterly*, 49(1), 179-204.

Shao, Y., Huang, C., Song, Y., Wang, M., Song, Y. H., & Shao, R. (2024). Using augmentation-based AI tool at work: A daily investigation of learning-based benefit and challenge. *Journal of Management*, *51*(8), 3352-3390.

Shollo, A., & Galliers, R. D. (2024). Chapter 11: Constructing actionable insights: The missing link between data, artificial intelligence, and organizational decision-making. *In Research Handbook on Artificial Intelligence and Decision Making in Organizations* (pp. 195–213). Edward Elgar Publishing.

Shrestha, Y., Ben-Menahem, S., & von Krogh, G. (2019). Organizational Decision-Making Structures in the Age of Artificial Intelligence. *California Management Review*, 61(4), 66–83.

Slavova, M., & Karanasios, S. (2018). When institutional logics meet information and communication technologies: Examining hybrid information practices in Ghana's agriculture. *Journal of the Association for Information Systems*, *19*(9), 775-812.

Smets, M., Jarzabkowski, P., Burke, G. T., & Spee, P. (2015). Reinsurance trading in Lloyd's of London: Balancing conflicting-yet-complementary logics in practice. *Academy of Management Journal*, *58*(3), 932-970.

Starke, G., & Ienca, M. (2024). Misplaced trust and distrust: How not to engage with medical artificial intelligence. Cambridge Quarterly of Healthcare Ethics, 33(3), 360-369.

Strauss, A., & Corbin, J. (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). Sage Publications, Inc.

Tams, S., Thatcher, J. B., & Craig, K. (2018). How and why trust matters in post-adoptive usage: The mediating roles of internal and external self-efficacy. The Journal of Strategic Information Systems, 27(2), 170-190.

Tang, L., Li, J., & Fantus, S. (2023). Medical artificial intelligence ethics: A systematic review of empirical studies. Digital Health, 9, 20552076231186064.

Teodorescu, M., Morse, L., Awwad, Y., & Kane, G. (2021). Failures of Fairness in Automation Require a Deeper Understanding of Human-ML Augmentation. MIS Quarterly, 45(3), 1483–1500.

Thornton, P. H. (2004). Markets *from Culture: Institutional Logics and Organizational Decisions in Higher Education Publishing*. Stanford University Press.

Thornton, P. H., & Ocasio, W. (2008). Institutional Logics. In *The SAGE Handbook of Organizational Institutionalism* (pp. 99–128). SAGE Publications Ltd.

Thornton, P. H., Ocasio, W., & Lounsbury, M. (2012). *The Institutional Logics Perspective*. Oxford University Press.

Timmermans, S., & Tavory, I. (2012). Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory*, *30*(3), 167-186.

van den Broek, E., Sergeeva, A., & Huysman, M. (2021). When the Machine Meets the Expert: An Ethnography of Developing AI for Hiring. *MIS Quarterly*, *45*(3), 1557–1580.

van Harreveld, F., Rutjens, B. T., Rotteveel, M., Nordgren, L. F., & van der Pligt, J. (2009). Ambivalence and decisional conflict as a cause of psychological discomfort: Feeling tense before jumping off the fence. *Journal of Experimental Social Psychology*, *45*(1), 167–173.

Weick, K. E. (1995). *Sensemaking in organizations*, vol. 3. Thousand Oaks, CA: Sage.

Yang, R., & Wibowo, S. (2022). User trust in artificial intelligence: A comprehensive conceptual framework. *Electronic Markets*, *32*(4), 2053-2077.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). SAGE.

Appendices

Appendix A: Review Summary: AI, Institutional Logics, and Automation/Augmentation

Theme	Key Insight	Institutional Perspective	Key Takeaway	Studies
Treatment of Automation vs. Augmentation	Existing studies generally treat automation and augmentation as distinct functional strategies or technical approaches.	Institutional underpinnings are typically absent or implied; emphasis is placed on task type, cognition, or technical fit.	Lack of theorization of automation and augmentation as enactments of institutional logics.	Raisch & Fomina (2025); Shao et al. (2024); Berente et al. (2021)
Level of Analysis	Studies offer rich insights at either micro (individual trust, learning) or macro (labor substitution) levels.	Micro-level studies emphasize psychological or behavioral dynamics; macro-level studies engage structural trends.	Insufficient attention to the organizational level, where institutional logic conflicts are enacted in practice.	Agrawal et al., (2023), Shao et al. (2024); Glikson & Woolley (2020); Lei & Kim (2024)
Use of Institutional Logics	Some studies acknowledge institutional considerations, often through governance or ethical frames.	References to institutional influence are present but typically not structured as competing or conflicting logics.	Institutional logics are underutilized as an explanatory mechanism for contradictions in AI deployment.	Berente et al. (2021, 2023); Lei & Kim (2024)

Appendix B: Interview Guide and Participant Information

Leading Questions • Could you please describe your role in your department? • What is BlueSky's overall attitude toward artificial intelligence (AI)? • How do you personally feel about AI?
Questions About AI Use • Why did your department decide to implement AI? • Has AI changed the way you work? If so, how? • What are the benefits and challenges have you experienced with AI in your work? Why do you think that is? • How does BlueSky manage the balance between the benefits and challenges of AI? How does this affect your attitude or behavior toward AI?
Questions About Trust and Distrust in AI • In what situation would you feel confident allowing AI to handle tasks with minimal or no human intervention? Why? • When would you hesitate to allow AI to handle tasks without human oversight? Why?
Questions About Reconciliation Practices • How do you reconcile the conflicts between your confidence and hesitation towards AI in your workplace? • Do you believe your approach to reconciling these conflicts aligns with BlueSky's overall stance on AI? Why or why not?

Participant Information

- Customer Retention Management — Key Duty: Understanding customer behavior, improving retention rates, and enhancing the overall customer experience.

  Interview Duration: Head of Department (63 min); Analyst 1 (47 min); Analyst 2 (35 min); Analyst 3 (40 min); Analyst 4 (33 min); Analyst 5 (50 min).

- Revenue Management — Key Duty: Optimizing income through strategic pricing, inventory control, and demand forecasting.

  Interview Duration: Head of Department (52 min); Team Leader (38 min); Analyst 1 (64 min); Analyst 2 (45 min); Analyst 3 (48 min); Analyst 4 (41 min); Analyst 5 (33 min).

- Data Science Lab — Key Duty: Working on various innovative projects and analyses to leverage data for improving operations.

  Interview Duration: Head of Department (78 min); Data Scientist 1 (47 min); Data Scientist 2 (48 min); Data Scientist 3 (35 min); Data Scientist 4 (40 min); Fraud Prevention Lead (39 min); Fraud Analyst (48 min).

  *One Data Analyst, to be assigned to department, with Piloting Background — (50 min)

Note: We encouraged participants to provide examples when answering our questions during the interviews.

Appendix C: Coding Structure

Open Coding	Axial Coding	Thematic Coding
• Prioritizing data-driven decisions over heuristics • Maintaining analytical independence • Grounding business analysis in academic research	Scientific objectivity (CRM)	Instrumental-analytic logic
• Recognizing AI's superior capabilities for pricing • Valuing AI pricing insights over human intuition • Minimizing human interference in AI-driven pricing	Revenue maximization (RM)	
• Standardizing processes for operational accuracy • Integrating data for improved efficiency • Pursuing efficiency through standardized AI process	Process standardization (DSL)	
• Aiming to design AI models grounded in statistical validity (CRM) • Aiming to produce unbiased, analytically robust outputs (CRM & DSL) • Aiming to outperform historical benchmarks through algorithmic recommendations (RM) • Aiming to establish replicable, data-driven decision procedures (CRM & RM) • Aiming to replace manual processes with integrated, automated systems (DSL)	To systematizing data-driven decision making	
• Framing as analytical purists committed to methodological rigor (CRM) • Seeing as commercial optimizers driving financial outcomes (RM) • Identifying as technical architects solving systemic inefficiencies (DSL)	Positioning as system-centric experts	
• Integrating contextual understanding with AI insights • Coordinating AI optimization with practical realities • Leveraging human intuition alongside AI analytics	Business-pragmatism (CRM)	Contextual-normative logic

• Ensuring ethical alignment in AI-driven decisions • Upholding customer values in AI applications • Applying ethical judgment in value-based decision	Social responsibility (RM)	
• Task complexity impedes AI standardization • Evolving environments require adaptability • Leveraging human adaptability in uncertain conditions	Adaptive flexibility (DSL)	
• Aiming to generate actionable insights aligned with campaign goals (CRM) • Aiming to ensure revenue strategies remain ethically appropriate and publicly defensible (RM) • Aiming to adapt AI systems to real-time disruptions and cross-functional constraints (DSL)	To ensure contextual and ethical fit	
• Framing as pragmatic enablers focused on campaign relevance and usability (CRM) • Seeing themselves as ethical guardians protecting brand reputation and public trust (RM) • Positioning themselves as experienced coordinators navigating operational complexity (DSL)	Positioning as context-sensitive stewards	
• Belief in AI's objectivity to deliver unbiased results • Belief in AI's analytical capacity to maximize revenues • Belief in AI's capacity to process and integrate large-scale multisource data for cargo standardization	Affirmed confidence in AI	Intensity of trust in AI
• Reliance on observed improvements to build confidence • Expectations shaped by experiential validation • Anticipation of limited performance improvement	Guarded confidence in AI	
• Concern about AI failing to produce actionable marketing strategies • Concern about AI ignoring ethical or societal consequences in pricing • Concern about AI failing to account for the contextual nuances in cargo management	Pronounced wariness of AI	Intensity of distrust in AI
• Limited vigilance toward AI misinterpreting irregularities or errors • Dormant concerns about AI embedding biases • Lack of concern about unnoticed AI errors causing operational issues	Latent concerns about AI	
• Reflectively interpreting AI outputs beyond surface accuracy • Enhancing and applying critical thinking in AI engagement • Drawing on experience to anticipate consequences of AI-driven actions	Mindful evaluation	Reconciliation practices
• Coordinating cross-functionally to preempt AI breakdowns • Running human–AI trials to define performance limits • Using feedback loops to uncover contingencies and refine models	Proactive safeguarding	

Appendix D: Methodological Design

❖ **Clarifying Conceptual Foundation**

Our initial inductive analysis revealed conflicting attitudes towards AI implementation within BlueSky, characterized by, for instance, both reliance on and hesitation about AI. Therefore, we anchored our data analysis in the thematic concepts of automation, augmentation, trust, and distrust in AI implementation, as outlined in our conceptual background. The institutional logics perspective was introduced later in the analysis as we sought to account for the patterned, cross-case regularities in how these attitudes were organized, rather than serving as an ex ante coding template.
❖ **Coding and Data Structure**
We followed Corbin and Strauss' (2008) guidelines for coding and developing a data structure, framing our analysis around our conceptual foundation. This means our data analysis was exploratory in nature. Our open coding remained true to the original meanings provided by our informants. As the analysis progressed, we identified similarities and differences among the open codes (Strauss & Corbin, 1998), allowing us to reduce the number of codes. Further analysis and extended discussions among the authors enabled a deeper level of abstraction, forming the basis of axial coding, which was then integrated into thematic dimensions (see Appendix C).
❖ **Conducting Cross-Case Synthesis**
Following Langley's (1999) strategy for making sense of qualitative data, we created a cross-case comparison table (see Table 2) to capture key institutional variations in AI implementation at BlueSky. The table illustrates the components of each institutional logic driving either trust or distrust in AI (e.g., scientific objectivity and business pragmatism), the AI implementation paradigms (i.e., automation and augmentation), and the reconciliation practices (i.e., mindful evaluation and proactive safeguarding). This step depicted how institutional logics were enacted across the three nested cases, providing a solid empirical foundation for further theorizing.
❖ **Theorization**
To theorize from our empirical analysis, we engaged with the institutional logics literature to link identified elements to broader theoretical reasoning (see Section 5). This process refined our findings into a higher level of abstraction and enabled analytical generalization. We developed a framework representing the constitutions and reconciliations of multiple competing institutional logics in AI implementation within a knowledge-intensive, business context.

Appendix E: Case Analysis—Chain of Evidence

Nested Case 1: AI Implementation in the Customer Retention Management Department
Instrumental-analytic logic: Scientific objectivity + to systematize data-driven decision-making + positioning as system-centric experts *Evidence*: CRM analysts designed statistically valid models and emphasized replicability, often distancing themselves from business-side input. They viewed AI as a rigorous alternative to heuristic decisions. *Key takeaway*: AI was leveraged as a vehicle to formalize analytical standards and reinforce the department's identity as autonomous experts in evidence-based decision-making. **Contextual-Normative logic:** Business pragmatism + to ensure contextual and ethical fit + positioning as context-sensitive stewards. *Evidence*: Analysts adjusted AI-generated segments based on marketing input, re-running models and supplementing outputs with descriptive analysis to improve business interpretability.

Key takeaway: Sustained relevance of AI depended on continuous negotiation between analytical outputs and practical business needs. **Trust in AI:** Affirmed, sustained *Evidence*: Analysts maintained belief in AI's scientific objectivity, even as implementation challenges emerged. *Key takeaway*: Affirmed trust in AI's analytical processing endured despite contextual misalignments. **Distrust in AI:** Latent, pronounced. *Evidence*: Analysts expressed minimal concern about AI misinterpreting data irregularities during customer segmentation. However, they expressed concerns about its inability to account for campaign-specific nuances. *Key takeaway*: Salient distrust arose not from failure, but from misfit with use-case needs **Mindful evaluation**: Interpreting AI through marketing relevance. *Evidence*: Analysts critically assessed clustering outputs for campaign fit, questioning assumptions and reframing results for business use. *Key takeaway*: Cognitive engagement enhanced output relevance by aligning technical logic with marketing needs. **Automation:** Delegation of complex analytic tasks to AI. **Augmentation:** Human–AI collaboration for contextualization and business alignment of results. **Proactive safeguarding**: Iterative redesign with business partners. *Evidence*: Marketing staff helped revise models and review segments; misaligned outputs were reworked collaboratively. *Key takeaway*: Cross-functional revisions kept AI aligned with evolving business goals.
Nested Case 2: AI Implementation in the Revenue Management Department
Instrumental-analytic: Revenue maximization + to systematize data-driven decision-making + positioning as system-centric experts. *Evidence*: RM analysts embraced AI as a tool to detect pricing opportunities that exceeded human assumptions. They tested algorithmic pricing on competitive routes and monitored margin improvements to justify system-wide deployment. *Key takeaway*: AI was accepted as a high-performance analytical partner when it could demonstrably outperform legacy approaches and support financial targets. **Contextual-Normative:** Social responsibility + to ensure contextual and ethical fit + positioning as context-sensitive stewards. *Evidence*: During regional disruptions, analysts paused automated pricing and adjusted thresholds to avoid public backlash. This response reflected organizational values around customer fairness and reputational care. *Key takeaway*: The deployment of AI was actively constrained to uphold broader social expectations, especially in morally sensitive situations. **Trust in AI: Transition from** guarded to affirmed and sustained *Evidence*: Initial route tests with mixed beliefs led to broader adoption after success. Trust was sustained even after algorithmic failure in ethical scenarios. *Key takeaway*: Trust had to be earned through empirical performance. **Distrust in AI:** Latent, pronounced. *Evidence*: Analysts expressed minimal unease about AI autonomously optimizing pricing decisions for routine situations. Analysts expressed heightened wariness in reaction to AI's unfair pricing adjustments during emergencies. *Key takeaway*: Even with trust, value-alignment failures triggered sharp pushback. **Automation:** Delegation of routine pricing decisions to AI.

Augmentation: Human AI-collaboration to optimize pricing during crises or unpredictable events.

Mindful evaluation: Ethical scrutiny of AI recommendations.
Evidence: Analysts reviewed AI outputs through ethical lenses, discussing fairness and reputational implications.
Key takeaway: Reflective judgment enabled moral calibration of automated decisions.

Proactive safeguarding: Adjustment protocols and scenario testing.
Evidence: Mentorship sessions and AI-human comparisons introduced to test and constrain AI pricing.
Key takeaway: Deliberate interventions shaped AI behavior to prevent value misalignment.

Nested Case 3: AI Implementation in Data Science Lab

Instrumental-analytic: Process standardization + to systematize data-driven decision-making + positioning as system-centric experts.

Evidence: DSL engineers sought to unify fragmented workflows through AI-driven forecasting models, emphasizing efficiency, replicability, and structured decision logic.
Key takeaway: AI was implemented as a structural solution to eliminate inconsistency and promote cross-functional alignment in decision processes.

Contextual-Normative: Adaptive flexibility + to ensure contextual and ethical fit + positioning as context-sensitive stewards
Evidence: When forecasts failed due to unmodeled operational contingencies (e.g., new aircraft or variable loader skills), the team introduced cross-department coordination and updated model assumptions.
Key takeaway: Human oversight and iterative adaptation were essential to maintain operational relevance in dynamic, multi-actor settings.

Trust in AI: Affirmed, sustained.
Evidence: DSL engineers believed in AI's capability to integrate complex datasets. Affirmed trust was sustained even after witnessing AI's inherent limitations in dynamic environments.
Key takeaway: AI was seen as a dependable engine for structured environments.

Distrust in AI: Latent, pronounced.
Evidence: Absence of worry over AI errors going undetected and causing operational issues. Acute concern about model inflexibility in unanticipated disruptions, as the AI failed to adapt to changing aircraft types or inter-team constraints.
Key takeaway: Distrust surged with exposure to exceptions AI couldn't model.

Automation: Delegating multi-source data integration to AI under stable operational conditions.

Augmentation: Human-AI collaboration for context-sensitive decision processes.

Mindful evaluation: Diagnostic interpretation of AI blind spots.
Evidence: Analysts treated anomalies as cues to reassess model assumptions and data scope.
Key takeaway: Reflective use of breakdowns deepened understanding of AI limitations.

Proactive safeguarding: Embedded cross-departmental feedback channels and model repair processes.
Evidence: Operational staff flagged edge cases; DSL teams iteratively refined model rules.
Key takeaway: Continuous input from frontline users strengthened system robustness.

Appendix F. Elemental categories of institutional logics

Institutional Logics	Categories	Characterization	Representative quote
Instrumental-analytic	**Principles**	Scientific Objectivity	"*(when deciding) we need to avoid social bias, and our own personal bias that we have."* (CRM Head analyst)
		Process Standardization	*"The whole organization is now in this process, the data scientists built a model, let's come together to see how we can improve it. Because how can the algorithm know as a parameter, who is loading the aircraft? Some loaders are highly skilled, stacking cargo like Lego, while others aren't as proficient, or who is doing the check in load control and by what conditions? This fragmentation is the black box, the uncertainty of the model right now." (Data Scientist)*
		Revenue Maximization	*This (the implementation of the AI platform), I'd say, is a true evolution, a turning point for us in terms of revenue growth. Previously, analysts were often more conservative in their pricing, potentially leaving money on the table.*
	Goals	To institutionalize data-driven decision making	*"Most of the rules used in previous campaigns were simple heuristics, and some were even worse speculations. Heuristics can be useful, but you cannot rely on such simplistic views for designing modern campaigns"* (CRM Head analyst).
	Identity	System steward, analytic authority	"Because we are a bit of scientists, we like things to... We don't look at the results but the methodology, we have also this fetish" (CRM Head Analyst)
	Principles	Social Responsibility	*"I tell them [the analysts on my team], it's their responsibility to monitor it [the AI]…We have a social responsibility."* (RM Team Leader)

Contextual -normative		Adaptive Flexibility	*"We don't have hundreds of people working in a standardized vertical structure where a few individuals have a full picture of the entire operation and can give the right instructions to everyone. We need to be more flexible."* (Data Scientist)
		Business Pragmatism	*"This work is not done to improve a certain statistical range. It is not done to satisfy our scientific libido and theoretical satisfaction. We maybe sometimes do it for that. But in reality, the company does not pay us for that. Neither the project nor the request came to satisfy academic curiosity… Because we are not a scientific institute, we are a company, we can't do these things."* (Head of CRM (Analysts)
	Goals	Ensuring contextual and ethical fit	*"We have a social responsibility. If the analyst isn't alert to readjust rules and thresholds - that is, to stop, close, and reopen pricing - the system can't understand all these nuances by itself. So, in reality, the analyst's role is still crucial."* (RM Team Leader)
	Identity	Ethical steward-context sensitive	*"I tell them [the analysts on my team], it's their responsibility to monitor it [the AI]…We have a social responsibility."* (RM Team Leader) *"Obviously on sustainability and reducing our environmental footprint… we are the only airline in the country and one of the few in Europe to carry out a systematic program of using SAF on our flights… This is an important step towards a more visible solution to improve the environmental footprint in aviation."* (CEO)