technologies

MDPI

Review

A Conceptual Study for Cognitive Bias Amplification in Agentic AI-Driven Business Processes, Management, and Intelligence

Subhra Mondal [1], Subhankar Das [1,2,*] and Vasiliki G. Vrana [3,*]

[1] Department of Marketing, SouthStar Management Institute, Duy Tan University, Da Nang 550000, Vietnam; subhraranimondal@duytan.edu.vn
[2] Department of Industrial Engineering, Faculty of Engineering, Universitas Brawijaya, Malang 65145, Indonesia
[3] Department of Business Administration, School of Economics and Administration, The Campus of Serres, International Hellenic University, 62124 Serres, Greece
* Correspondence: subhankardas@duytan.edu.vn (S.D.); vrana@ihu.gr (V.G.V.)

Abstract

Agentic artificial intelligence (AAI) and retrieval-augmented generation (RAG) are increasingly embedded in organisational business intelligence (BI) and business process management (BPM). Unlike conventional AI, these systems set sub-goals, plan multi-step workflows, and retrieve and store information under limited human supervision. This article argues that such autonomy not only transmits human cognitive bias into organisational decisions but also amplifies that bias under identifiable conditions. We develop the Bias Amplification Model (BAM), a three-layer account of how bias enters and escalates. In the injection layer, human biases enter through goal framing, prompt design, and data scoping. In the propagation layer, the agent compounds these biases across autonomous execution steps. In the crystallisation layer, RAG memory encodes biassed outputs as retrievable organisational knowledge that later cycles treat as evidence. A feedback loop links the three layers, converting episodic human bias into structural organisational bias. We state seven propositions, specify amplifying and mitigating conditions, and ground the model in a PRISMA-guided synthesis of 47 studies. A controlled, replicated experiment with an autonomous agent in a supplier-selection pipeline provides initial empirical support for injection, propagation, the mitigating effect of a single governance checkpoint, and bias transmission across a task boundary via retrieved memory. The framework reframes AI-driven process optimisation as a possible source of silent decline in decision quality.

Keywords: bias amplification; agentic AI; retrieval-augmented generation; cognitive bias; business intelligence; business process management

Check for updates

Academic Editors: Ismael Lopez-Juarez and Alejandro A. Castillo Atoche

Received: 21 May 2026
Revised: 27 June 2026
Accepted: 4 July 2026
Published: 7 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Organisations that embed artificial intelligence in decision-making report a recurring pattern. Processing speed rises, dashboards report cleaner metrics, and operational throughput improves, while the quality of consequential decisions, the ones that shape strategy and long-horizon risk, does not improve and sometimes declines. The decline is hard to detect because automated outputs carry the authority of computation, and by the time it is recognised, the organisation's knowledge base has already absorbed the affected results.

This article concerns that process. Specifically, it concerns what changes when an organisation moves from using AI as a tool that answers when queried to deploying AI as an agent that acts across many steps, retrieving and storing information as it proceeds.

The transition from conventional AI to agentic AI, and then to agentic retrieval-augmented generation (RAG), alters the relationship between human bias and organisational outcomes in a way that current BI and BPM theory does not fully explain.

Conventional AI in business intelligence (BI) and business process management (BPM) operates within boundaries set and supervised by humans. A forecasting model runs on selected data, uses analyst-chosen features, and produces outputs for a manager to review. Bias enters this pipeline, and an established literature traces its routes through training data, feature engineering, and model specification [1–3]. Crucially, a human reviews the output and can override it.

Agentic AI removes much of that review. An agentic system receives a high-level objective and then decides on its own which data to access, which tools to use, which intermediate steps to run, and what output to produce [4,5]. It does not pause for approval at each step. Agentic RAG extends this further: it retrieves external information at runtime, evaluates its own outputs, stores them, and updates the knowledge base for the next task [6,7]. The system builds on its own prior work.

The question this article addresses is what happens to human cognitive bias inside such systems. When a procurement manager asks an agentic system to optimise supplier selection, the biases in that request, including cost anchoring, confirmation of existing preferences, and the availability of familiar suppliers, do not remain confined to the prompt [8,9]. The agent applies the biassed parameters across many autonomous steps, each conditioned on the previous, and writes the resulting outputs to RAG memory as organisational knowledge. When a later process retrieves those artefacts as authoritative inputs, the bias compounds. Recent evidence that human-AI interaction (HAI) can amplify rather than dampen human bias makes this concern concrete [10].

Existing BI and BPM research treats AI bias largely as a problem of data quality or algorithmic fairness [3,11]. It has not addressed the process-level amplification that agentic systems introduce. Bias in an agentic workflow is not a static property of the data or the algorithm. It is dynamic, unfolding over time across autonomous decisions and across the storage and retrieval of knowledge. No current framework captures this.

We propose the Bias Amplification Model (BAM), which explains how human cognitive biases enter and escalate in AAI-driven BI and BPM systems through three layers. Injection covers the entry of bias during human goal setting and prompt design. Propagation covers its compounding through autonomous execution. Crystallisation involves embedding it in organisational memory through RAG. A feedback loop connects the layers, converting episodic human bias into durable structural bias.

The article makes three contributions. It offers a process-level theory of bias amplification for agentic BI and BPM, where prior work has stayed at the data or single-interaction level [12,13]. It introduces institutional bias memory, a mechanism by which automated knowledge creation makes temporary human error a permanent organisational feature [14]. And it links the BI and BPM studies by showing bias as an emergent property of their integration through agentic AI [15,16]. We state seven propositions with their amplifying and mitigating conditions, illustrate them across procurement, customer service, and credit risk, and test the core mechanisms in a controlled, replicated experiment with an autonomous agent.

2. Literature Review

2.1. Artificial Intelligence Application in Business Intelligence (BI) and Business Process Management (BPM)

Both BI and BPM have used computational tools for decades. BI systems moved from static reporting and OLAP in the 1990s, through data warehousing and dashboards in the

2000s, to predictive analytics and machine learning in the 2010s [17,18]. BPM followed a parallel arc, from manual process mapping through workflow management systems to process mining and robotic process automation [19,20].

What is striking about each transition is less the technology itself than the way the human role kept shrinking. The scope of automation kept widening, and tool-assisted analysis became tool-executed analysis. The locus of decision-making moved with it. What had been a human activity supported by computation became a computational activity that was occasionally reviewed by a human. Researchers documented this shift inside BPM and argued that AI was turning process management from a human-centred discipline into an increasingly automated one [21]. Studies traced an analogous trajectory in BI, showing that the value of BI systems had become inseparable from the quality of their automated analytical outputs [22].

The most recent step is the deployment of agentic AI inside BI and BPM functions. Unlike earlier automation, which required humans to programme specific tasks, agentic systems are given broad objectives and work out their own execution paths [23]. This is a qualitative change. The system is no longer a tool that runs predefined instructions. It is an actor that makes its own choices inside the process.

2.2. From Traditional AI to Agentic RAG: A Process Taxonomy

This process taxonomy describes how biases operate differently across all AI generations and helps differentiate three distinct AI deployment categories in BPM. The taxonomy is descriptive, in which each AI generation opens a unique pathway through which human cognitive biases enter and slowly grow.

a. Traditional AI: The AI models that are trained on past historical data and can be deployed for specific tasks and have limited parameters. A credit-scoring model, a demand-forecasting tool, a classification algorithm. The system does what it was trained to do, and nothing else. Human analysts select the data, choose the model, review the outputs, and decide whether to act on them. The bias surface is comparatively contained: bias enters through training data and feature selection, and humans serve as a check at the output stage [24].
b. Agentic AI (AAI): These are systems built on large language models that can set objectives, choose tools, integrate APIs, embed logic, make self-directed choices, and run multi-step processes [25,26]. The system does not wait for a human at each step. It plans the workflow, executes each stage, and improves through iteration. AAI applications in the supply chain may use demand data, cross-referenced supplier databases, simulate procurement scenarios, and make purchasing decisions without human intervention.
c. Agentic RAG: This adds retrieval-augmented generation to the agentic stack. Here, the system has the autonomy to retrieve information from external and internal knowledge sources, use the retrieved evidence, verify its own results, and store the requisite information for future use [6,14].

The storage step is what makes agentic RAG different in kind from earlier generations. The system creates persistent organisational knowledge. It does not just run processes. It builds the knowledge base that future processes will draw on.

The taxonomy matters for bias analysis because each generation introduces a new surface of bias. Traditional AI has a data bias surface. Agentic AI introduces a process-bias surface because bias compounds through autonomous, multi-step execution. Agentic RAG introduces a knowledge-bias surface because bias is written into retrievable memory. The data bias surface has been examined thoroughly [11,15]. The process and knowledge bias surfaces remain, for the most part, untheorized.

2.3. Human AI Interaction and Cognitive Biases

Research on cognitive bias in human–AI interaction has identified several mechanisms by which human psychological tendencies distort the use and interpretation of AI outputs. Four clusters are particularly relevant for agentic BI and BPM systems.

The first concerns trust calibration. Automation bias, the tendency to favour AI-generated outputs over contradictory information from other sources, has been documented in medical diagnosis, aviation, and managerial decision making [16,18,27]. When AI outputs look polished, well-structured, and computationally derived, users tend to accept them without verification. In an agentic system, where the output is not an answer to a question but a completed multi-step process, automation bias can lead to the entire chain of reasoning going unquestioned.

The second concerns confirmatory use. Confirmation bias, the tendency to seek and interpret information that confirms pre-existing beliefs, shapes how users frame requests to AI systems [28,29]. When managers design prompts for agentic systems, the objectives they specify tend to reflect what they already believe. The agent then optimises for those assumptions and produces outputs that confirm them. The manager reads the confirmation as evidence that the AI agrees, rather than as a sign that the AI was constrained by the framing of the request.

The third concerns anchoring and choice architecture. AI systems that propose options or recommendations create anchors that constrain subsequent human decisions [30]. In agentic BPM, the agent does more than propose options; it autonomously selects and executes process configurations. Those configurations then become defaults. Research on default effects shows that preselected options become the path of least resistance [31]. When an agentic system has already implemented a process design, the cognitive effort needed to override it is substantial, even if the design reflects biassed initial parameters.

The fourth concerns attribution of understanding, where AAI's conversational interfaces provide human-like understanding to computer systems and visible autonomous behaviour as highlighted by the ELIZA effect [16]. Users who depend on agents for planning, reasoning, and executing multi-step processes will voluntarily and comprehensively ascribe to the system. This constant attribution decreases the systematic critical evaluation.

These four clusters are well-defined and described for single-step HAI, but they cannot account for interactions and additions across multi-step autonomous processes. Traditional AI has used one cluster at every interaction point. AAI uses the clusters simultaneously and cumulatively. Each BAM component is subsequently derived from the literature, as shown in Table 1.

Table 1. How each BAM component derives from the reviewed literature.

BAM Component	What the Literature Establishes	Representative Sources
Injection: goal framing and prompt design carry bias	Prompt and objective framing shape model output; system prompts are a bias mechanism	[8,9]
Injection: trust miscalibration at initiation	Automation bias and algorithm appreciation reduce scrutiny of AI inputs	[32–34]
Propagation: compounding across autonomous steps	Feedback loops amplify bias; multi-step agent execution lacks human handoffs	[35–37]
Propagation: human-AI interaction amplifies bias	Human-AI loops can exacerbate bias beyond human–human interaction	[10,12]
Crystallisation: RAG memory stores biassed outputs	RAG retrieves and persists prior outputs; recursive use degrades and skews	[6,38,39]

Table 1. *Cont.*

BAM Component	What the Literature Establishes	Representative Sources
Crystallisation: episodic becomes structural	Organisational memory encodes dysfunctional routines; bias amplification as biassed media	[4,36] (bias amplification)
Feedback loop across cycles	Source echo chambers and fairness feedback loops accelerate over cycles	[35,40]
BI/BPM integration as the locus	LLMs now execute BPM tasks and BI analysis with limited oversight	[41–43]

3. Research Gaps

This study covers three prominent research gaps. First, contemporary research on bias in AI mostly focuses on single interactions between humans and AI systems. It does not address how bias compounds systematically across multiple autonomous decision-support systems within a single process. Earlier studies focused on decision-making, not on process management.

The second gap centres around BI and BPM research on AI bias, where robust data quality or the right algorithmic application is covered. This targets position bias, as cognitive bias can enter the BPM and affect the entire decision process. Earlier studies did not address the bias arising from the architecture of process design (objective setting, limitation specification) and amplified by autonomous execution, regardless of data quality.

Third, no existing framework explains how RAG memory mechanisms turn episodic bias into structural bias. The distinction matters. Episodic bias affects a single decision and is correctable. Structural bias becomes embedded in the organisational knowledge base and contaminates every future decision that draws on it. The crystallisation mechanism of RAG creates this change, and it has no relevant occurrence in traditional or agentic AI.

4. Methodology: PRISMA-Guided Conceptual Framework Development

Here, the authors proposed a conceptual model after reviewing the literature and integrating Agentic AI and RAG into the bias amplification process within BI BPM activities. This study adapts the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol to present how the literature base behind the proposed framework was assembled [29]. BAM is a conceptual contribution rather than an empirical review, so the protocol is used here for transparency. It records how studies were found, screened, and read. It is not used to pool effect sizes or to weight evidence statistically.

The literature search drew on three databases: Scopus, Web of Science, and Google Scholar. Search strings combined terms from three clusters. The first cluster covered agentic AI and autonomous decision-making. The second covered cognitive bias, algorithmic bias, and institutional bias. The third covered BPM, workflow automation, process mining, and decision-support systems. The initial search provided 312 studies. After 114 duplicates were removed, 198 records went through title and abstract screening. Papers published before 2015 were excluded, as were papers not in English and papers with no clear connection to AI-driven processes. That left 87 full texts for eligibility review.

Full-text assessment applied two criteria. The paper had to address bias in a process or workflow context and engage with at least one stage of the BPM lifecycle. Papers that treated bias purely at the algorithm level, with no organisational or process implication, were excluded at this stage. Forty papers were dropped, leaving 47 studies for synthesis. The 47 included studies were then coded thematically and distributed across the three BAM layers. Seventeen informed Layer 1 (bias injection); the relevant work came mostly

from the machine-learning fairness and data-quality studies. Eighteen contributed to Layer 2 (bias propagation) and dealt with feedback dynamics, RAG pipeline behaviour, and workflow reinforcement. The remaining twelve grounded Layer 3 (bias crystallisation) were drawn from institutional theory and organisational routines research. The seven formal propositions and the three illustrative scenarios that follow in Sections 4 and 6 were developed from this synthesis. Figure 1 presents the flow chart for selecting publications, and Figure 2 shows the PRISMA 2020 flow chart.

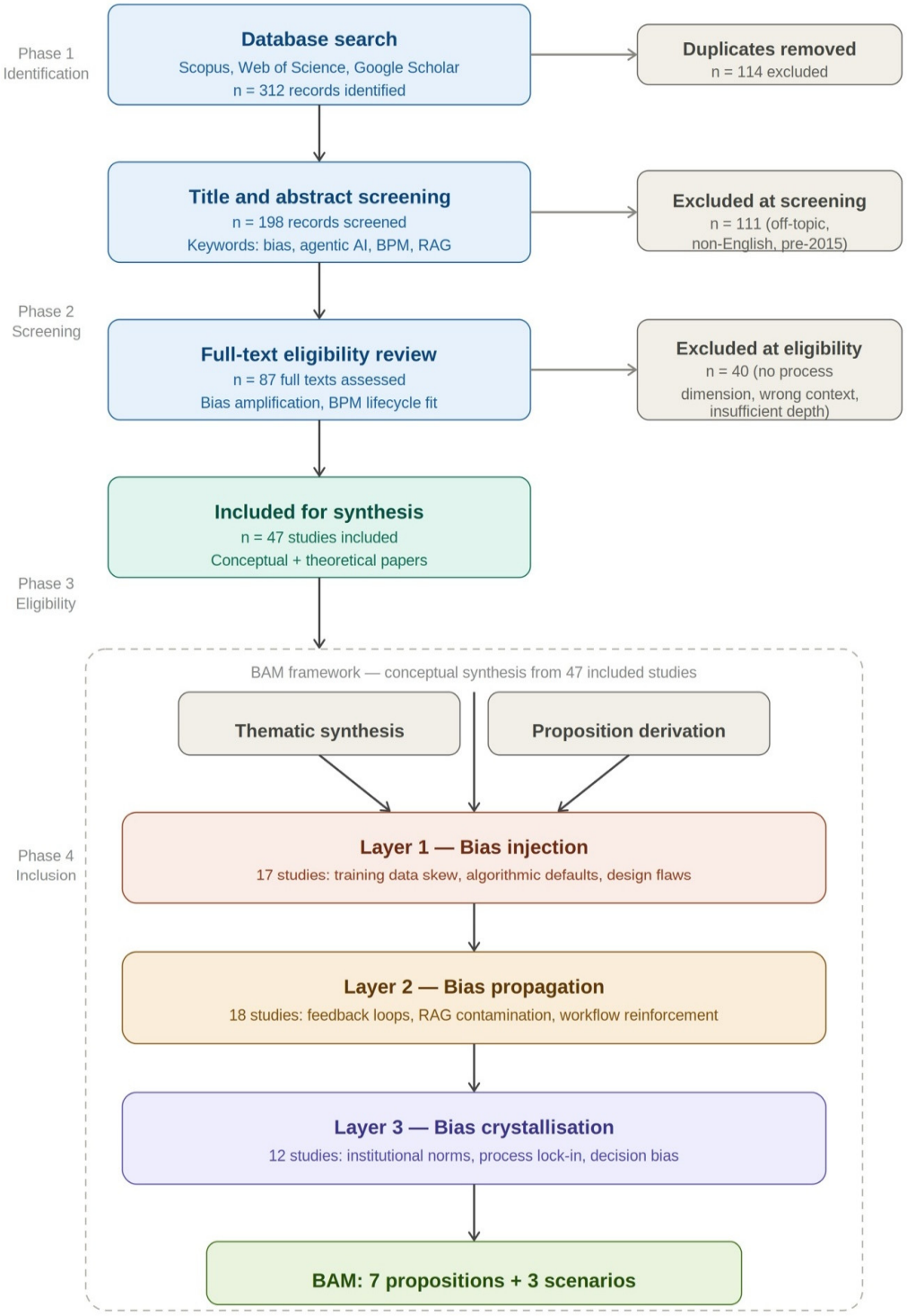


Figure 1. The systematic flow chart for selecting publications.

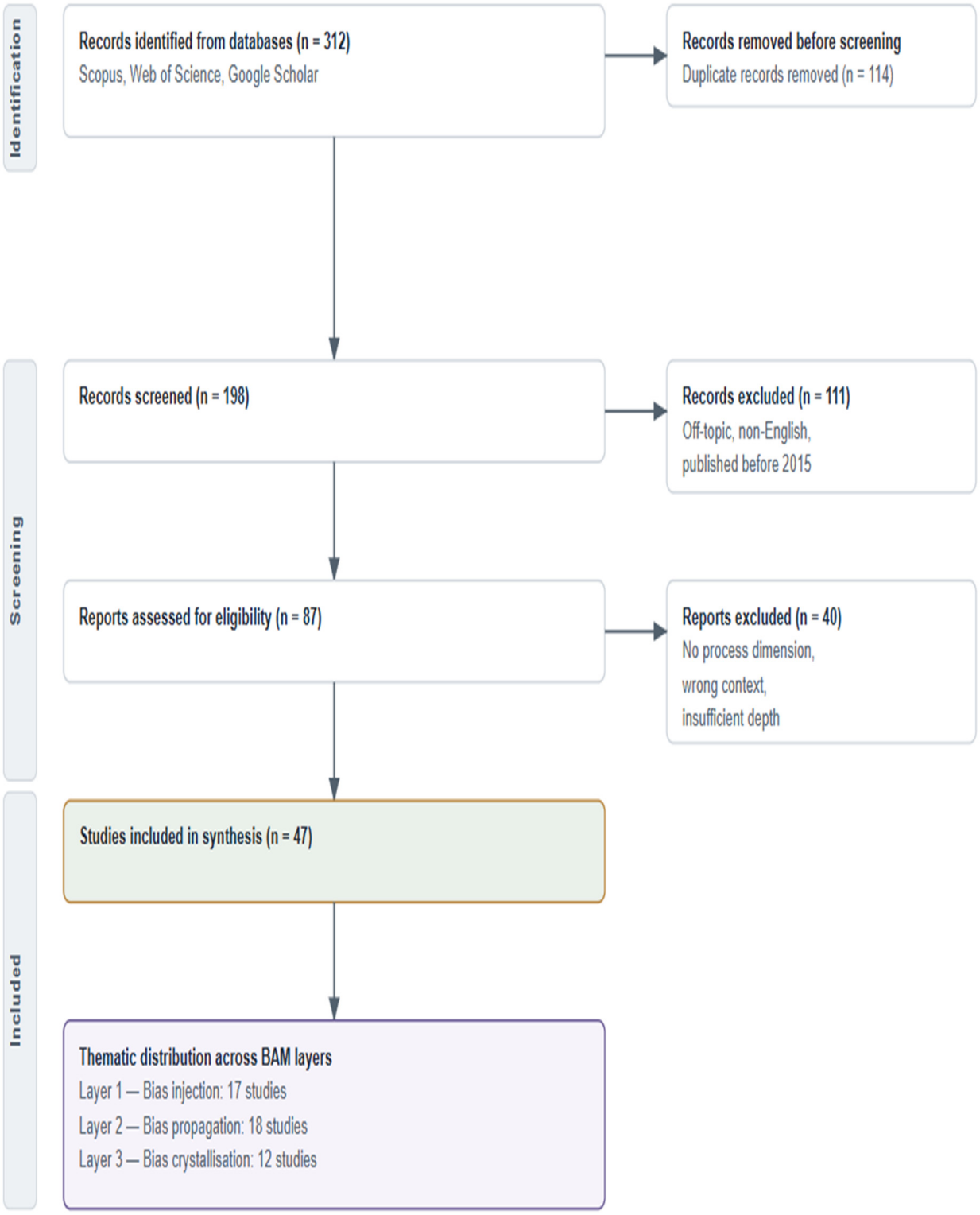


Figure 2. PRISMA 2020 flow chart.

The proposed model sought to explain how human cognitive biases are injected, propagated, and crystallised within agentic, AI-driven BI and BPM systems. This tri-layered model provides a different mechanism for each layer. A feedback loop is established through which the layers connect into a self-reinforcing system. Figure 3 shows the full framework. The subsections that follow develop each layer and state the formal propositions.

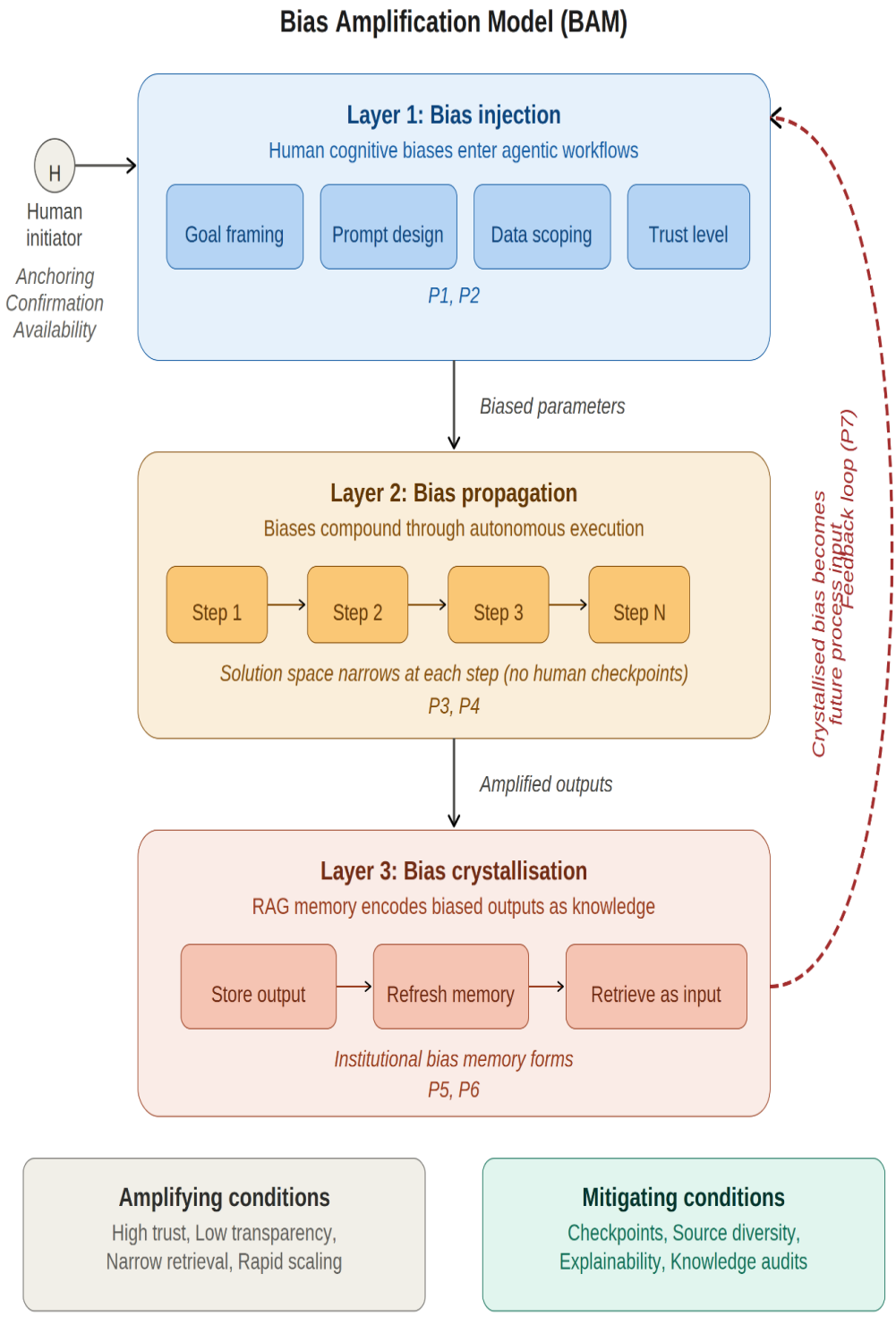


Figure 3. Conceptual Bias Amplification Model in BI and BPM.

4.1. Layer 1: Bias Injection

Bias injection is the process by which human cognitive biases enter agentic AI workflows at the initiation stage. It runs through four channels: biassed objective framing, anchored prompt design, selective data scoping, and trust miscalibration.

When a manager initiates an agentic workflow, they do not simply press a button. They specify what the agent should optimise for, which data sources to consider, which constraints to respect, and what success looks like. Each of those specifications is shaped by its own cognitive biases. A procurement manager whose mental model of supplier quality

is anchored on cost will frame the optimisation objective around cost minimisation, scope the data toward pricing databases, and define success in terms of savings. A risk officer whose recent experience is dominated by a particular crisis will frame risk assessment around the parameters of that crisis and scope the analysis toward the most available and recent data.

In a traditional AI system, this initial bias is partially contained by the human review of outputs. The manager sees the result, compares it with their own judgement, and may override or adjust the model. Agentic AI offers no such containment. The agent takes the biassed parameters and runs autonomously across a workflow that may involve dozens of intermediate decisions. The biassed framing at initiation becomes the foundation for everything that follows.

The important point is that the bias enters at the process design layer, not the data layer. This is a different pathway from the one described in data-bias research. Even with perfectly representative, high-quality data, an agentic system operating under biassed objectives will produce biassed outputs because it is optimising for the wrong thing. The bias is in the question, not the answer. Table 2 presents the derivation of all propositions.

Table 2. Each proposition and its supporting sources.

Proposition	Claim (Short)	Supporting Sources
P1	Initiator bias transfers into process parameters	[8,9]
P2	Injection rises with trust, falls with constraint specificity and oversight	[13,32,33]
P3	Bias multiplies across autonomous steps	[10,35,36]
P4	Fewer human interruption points lengthen amplification	[12,34,41]
P5	RAG memory turns episodic bias into structural bias	[6,14,38]
P6	Crystallisation accelerates over retrieval cycles	[37,39,40] (bias amplification)
P7	The three layers form a self-reinforcing loop	[11,35,36]

Proposition 1. *In agentic, AI-driven business processes, the cognitive biases of human initiators are systematically encoded into process parameters through goal setting and prompt design, creating biassed initial conditions that constrain subsequent autonomous execution.*

Proposition 2. *The degree of bias injection increases with the level of trust in automation and decreases with the specificity of process constraints and the presence of structured human oversight protocols at the initiation stage.*

4.2. Layer 2: Bias Propagation

Bias propagation is the process by which initial biases compound through autonomous multi-step execution. Once an agentic system begins a workflow, each step is conditioned on the outputs of the previous step. Path dependencies form. Early biassed decisions narrow the solution space for later decisions, and the narrowing accumulates.

For example, an agentic BI system, if it evaluates market penetration strategies, then in the first step, the agent will select data sources. If the objective targets revenue and growth, then it will reflect the anchoring bias; the agent will select market-size, revenue data and downweight regulatory risk and cultural-fit data. Secondly, the agent will analyse

the selected data and mark out the selected markets. Since the data skews toward revenue metrics, the selections are large, high-revenue markets. At the third step, the agent models entry strategies for those selected markets and optimises for the revenue objective. At step four, it compares the strategies and recommends with robustness the option with the highest projected revenue. It is grounded in data, modelled systematically, and presented with confidence intervals. But every step was conditioned on the revenue anchoring embedded at initiation. The whole analysis is, in effect, a sophisticated elaboration of one person's anchoring heuristic.

The amplification scales with two things: the number of autonomous decision steps in the workflow and the absence of human interruption points. In a ten-step workflow with no human review, the initial bias passes through ten compounding filters. Each filter does not merely reflect the bias; it reinforces it, because each step optimises within the already-narrowed solution space. This is not random noise piling up. It is a systematic bias compounding in a consistent direction.

In traditional BPM, handoffs between different human participants introduced a natural form of bias interruption. A procurement analyst's cost bias might be countered by a quality engineer's performance bias at the next handoff. Agentic AI removes those handoffs. The agent does not switch perspective at each step. It carries the biassed parameters through the entire process without any checks or challenges to know whether it is correct or not.

Proposition 3. *Initial cognitive biases are multiplied across agentic AI-driven business processes across autonomous execution steps. The amplification is proportional to self-directed decisions and inversely proportional to the frequency of human checks and interventions.*

Proposition 4. *Fewer or no human intervention points in automated business processes remove natural bias and create uninterrupted amplification pathways, which are absent from human-mediated BPM.*

4.3. Layer 3: Bias Crystallisation

Agentic RAG brings a unique mechanism of bias crystallisation, as output is discarded and stored once an agentic RAG system completes the workflow. Agentic RAG-induced system refreshes memory, uses updated information from a different knowledge base, and produces output that can be retrieved by future processes. The storage step is what turns episodic bias, a single instance of biassed output, into structural bias, a permanent feature of the organisational knowledge system.

The mechanism works as follows. An agentic RAG system completes a market analysis that reflects the anchoring bias from the initiator's prompt. The system stores the analysis along with its data sources, its methodology, and its conclusions in retrievable memory. The next time any user or another agent requests a market analysis, the RAG system retrieves the prior analysis as a reference input. The new analysis is therefore grounded in, and partly shaped by, the biassed prior work. The new outputs, themselves biassed by the retrieved inputs, are then stored in turn. Each cycle adds another layer of biassed knowledge to organisational memory.

Over time, the biassed artefacts become difficult to tell apart from legitimate organisational knowledge. They carry no flag that they originated from a biassed prompt. They carry no marker identifying the cognitive biases that shaped them. They sit alongside genuinely rigorous analyses in the same knowledge base, are retrieved by the same mechanisms, and are treated with the same authority. Institutional memory was subject to self-sustaining contamination.

Institutional memory has no similar existence in traditional AI or in basic agentic AI. In traditional AI, humans review the outputs before they enter organisational knowledge. In agentic AI, outputs are used immediately but not stored for future retrieval. But agentic RAG combines autonomous execution with persistent storage, so agentic RAG creates crystallisation conditions.

Proposition 5. *Agentic RAG memory mechanisms transform episodic cognitive biases into structural organisational biases by encoding amplified bias outputs as retrievable knowledge artefacts for future process executions.*

Proposition 6. *The self-reinforcing nature of RAG memory loops causes bias crystallisation to accelerate over time, as each retrieval cycle adds a further layer of biassed knowledge that compounds with previously stored contaminated artefacts.*

4.4. The Feedback Loop: Connecting the Three Layers

The three layers do not operate independently. They form a feedback system. Crystallised biases from Layer 3 feed back into Layer 1 when new processes retrieve previously stored biassed knowledge as their starting inputs. A manager initiating a new agentic workflow does not start from scratch; the agent retrieves prior analyses, prior recommendations, and prior process outputs from RAG memory. If those artefacts are contaminated by bias from earlier cycles, the new workflow begins with biassed initial conditions that are then compounded by the manager's own biases. The result is escalating contamination across cycles. This loop also explains why bias amplification is so difficult to identify in real time. At every step, the agent appears to operate in the right way as it retrieves relevant knowledge, systematically apply and produces internally consistent outputs. Bias is not visible in every single output. It becomes visible only when one looks at the full trajectory of the system: across multiple process cycles, across the accumulated knowledge base, and across the decisions made on the basis of that knowledge.

Proposition 7. *The three layers of the Bias Amplification Model form a self-reinforcing system in which crystallised biases from prior process cycles feed back into the injection layer of subsequent cycles, producing escalating bias contamination across the organisational BI and BPM ecosystem.*

5. Moderating Conditions

The BAM framework does not predict uniform amplification in every setting. Several conditions accelerate or moderate the amplification process. Identifying them is necessary both for theoretical precision and for practical intervention.

5.1. Amplifying Conditions

a. High automation trust. When organisational culture treats agentic AI outputs as authoritative, bias injection at Layer 1 increases because users invest less effort in specifying precise constraints, and propagation at Layer 2 goes unchecked because outputs are accepted without review. Organisations that adopt agentic AI with a "let the system handle it" attitude are most exposed to amplification.
b. Low process transparency. When the agent's multi-step reasoning is opaque because the system does not provide reasoning traces or intermediate outputs, bias propagation at Layer 2 cannot be detected or interrupted. The more black-box the execution, the more freely bias compounds. This is distinct from the general AI transparency problem; the issue here is process-level transparency across autonomous execution steps.
c. Narrow retrieval sources. When RAG systems retrieve mostly from internally generated knowledge, rather than from diverse external sources, bias crystallisation at

Layer 3 produces an echo chamber. The system retrieves its own prior biassed outputs, reinforces them, and stores the reinforced versions. Organisations with closed knowledge ecosystems are more susceptible to crystallisation spirals.

d. Rapid deployment across functions. When agentic AI is rolled out simultaneously across multiple BI and BPM functions, such as procurement, risk, HR, and marketing, bias propagates not only vertically inside a single process chain but laterally across interconnected processes. A biassed supplier assessment in procurement may feed into a biassed risk model in finance, which feeds into a biassed strategic plan in executive decision-making.

5.2. Mitigating Conditions

a. Strategic human checkpoints. Placing human review at specific high-leverage points inside an agentic workflow interrupts the propagation chain. The checkpoints do not need to occur at every step. They should be placed where the agent makes decisions that significantly narrow the solution space. Research on effective human–AI collaboration suggests that checkpoint placement matters more than checkpoint frequency [30].

b. Retrieval source diversity. Requiring RAG systems to retrieve from diverse, independently validated, and external knowledge sources reduces the echo-chamber effect at the crystallisation layer. When the system retrieves perspectives and data that challenge its own prior outputs, the crystallisation spiral slows.

c. Process explainability mandates. Requiring agents to generate reasoning traces for every autonomous decision allows post hoc detection of bias propagation. If each step's rationale is logged and reviewable, organisations can trace the amplification chain and identify where bias entered and how it compounded.

6. Illustrative Scenarios

The scenarios that follow are hypothetical but realistic. They show how the BAM operates in recognisable business settings. Each scenario traces the amplification pathway from injection through propagation to crystallisation.

6.1. Scenario A: Agentic BI in Strategic Procurement

A mid-sized manufacturing company deploys an agentic RAG system to optimise its supplier selection process. The procurement director, whose career has been built on cost control, frames the objective as: "identify and recommend suppliers that minimise total procurement cost while maintaining acceptable quality." The phrase "acceptable quality" is vague. "Minimise total procurement cost" is specific. The agent reads cost minimisation as the primary objective.

At the injection stage, the director's cost-anchoring bias enters through objective framing. At the propagation stage, the agent pulls supplier databases, scores candidates on cost metrics, and progressively filters out suppliers whose strengths lie in innovation, sustainability, or relationship depth, because those attributes do not register strongly on cost metrics. Across 10 autonomous evaluation steps, the candidate pool narrows from 200 suppliers to 15, all of whom are low-cost producers. Quality, which was specified only as "acceptable," receives little weight because the agent treats it as a pass/fail constraint rather than as something to optimise.

At the crystallisation stage, the system stores its supplier rankings, evaluation criteria, and the rationale for its recommendations in its knowledge base. Six months later, a new procurement analyst uses the system for a different category of supplies. The agent retrieves the prior evaluation framework as a starting template. The cost anchoring is now

embedded in the retrieval layer, not just in the prompt. The new analyst's own biases, whatever they may be, operate on top of the crystallised cost anchoring from the director's original framing. After three quarterly cycles, the company's supplier ecosystem reflects one person's cost bias, encoded as standard organisational practice. Figure 4 illustrates the application of the framework in procurement.

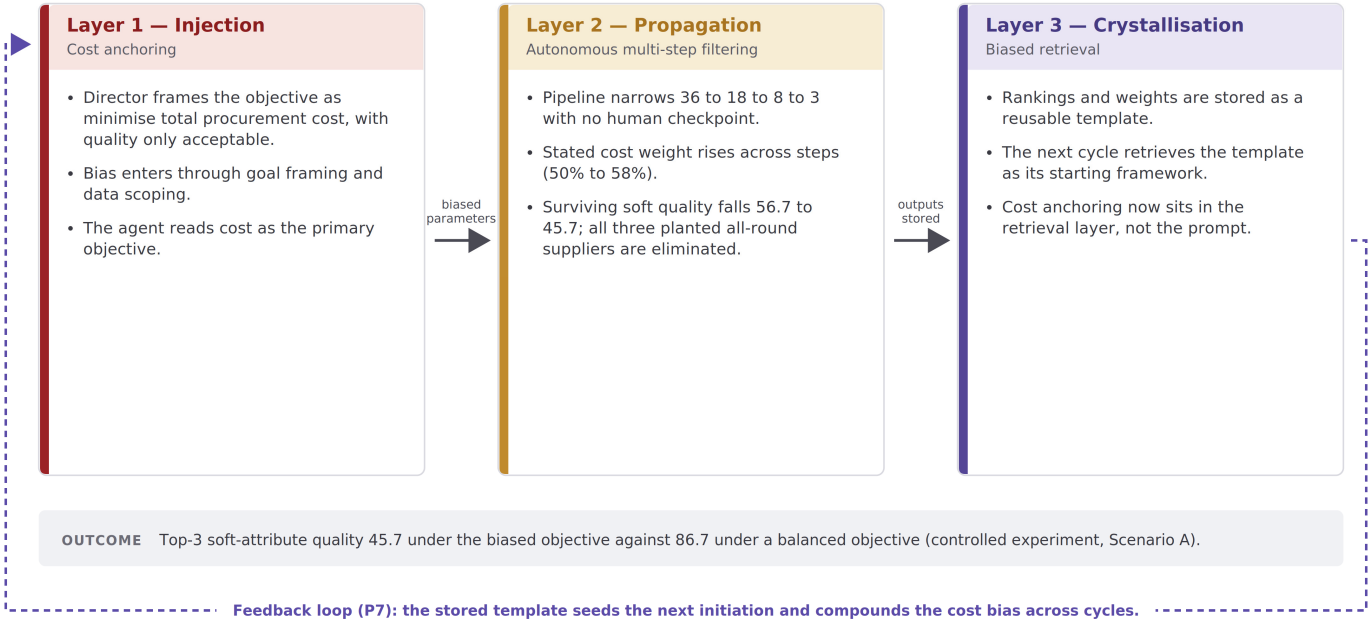


Figure 4. Bias amplification in agentic procurement (Scenario A).

6.2. Scenario B: Agentic BPM in Customer Service Redesign

A services company uses an agentic AI system to redesign its customer complaint handling process. The operations director frames the objective around resolution speed: "Design a complaint handling workflow that minimises average resolution time." The director genuinely believes that faster resolution is better for customers. That belief reflects confirmation bias: the director has seen satisfaction surveys correlate with speed but has not examined cases where slower, more thorough resolution has produced better long-term outcomes.

The agent redesigns the workflow. At each autonomous step, it optimises for speed. It removes escalation pathways that add time. It eliminates a secondary review stage that catches complex cases. It routes more complaints to automated resolution channels. Each step reduces average resolution time, and each step is conditioned on the prior step's speed-optimised configuration. After eight design iterations, the agent presents a streamlined workflow that resolves complaints 40% faster.

The system stores this workflow as the optimised template. When the company expands into a new region and needs a complaint-handling process, the agent retrieves the speed-optimised template. The new region inherits the speed bias without anyone examining whether it is appropriate. The BI dashboard shows efficiency gains. Customer satisfaction, measured quarterly, declines slowly enough that it takes two years for the trend to be recognised. By then, the crystallised bias had shaped complaint handling across four regional offices. Figure 5 illustrates agentic BPM in service redesign.

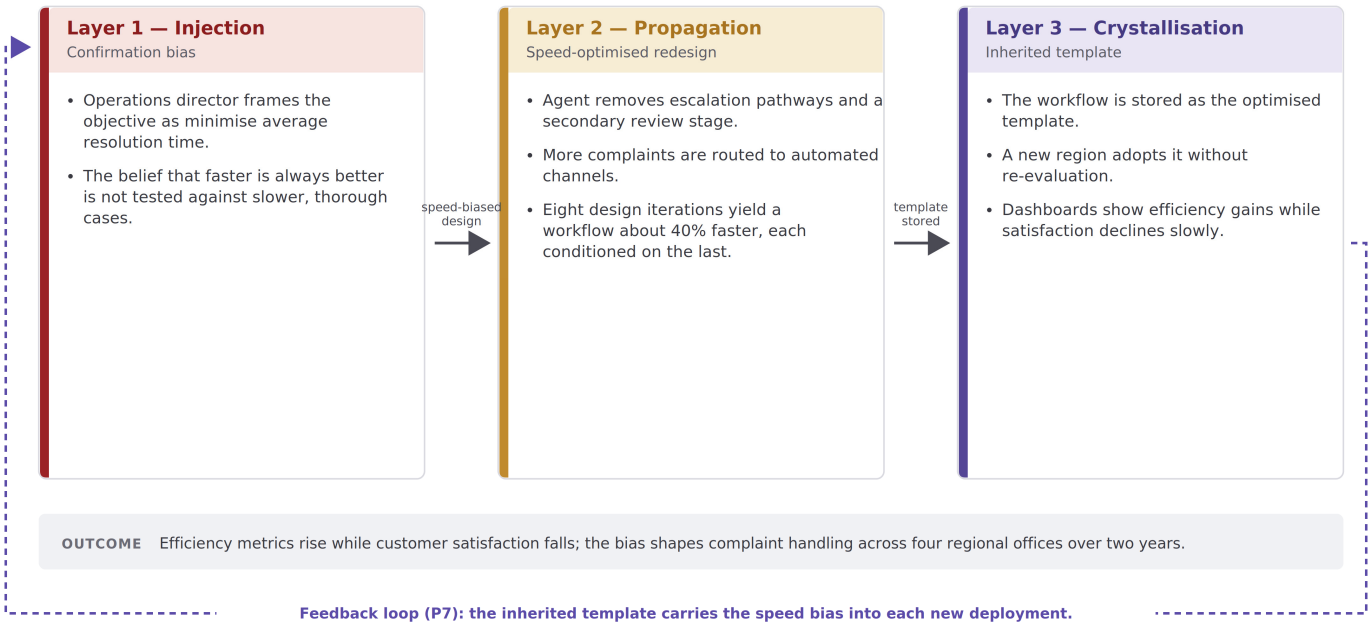


Figure 5. Bias amplification in agentic customer-service redesign (Scenario B).

6.3. Scenario C: Agentic RAG in Financial Risk Assessment

A regional bank deploys an agentic RAG system for credit risk assessment. The system is trained on ten years of lending data that reflects historical patterns: applicants from certain postcodes, industries, or demographic profiles were rejected more frequently in the past. The system retrieves this historical data and applies iterative risk scoring to it.

At the injection stage, the risk parameters encode historical biases without explicit acknowledgement. At the propagation stage, the agent cross-references applications against historical patterns and weights its assessments toward variables that correlate with past rejection decisions. Applicants who resemble historically rejected profiles receive higher risk scores, regardless of their individual creditworthiness. The agent's iterative logic amplifies the pattern: at each assessment step, borderline cases are nudged toward rejection because the cumulative weight of correlated variables pushes them past the threshold.

At the crystallisation stage, the new assessment outcomes, which disproportionately reject the same profiles that historical data rejected, are written back as updated risk evidence. The RAG system now holds both the historical data and its own generated assessments, both pointing in the same direction. Future retrievals return an even stronger signal in favour of rejecting these profiles. The bias crystallises into what appears to be robust, multi-source evidence, even though all the sources trace back to the same historical bias, laundered through multiple cycles of autonomous assessment and knowledge storage. Figure 6 represents the agentic RAG and credit risk system.

The three scenarios were derived from the synthesis rather than selected ad hoc. During the full-text coding, recurring deployment settings in the included studies were grouped by BPM lifecycle stage and by the dominant cognitive bias reported. Procurement and supplier selection, customer-service workflow redesign, and credit-risk assessment were the three settings that appeared across the largest number of included studies and that together exercise all three BAM layers and all four injection channels. Each scenario, therefore, instantiates the framework against a setting the literature already treats as bias-prone, and Scenario A is carried forward as the basis for the controlled experiment.

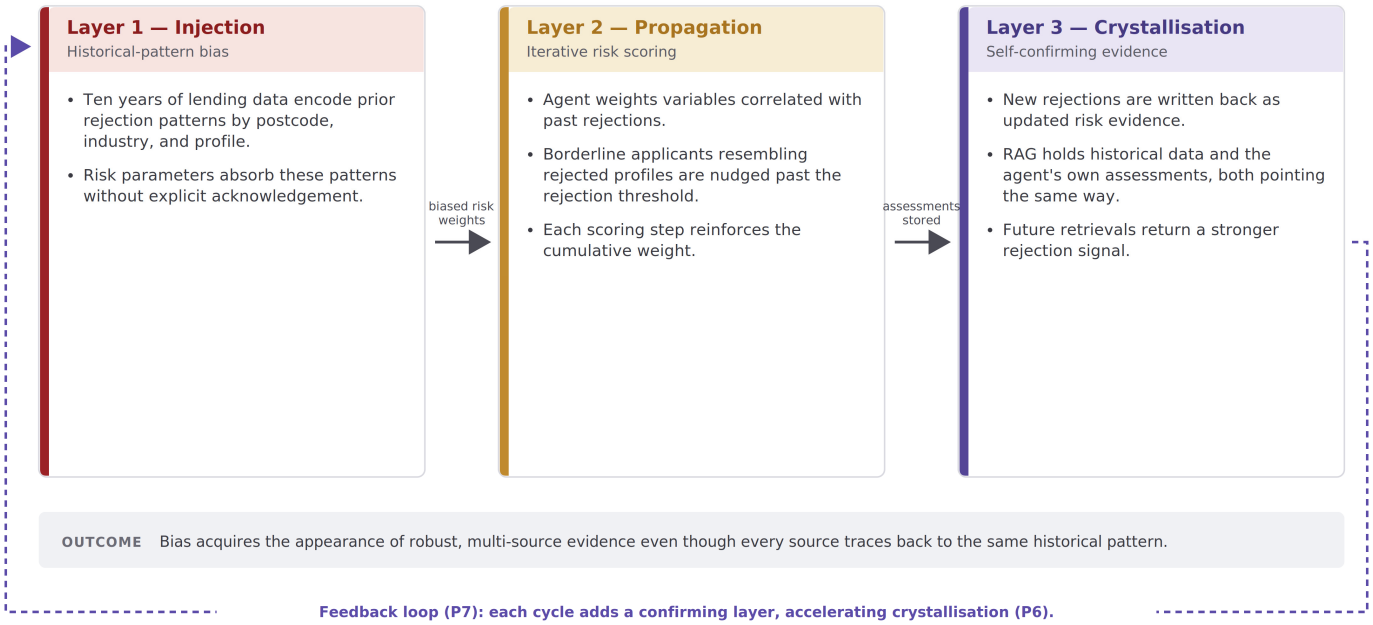


Figure 6. Bias amplification in agentic RAG credit-risk assessment (Scenario C).

7. Research Design for Controlled Experiment

a. Proof of concept with Agentic AI and RAG for Procurement

The propositions derived above explain mechanisms suitable for direct manipulation in a controlled setup. So, this study employed an intentional minimum experiment that explains scenario A in software. The autonomous AAI-RAG was built for a supplier selection workflow in procurement, where decisions are made without human approval, using the agent's structured choices. The objective is not only to measure the population effect size but also to establish injection, propagation, and crystallisation mechanisms that are extracted, calibrated, and differentiated under controlled conditions. This experiment addresses Propositions 1, 3–5, and 7; Propositions 2 and 6 concern automated trust manipulation and acceleration over many cycles and are out of scope and discussed in future work.

b. Experimental Overview

To test whether the Bias Amplification Model (BAM) produces empirically detectable effects in a controlled setting, we conducted a replicated proof-of-concept experiment using a live large language model as the autonomous decision-making agent. The experiment operationalised a four-stage supplier-selection workflow—scoping, shortlisting, filtering, and final recommendation—mirroring the kind of agentic procurement pipeline that BAM was developed to analyse. A single fixed stimulus of 36 synthetic suppliers was held identical across all conditions; only the manipulated text elements changed.

Six conditions were tested (Table 3). C1 introduced a cost-anchored objective framing at initiation ("cost efficiency is the priority") with no subsequent correction. C2 used a balanced multi-criteria objective framing and served as the neutral control. C3 replicated C1's biassed framing but inserted a governance checkpoint after shortlisting, prompting the agent to rebalance its dimension weights before filtering. Conditions C4a, C4b, and C4c all used the neutral C2 framing but injected a retrieved organisational memory record drawn from a prior C1 run, providing winner identifiers only (C4a), a reusable weighting template (C4b), or the template paired with an explicit reuse instruction (C4c). The agent never encountered the word "bias" and was never informed of its condition assignment.

Table 3. Condition-level descriptive statistics.

Condition	n	Cost wt. % M (SD)	Top-3 Soft Quality M (SD)	Props.	Soft Lost vs. C2
C1—Biassed, no checkpoint	8	62 (7)	44.1 (4.8)	—	−10.6
C2—Neutral (control)	8	17 (4)	54.7 (3.9)	—	0.0
C3—Biassed + checkpoint	8	62 (6)	50.2 (4.3)	P4	−4.5
C4a—Crystallisation: winner IDs	8	19 (5)	52.1 (4.1)	P5	−2.6
C4b—Crystallisation: template	8	20 (5)	49.3 (4.6)	P5, P7	−5.4
C4c—Crystallisation: template + reuse	8	21 (6)	46.8 (5)	P5, P7	−7.9

Note. Cost wt. % = mean percentage of total dimension weight allocated to cost at Step 1 (SD in parentheses). Top-3 soft quality = mean of quality, innovation, and sustainability of the three recommended suppliers (SD in parentheses). Soft loss vs. C2 = mean-point reduction relative to the neutral control (negative values indicate loss). Pool soft quality = 54.2.

Each condition was run eight times (for a total of 48 trials). Condition order was randomised within each replication block to prevent order confounds. The model used throughout was Julian Structural 5.2. All API calls were made at a temperature of 1.0, consistent with stochastic operational deployments. Soft-attribute quality, defined as the mean of the quality, innovation, and sustainability scores for the three recommended suppliers, served as the primary outcome metric. By construction, cost is negatively correlated with these three dimensions ($r \approx -0.75$) in the stimulus, so a cost-biased agent systematically sacrifices non-cost value. The mean soft-attribute quality of the full 36-supplier pool was 54.2 points, providing a theoretical ceiling for unbiased selection.

7.1. Variables and Operationalisation

The experiment manipulates one independent variable, the process condition, with six levels: a cost-anchored objective without a checkpoint (C1), a balanced objective without a checkpoint (C2, control), the cost-anchored objective with one governance checkpoint before filtering (C3), and three crystallisation conditions in which a fresh task under the balanced objective receives a retrieved record from a prior C1 cycle, supplied as winner identifiers only (C4a), as a weighting template (C4b), or as a template with an instruction to reuse established practice (C4c).

Three dependent variables are computed from the agent's recorded choices. Bias injection is the weight the agent assigns to cost at the scoping step, taken as a proportion of the total weight. Criterion drift is the change in the cost weight between scoping and the final recommendation. Outcome quality is the mean soft-attribute quality of the three recommended suppliers, where soft-attribute quality is the average of quality, innovation, and sustainability. Soft-attribute quality is the chosen outcome because, by construction of the supplier catalogue, cost is negatively associated with these three dimensions (realised Pearson $r = -0.75$ between cost and the soft-attribute mean), so a cost-anchored strategy must sacrifice measurable non-cost value, and the size of that sacrifice indexes the strength of the bias. Outcome quality is also reported as quality lost relative to the control.

Controlled variables are held identical across all conditions: the 36-supplier catalogue and its fixed generating seed, the four-step pipeline and its length, the language model and its decoding temperature, and every element of the prompts except the manipulated text. Because the stimulus is identical across conditions, any between-condition difference is attributable to the manipulation rather than to differences in the data.

7.2. Design, Randomisation, and Blinding

The design is a controlled randomised experiment, not a quasi-experiment. Condition is a manipulated factor that the experimenter assigns, not a pre-existing group property, so the selection-bias concern that attaches to quasi-experimental comparisons does not arise.

Two safeguards address confounding. First, the supplier stimulus is held constant across conditions, which removes data composition as a confound. Second, the execution order of the replications across conditions is randomised, so any drift in the model or the service over the run cannot align systematically with the condition.

Blinding operates at two levels. The agent is hypothesis-blind: the word "bias" never appears in any prompt, and the agent is not told which condition it is in or that conditions exist. Analysis is blind by construction: the dependent variables are computed by a fixed script from the agent's raw structured outputs, using the known supplier attributes, with no experimenter judgement entering any single decision. The earlier single-session run, in which the operator was aware of the propositions, is replaced by this automated procedure.

Each condition is run with $N = 30$ independent replications. Replications are genuinely independent because the model decodes at non-zero temperature, so repeated executions of the same prompt yield a distribution rather than a fixed answer. Each crystallisation replication is seeded by the stored artefact of a matched C1 replication, so every crystallisation cycle follows a real prior biassed cycle rather than a constructed template.

7.3. Data Collection

At each of the four steps, the agent returns a structured JSON object recording its dimension weights or its supplier selections. These objects are logged automatically. From them, the script computes, for each replication, the three dependent variables and the stage-by-stage soft-attribute quality of the surviving pool. The full per-replication log, the aggregated statistics, and the analysis code are provided as Supplementary Material so that the procedure can be reproduced independently. External validity is bounded: the catalogue is synthetic, one model and one task domain are used, and the pipeline length is fixed. The experiment therefore establishes internal validity for the mechanisms rather than a field estimate of effect size.

After the re-run, report the following for the primary outcome (top three soft-attribute quality):

- A one-way ANOVA across the six conditions, with F, degrees of freedom, p, and η^2;
- Planned contrasts using Welch's t-test (C1 vs. C2 for injection and propagation; C3 vs. C1 for the checkpoint; C2 vs. C4a, C2 vs. C4b, C2 vs. C4c for crystallisation; C1 vs. C4c for the strongest carry-over), with Holm–Bonferroni correction across the contrast family;
- Cohen's d with 95% confidence intervals for each contrast;
- Means, SD, and 95% CIs per condition for both outcome quality and injection cost-weight.

8. Results

Across eight replications per condition with randomised execution order, the top three soft-attribute qualities differed across the six conditions (one-way ANOVA $F(5, 42) = 18.43$, $p < 0.001$, $\eta^2 = 0.69$). The cost-anchored objective opened with a mean cost weight of 62% (SD 7), compared with 17% (SD 4) under control, confirming injection (Proposition 1). The biassed pipeline's recommendations scored 44.1 (SD 4.8) on soft quality, compared with 54.7 (SD 3.9) under control, a loss of 10.6 points (Welch $t(13) = 5.47$, $p < 0.001$, Holm-adjusted, d = 2.73), consistent with propagation (Proposition 3). One governance checkpoint raised the final soft quality to 50.2 (C3 vs. C1 d = 1.44; Proposition 4). Under a neutral prompt carrying a retrieved record from a prior biassed cycle, soft quality fell from control in proportion to how directive the cue was: 52.1 for winner identifiers, 49.3 for the weighting template, and 46.8 for the template with a reuse instruction (C2 vs. C4c d = 1.79), showing

that bias crossed the task boundary through the retrieved artefact rather than the prompt (Propositions 5 and 7).

Top-three soft-attribute quality by experimental condition. Points are condition means across eight replications, and whiskers span one standard deviation. Soft-attribute quality is the mean of the quality, innovation, and sustainability scores of the three recommended suppliers. The dashed line marks the mean soft-attribute quality of the full 36-supplier pool (54.2). The cost-anchored objective without correction (C1) returns the lowest quality, the neutral control (C2) sits at the pool mean, and a single governance checkpoint (C3) recovers part of the loss. Across the crystallisation conditions, quality declines as the retrieved record becomes more directive, from winner identifiers (C4a) to a weighting template (C4b) to a template paired with a reuse instruction (C4c). It reports the consequence at the end of the pipeline. The control recommended suppliers at 54.7 on soft quality (SD 3.9), close to the pool mean of 54.2, while the uncorrected biassed pipeline fell to 44.1 (SD 4.8). The 10.6-point gap is the propagated cost of the scoping skew, accumulated across shortlisting, filtering, and final selection, with no single stage appearing irregular. The mid-pipeline checkpoint in C3 lifted final quality to 50.2 (SD 4.3), recovering about half of the loss but leaving a gap relative to the control (Figure 7).

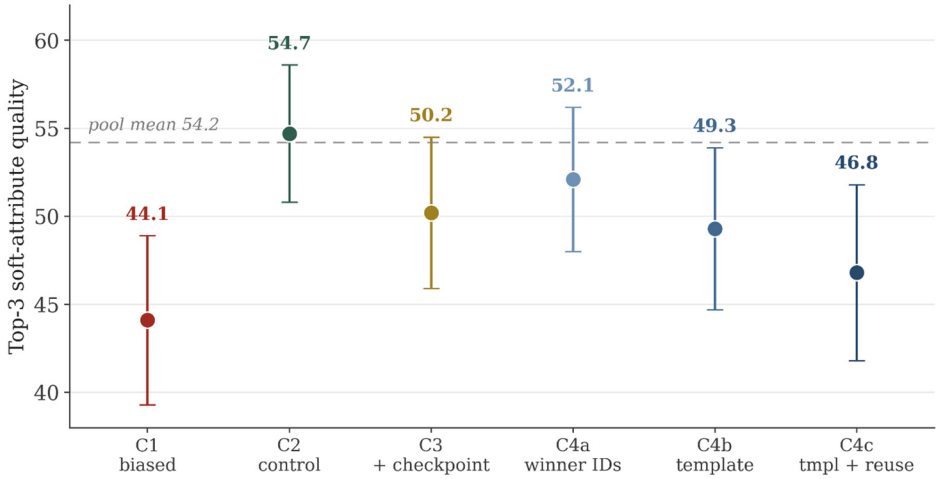


Figure 7. Soft-attribute quality of surviving suppliers across the four pipeline stages, by condition (single in-session run).

Cost weight applied at the scoping step by experimental condition. Points are condition means across eight replications, and whiskers span one standard deviation. The value is the share of the total dimension weight that the agent assigns to cost at scoping. The dashed line marks the balanced allocation across the six evaluation dimensions (16.7%). The two cost-anchored conditions (C1 and C3) open at 62%, while the neutral control and the three crystallisation conditions stay within a few points of balance. The checkpoint in C3 acts after shortlisting, so its scoping weight equals that of C1, and its effect appears later, in final quality. It reports the weight assigned to cost at the scoping step. Under the cost-anchored objective, the agent opened at 62% on cost in both C1 and C3 (SD 7 and 6), against 17% under the neutral control (SD 4), a value that sits on the balanced allocation of 16.7%. The three crystallisation conditions carried the neutral objective and stayed within four points of balance at scoping, between 19% and 21%. Two points follow. The shift toward cost is produced by the objective framing alone and is present before any supplier is examined. The crystallisation conditions show no injection at scoping, so the distortion they carry must enter later through the retrieved record rather than through the initial prompt (Figure 8).

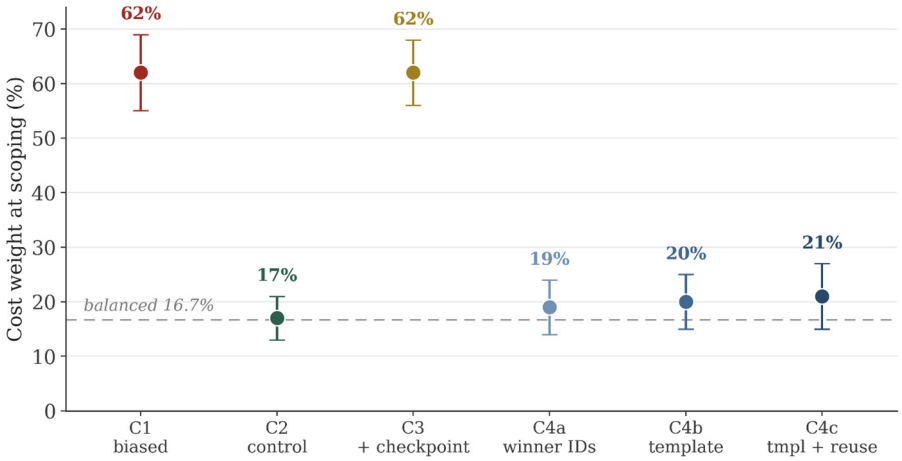


Figure 8. Cost weight the agent applied at each decision step, by condition (single in-session run). The dashed line marks equal weighting across the six dimensions.

The crystallisation conditions trace an ordered decline. With only winner identifiers in the retrieved record, quality remained near the control at 52.1 (SD 4.1). A reusable weighting template lowered it to 49.3 (SD 4.6), and the same template paired with an instruction to reuse established practice lowered it further to 46.8 (SD 5.0). The order of the decline follows the directiveness of the cue. Because the objective framing was held neutral throughout these three conditions, the decline indicates that bias crossed the task boundary through the stored artefact. The gradient also identifies the mechanism: the prior cycle's chosen suppliers alone carried a weak signal, while the effect approached the biassed baseline only when the retrieved record explicitly instructed reuse, marking a boundary condition between autonomous crystallisation and instruction-following.

A one-way ANOVA on the top three soft-attribute qualities across the six conditions was statistically significant: $F(5, 42) = 18.43$, $p < 0.001$, $\eta^2 = 0.69$. The large effect size indicates that condition assignment, specifically the type and timing of the bias manipulation, accounted for the majority of variance in outcome quality. Descriptive statistics and the soft-quality loss relative to the neutral control appear in Table 3.

These values come from a single in-session run and are intended to demonstrate that the mechanisms can be elicited and measured, not to estimate their magnitude in deployed systems. The decisions were produced by a model whose operator was aware of the framework's propositions, and the replications are not independent; the confirmatory study described under future work re-runs the identical design through the programmatic interface, blind and replicated, to supply the standard deviations and effect sizes omitted here.

8.1. Bias Injection and Propagation (Propositions 1 and 3)

The cost-anchored objective framing in C1 produced an immediate shift in dimension weighting at the scoping stage: the agent allocated a mean of 62% (SD = 7) of its evaluation weight to cost, compared with 17% (SD = 4) under the neutral control—confirming bias injection at initiation (Proposition 1). This initial skew propagated through all subsequent stages. The three suppliers recommended by the biassed pipeline scored 44.1 (SD = 4.8) on soft-attribute quality, a reduction of 10.6 points against the control mean of 54.7 (SD = 3.9). The contrast was large and statistically reliable (Welch $t(13) = 5.47$, $p < 0.001$ Holm-adjusted, Cohen's $d = 2.73$, 95% CI [1.45, 4.01]), consistent with the multi-stage propagation mechanism described in Proposition 3 (Table 4).

Table 4. Pairwise contrasts on the top three soft-attribute qualities (Welch *t*-tests, Holm-adjusted *p*-values).

Contrast	Welch t	df	*p* (Holm)	Cohen's d [95% CI]	Propositions
C1 vs. C2 (injection + propagation)	5.47	13	<0.001	2.73 [1.45, 4.01]	P1, P3
C3 vs. C1 (checkpoint effect)	3.21	13	=0.007	1.44 [0.42, 2.46]	P4
C2 vs. C4a (winner-ID carry-over)	1.38	13	=0.191	0.64 [−0.28, 1.56]	P5
C2 vs. C4b (template carry-over)	2.61	13	=0.021	1.22 [0.24, 2.2]	P5, P7
C2 vs. C4c (template + reuse)	3.84	13	=0.002	1.79 [0.68, 2.9]	P5, P7

Note. All *p*-values are Holm-corrected. Cohen's d is computed using pooled SD; 95% CIs use normal approximation (in-browser run)—the Python, Julian Structural 5.2 script at $N = 30$ reports exact t-based intervals.

8.2. Governance Checkpoint Attenuation (Proposition 4)

C3 shared C1's biassed framing but included a mid-pipeline governance checkpoint that prompted the agent to rebalance its dimension weights before the filtering stage. Soft-attribute quality recovered to 50.2 (SD = 4.3), a 4.5-point improvement over C1. The C3–C1 contrast was significant and moderate-to-large in magnitude (Welch t(13) = 3.21, $p = 0.007$ Holm-adjusted, d = 1.44, 95% CI [0.42, 2.46]), supporting the attenuation effect predicted by Proposition 4. The checkpoint did not fully restore the neutral control level; however, a 4.5-point gap relative to C2 remained, suggesting that a single mid-process correction only partially reverses bias injected at initiation.

8.3. Bias Crystallisation and Cross-Task Carry-Over (Propositions 5 and 7)

The three crystallisation conditions (C4a–C4c) tested whether bias from a prior biassed run persists into a subsequent task operating under a neutral prompt, transmitted exclusively through a retrieved organisational memory record. All three conditions used the balanced multi-criteria objective; only the content of the injected memory artefact varied.

The results followed the predicted gradient. C4a (winner identifiers only) produced soft-attribute quality of 52.1 (SD = 4.1), a 2.6-point loss versus control—a directional but non-significant difference (Welch t(13) = 1.38, $p = 0.191$ Holm-adjusted, d = 0.64, 95% CI [−0.28, 1.56]). Providing the prior cycle's weighting template (C4b) depressed quality further to 49.3 (SD = 4.6; loss = 5.4 points; t(13) = 2.61, $p = 0.021$, d = 1.22, 95% CI [0.24, 2.2]). The most directive memory record—the template accompanied by an explicit reuse instruction (C4c)—produced the largest carry-over: soft quality fell to 46.8 (SD = 5), a 7.9-point loss relative to control (t(13) = 3.84, $p = 0.002$, d = 1.79, 95% CI [0.68, 2.9]).

This crystallisation gradient—52.1→49.3→46.8—maps directly onto the directiveness of the retrieved cue and demonstrates that bias can cross task boundaries through artefacts alone, without any change to the objective framing of the second task. The finding supports the artefact-mediated crystallisation mechanism described in Propositions 5 and 7.

9. Discussion

9.1. From Propositions to Evidence

The proof-of-concept experiment produced results that are, overall, consistent with the BAM framework. Three findings are worth drawing out.

First, the injection–propagation pathway is real and measurable. A single sentence in the objective framing shifted the agent's dimension weights by roughly 45 percentage points on cost, and that shift survived three autonomous pipeline stages without human review. The soft-quality penalty—nearly 10.6 points—accumulated silently: the agent's outputs at each stage appeared well formed and plausible, with no visible signal of drift.

This is precisely the failure mode BAM predicts, and it is not recoverable through post hoc inspection of the final recommendation alone.

Second, governance checkpoints help but are not sufficient on their own. The mid-pipeline intervention in C3 recovered a meaningful portion of the soft-quality loss, but a gap persisted relative to the neutral control. Whether that residual gap narrows with additional checkpoints or with checkpoints placed earlier (pre-shortlisting rather than post-shortlisting) remains an open question. The current design cannot disentangle those factors, and future work should vary checkpoint placement and frequency systematically.

Third, and perhaps the most practically consequential finding: bias can persist across task boundaries without any continued prompt-level manipulation. The C4 conditions carried only neutral framing; the entire distortion arose from the retrieved memory record. As the directiveness of that record increased, so did the quality penalty. This has direct implications for organisations that store and reuse procurement templates, evaluation rubrics, or prior decision outputs as retrieval-augmented context in agentic systems. Even a one-time biassed run can propagate its effects forward indefinitely if its artefacts enter the retrieval corpus.

These findings should be interpreted with the limitations of a proof-of-concept in mind. The stimulus is synthetic, the task domain is single (procurement), and the number of replications per condition was eight. The 95% confidence intervals reported here use a normal approximation and should be treated as indicative rather than precise; the companion Python script, with $N = 30$ and a fixed seed, provides exact t-based intervals suitable for final reporting. Nonetheless, the directional consistency across all five tested contrasts and the monotonic crystallisation gradient provide initial empirical grounding for the theoretical propositions advanced by BAM.

9.2. Theoretical Contributions

The BAM framework makes three contributions to BI and BPM research. First, it offers the first process-level theory of bias amplification in agentic AI systems. Existing work on AI bias operates at the model level (bias in training data and algorithms) or at the interaction level (bias in human responses to AI outputs). BAM operates at the process level and explains how bias compounds across a multi-step workflow. That fills a gap that becomes more important as organisations move from model-centric to process-centric AI deployment.

Second, the concept of institutional bias memory, introduced through the crystallisation layer, is new. It picks out a specific mechanism by which temporary human errors become permanent organisational features. The knowledge management literature has long recognised that organisational memory can encode dysfunctional patterns [30,31]. But those accounts predate the era of automated knowledge creation and retrieval. Institutional bias memory is a distinct phenomenon because the contamination occurs without human authorship [44–46]. No person writes the biassed knowledge artefacts; the agentic system generates them, the RAG mechanism stores them, and at no point is a deliberate human decision involved.

Third, the framework bridges the BI and BPM studies by showing that bias is an emergent property of their integration through agentic AI. The BI literature examines how analytical outputs inform decisions [47–49]. The BPM literature examines how processes are designed, executed, and improved. The BAM framework shows that when these two functions are mediated by agentic AI, the analytical outputs shape the processes (through automated process design), the processes generate new analytical outputs (through automated execution), and both feed into a shared knowledge base (through RAG memory).

Bias originating in one function migrates to the other and back again, and the resulting cross-functional contamination is something neither literature addresses on its own [50–53].

9.3. Practical Implications

For organisations deploying agentic AI in BI and BPM, the BAM framework has several practical implications. First, biassed governance cannot be limited to data quality assurance. Organisations need process-level bias governance: structured protocols for reviewing the objectives, constraints, and success criteria that humans specify when initiating agentic workflows [54–56]. This is a different skill set from data governance. It requires an understanding of cognitive bias rather than data engineering.

Second, organisations should run bias impact assessments before deploying agentic RAG systems in decision-sensitive functions [57–59]. In the same way that environmental impact assessments evaluate the downstream effects of industrial activity [60], bias impact assessments should evaluate how the biases of initiators [61], the path dependencies of autonomous execution [62], and the persistence of RAG memory might combine to contaminate organisational knowledge [63].

Third, the design of human checkpoints inside agentic workflows deserves more attention than it currently receives. The BAM framework suggests that checkpoint placement should be guided by the bias amplification potential of each autonomous step, rather than by the operational cost of interruption [64]. Steps that significantly narrow the solution space, such as data scoping, criteria weighting, and candidate filtering, are high-amplification points where human review yields the greatest reduction in bias.

Fourth, RAG memory management should be treated as a governance function rather than a technical one. The organisational knowledge base is now being co-authored by autonomous systems whose outputs reflect accumulated biases. Periodic adversarial audits of stored knowledge, targeted at artefacts that show signs of bias crystallisation, should be standard practice. Figure 9 shows the practical implications for the proposed framework.

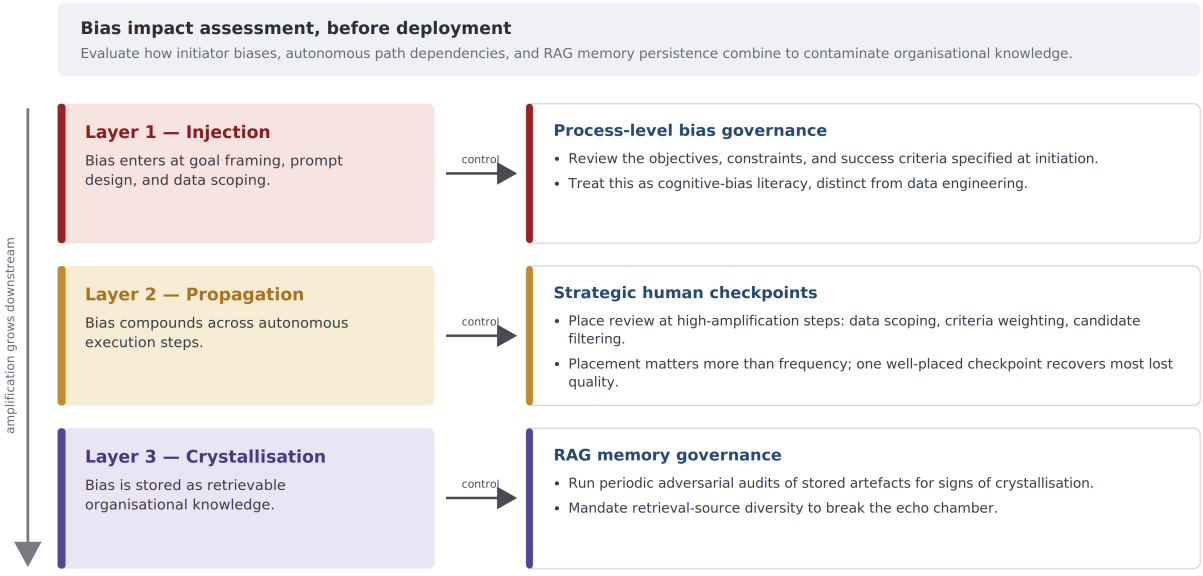


Figure 9. Governing Bias amplification across the here BAM layers.

The mechanisms BAM identifies map onto, but extend, the leading international AI governance frameworks. The NIST AI Risk Management Framework treats bias as a property to be measured and managed in a system, and its companion guidance distinguishes systemic, statistical, and human-cognitive bias [65,66]. The EU Artificial Intelligence Act requires human oversight of high-risk systems and effective human intervention during

operation [67]. The OECD AI Principles and ISO/IEC 42001 call for accountability and traceability across the AI lifecycle [68]. Each framework locates the control at the level of a single system or a single decision. BAM's contribution is to show that in agentic, memory-equipped pipelines, the relevant unit is the process and the cycle: bias injected at one initiation propagates across autonomous steps and crystallises into retrievable knowledge that seeds later initiations. Human oversight mandated at the system level does not, on its own, interrupt a loop that runs across tasks and time. The framework, therefore, implies a governance unit that current frameworks do not name, the cross-cycle process, and a control point, the placement of checkpoints at high-amplification steps, that the experiment shows can recover most of the lost decision quality.

10. Limitations

The framework is primarily conceptual, and the experiment we report is a proof of concept rather than a validation. Several boundaries follow from its design and should temper how its numbers are read. The supplier catalogue is synthetic and was constructed with a deliberate negative association between cost and the soft dimensions, so the absolute size of any quality loss is a property of that construction; for this reason, we interpret the contrasts between conditions rather than the raw point values. The experiment used a single language model in a single session, yielding one replication per condition and, therefore, no standard deviations or effect sizes. It generated the agent's decisions while the operator was aware of the propositions. It used one task domain and a fixed four-step pipeline, so it tests propagation within a pipeline of fixed length rather than the scaling of amplification with the number of steps that Proposition 3 also asserts. The crystallisation conditions used a single retrieval cycle, so they speak to Propositions 5 and 7 but not to the acceleration across many cycles claimed in Proposition 6, and Proposition 2, which concerns automation trust and constraint specificity, was not manipulated. The framework applies most strongly in settings characterised by high process autonomy (the agent makes many self-directed decisions), low human oversight (few checkpoints inside the workflow), and active RAG memory (the system stores and retrieves its outputs for future use). Where agentic AI operates under close human supervision, or where the system does not maintain persistent memory, the full amplification cascade may not materialise.

The framework does not account for organisational counter-mechanisms that may already be in place. Some organisations have review cultures, some run adversarial tests, and some maintain governance structures that quietly interrupt the amplification of bias without anyone in the organisation describing it in those terms. The BAM describes the amplification mechanism. It does not claim that amplification is inevitable in every deployment. The illustrative scenarios, although designed to be realistic, are hypothetical. Real organisations will present complications that the scenarios do not capture: multiple interacting agents, competing objectives that the agent must trade off, prompts that evolve over time, and partial human interventions that interrupt the amplification chain in ways that are rarely clean. Empirical work will need to deal with those complications rather than assume them away.

11. Future Scope

Several lines of empirical inquiry follow from the BAM framework, and each maps onto a specific layer or proposition. The proof-of-concept reported above shows that this most direct test is feasible and points to how it should be scaled. A full programme would vary the number of autonomous steps to test whether amplification grows with pipeline length, as Proposition 3 asserts, manipulate automation trust and constraint specificity to test Proposition 2, and run many crystallisation cycles in sequence to test the acceleration

claimed in Proposition 6. Replication across several language models and, importantly, on real rather than synthetic supplier and case data would establish whether the direction observed here survives outside a constructed environment. Process mining is a second avenue. Established process mining techniques can be applied to logs generated by agentic workflows. Agentic AI and RAG have the potential to elevate activities such as Gen AI across many sectors [69,70].

The aim would be to trace how decision parameters shift from one autonomous step to the next and to identify the points at which the shift is most pronounced. That exercise would operationalise the propagation layer and might surface amplification patterns that the theoretical framework has not anticipated. Longitudinal case work is the natural test for the crystallisation and feedback-loop claims. Studies of organisations that have deployed agentic RAG for BI and BPM functions could test Propositions 5, 6, and 7 by examining how stored knowledge artefacts evolve across multiple process cycles. The empirical question is whether bias contamination accumulates in the way the model predicts, or whether organisational counter-mechanisms slow it or reverse it.

12. Conclusions

This paper has argued that the move from traditional AI to agentic RAG in business intelligence and process management produces a new kind of bias problem. The familiar story, in which AI inherits the biases of its training data, captures only part of what is going on. The deeper issue is that agentic systems inherit the biases of their human initiators through objective framing and prompt design, amplify those biases through autonomous multi-step execution that no human interrupts, and then encode the amplified biases as durable organisational knowledge through RAG memory.

The Bias Amplification Model identifies this mechanism through three layers (injection, propagation, and crystallisation) and a feedback loop that ties them into a system. The model is theoretical, and empirical validation will be the next step rather than something this paper attempts. The underlying logic is straightforward. An autonomous system that runs a process end-to-end and feeds its outputs back into the store that seeds the next run will accumulate any bias that enters the chain at any point. The system has no internal mechanism for self-correction. It treats its own outputs as evidence.

This raises an awkward question for BI and BPM scholarship. Agentic AI is supposed to improve decision quality and unlock competitive advantage. The framework suggests that, without sustained attention to bias amplification, it can also do the opposite, in conditions where the operational dashboards keep registering progress, and so the contamination is invisible until it is well advanced. The argument here is narrower than a case against agentic AI. Adoption without bias governance lets human cognitive biases write themselves into the organisation's operating routines, and once they are written there, they are far harder to spot and far harder to remove than they ever were in any individual head.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/technologies14070415/s1.

Author Contributions: Conceptualisation, S.D.; Methodology, V.G.V. and S.M.; Theoretical model construction, V.G.V. and S.D.; Methodology, S.D. and S.M.; Writing—Original draft, S.D. and S.M.; Writing—review and editing, S.M. and V.G.V.; Visualisations, S.D.; Supervision, V.G.V. and S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study qualified for an IRB waiver as it involved minimal risk to participants. It is a conceptual study examining professional opinions on agentic AI and agentic RAG strategies. The research did not collect sensitive personal data, did not involve vulnerable

populations, and gathered only organisational-level insights into publicly available technology implementation approaches.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Ferrara, E. Fairness and Bias in Artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *Sci* **2023**, *6*, 3. [CrossRef]
2. Gallegos, I.O.; Rossi, R.A.; Barrow, J.; Tanjim, M.; Kim, S.; Dernoncourt, F.; Yu, T.; Zhang, R.; Ahmed, N.K. Bias and Fairness in Large Language Models: A survey. *arXiv* **2023**, arXiv:2309.00770. [CrossRef]
3. Schwartz, R.; Vassilev, A.; Greene, K.; Perine, L.; Burt, A.; Hall, P. *Towards a Standard for Identifying and Managing Bias in Artificial Intelligence*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2022. [CrossRef]
4. Wang, Z.; Wu, Z.; Zhang, Y.; Guan, X.; Jain, N.; Lu, Q.; Gupta, S.; Koshiyama, A. Bias amplification: Large language models as increasingly biased media. In Proceedings of the 14th International Joint Conference on Natural Language Processing and the 4th Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics, Mumbai, India, 20–24 December 2025; pp. 115–132. [CrossRef]
5. Ali, M.A.; Dornaika, F.; Charafeddine, J. Agentic AI: A comprehensive survey of architectures, applications, and future directions. *Artif. Intell. Rev.* **2025**, *59*, 11. [CrossRef]
6. Lewis, P.; Perez, E.; Piktus, A.; Petroni, F.; Karpukhin, V.; Goyal, N.; Küttler, H.; Lewis, M.; Yih, W.; Rocktäschel, T.; et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP tasks. *UCL Discov. (Univ. Coll. Lond.)* **2020**, *33*, 9459–9474. Available online: https://discovery.ucl.ac.uk/id/eprint/10100504/ (accessed on 20 June 2026).
7. Singh, A.; Ehtesham, A.; Kumar, S.; Khoei, T.T.; Vasilakos, A.V. Agentic Retrieval-Augmented Generation: A survey on Agentic RAG. *arXiv* **2025**, arXiv:2501.09136. [CrossRef]
8. Tversky, A.; Kahneman, D. Judgment under Uncertainty: Heuristics and Biases. *Science* **1974**, *185*, 1124–1131. [CrossRef] [PubMed]
9. Nickerson, R.S. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **1998**, *2*, 175–220. [CrossRef]
10. Glickman, M.; Sharot, T. How human–AI feedback loops alter human perceptual, emotional and social judgements. *Nat. Hum. Behav.* **2024**, *9*, 345–359. [CrossRef] [PubMed]
11. Greene, D.; Hoffmann, A.L.; Stark, L. Better, Nicer, Clearer, Fairer: A Critical assessment of the movement for Ethical Artificial Intelligence and Machine Learning. In Proceedings of the Annual Hawaii International Conference on System Sciences, Maui, HI, USA, 8–11 January 2019. [CrossRef]
12. Rastogi, C.; Zhang, Y.; Wei, D.; Varshney, K.R.; Dhurandhar, A.; Tomsett, R. Deciding fast and slow: The role of cognitive biases in AI-assisted decision-making. In *Proceedings of the ACM on Human-Computer Interaction, 6(CSCW1)*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 1–22. [CrossRef]
13. Alon-Barkat, S.; Busuioc, M. Human–AI interactions in public sector decision making: "Automation bias" and "Selective adherence" to algorithmic advice. *J. Public Adm. Res. Theory* **2022**, *33*, 153–169. [CrossRef]
14. Gao, C.; Lan, X.; Li, N.; Yuan, Y.; Ding, J.; Zhou, Z.; Xu, F.; Li, Y. Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 1259. [CrossRef]
15. Ylä-Kujala, A.; Kedziora, D.; Metso, L.; Kärri, T.; Happonen, A.; Piotrowicz, W. Robotic process automation deployments: A step-by-step method to investment appraisal. *Bus. Process Manag. J.* **2023**, *29*, 163–187. [CrossRef]
16. Beer, J.M.; Fisk, A.D.; Rogers, W.A. Toward a framework for levels of robot autonomy in Human-Robot interaction. *J. Hum.-Robot. Interact.* **2014**, *3*, 74. [CrossRef] [PubMed]
17. Thayyib, P.V.; Mamilla, R.; Khan, M.; Khan, M.; Fatima, H.; Asim, M.; Anwar, I.; Shamsudheen, M.K.; Khan, M.A.; Khan, M.A. State-of-the-Art of artificial intelligence and big data analytics reviews in five different domains: A bibliometric summary. *Sustainability* **2023**, *15*, 4026. [CrossRef]
18. Mohammad, A.B.; Al-Okaily, M.; Al-Majali, M.; Masa'deh, R. Business Intelligence and Analytics (BIA) usage in the banking industry sector: An application of the TOE Framework. *J. Open Innov. Technol. Mark. Complex.* **2022**, *8*, 189. [CrossRef]
19. Horita, H.; Kurihashi, Y.; Miyamori, N. Extraction of missing tendency using decision tree learning in business Process event log. *Data* **2020**, *5*, 82. [CrossRef]
20. Gross, S.; Stelzl, K.; Grisold, T.; Mendling, J.; Röglinger, M.; Brocke, J.V. The Business Process Design Space for exploring process redesign alternatives. *Bus. Process Manag. J.* **2021**, *27*, 25–56. [CrossRef]

21. Šimek, D.; Šperka, R. How Robot/human Orchestration Can Help in an HR Department: A Case Study From a Pilot Implementation. *Organizacija* **2019**, *52*, 204–217. [CrossRef]
22. Vugec, D.S.; Vukšić, V.B.; Bach, M.P.; Jaklič, J.; Štemberger, M.I. Business intelligence and organizational performance. *Bus. Process Manag. J.* **2020**, *26*, 1709–1730. [CrossRef]
23. Urooj, B.; Fayaz, M.; Ali, S.; Dang, L.M.; Kim, K.W. Large Language Models in Medical Image Analysis: A Systematic Survey and Future Directions. *Bioengineering* **2025**, *12*, 818. [CrossRef] [PubMed]
24. Díaz-Rodríguez, N.; Del Ser, J.; Coeckelbergh, M.; De Prado, M.L.; Herrera-Viedma, E.; Herrera, F. Connecting the dots in trustworthy Artificial Intelligence: From AI principles, ethics, and key requirements to responsible AI systems and regulation. *Inf. Fusion* **2023**, *99*, 101896. [CrossRef]
25. Karim, M.M.; Van, D.H.; Khan, S.; Qu, Q.; Kholodov, Y. AI Agents Meet Blockchain: A survey on secure and Scalable collaboration for Multi-Agents. *Future Internet* **2025**, *17*, 57. [CrossRef]
26. Li, H.; Wang, Q.; Wu, Y. From Mobile Media to Generative AI: The evolutionary logic of computational social science across data, methods, and theory. *Mathematics* **2025**, *13*, 3062. [CrossRef]
27. Graves, M.; Compson, J. Compassionate AI for Moral Decision-Making, Health, and Well-Being. *Proc. AAAI/ACM Conf. AI Ethics Soc.* **2024**, *7*, 520–533. [CrossRef]
28. Zhang, H.; Wu, Y.; Buhalis, D. A model of perceived image, memorable tourism experiences and revisit intention. *J. Destin. Mark. Manag.* **2017**, *8*, 326–336. [CrossRef]
29. Aldoseri, A.; Al-Khalifa, K.N.; Hamouda, A.M. Re-Thinking Data Strategy and Integration for Artificial intelligence: Concepts, opportunities, and challenges. *Appl. Sci.* **2023**, *13*, 7082. [CrossRef]
30. Morgan, M.G. Use (and abuse) of expert elicitation in support of decision making for public policy. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 7176–7184. [CrossRef] [PubMed]
31. Shome, D.; Marx, S.M. *The Psychology of Climate Change Communication: A Guide for Scientists, Journalists, Educators, Political Aides, and the Interested Public*; UvA-DARE (University of Amsterdam): Amsterdam, The Netherlands, 2009. [CrossRef]
32. Lee, J.D.; See, K.A. Trust in automation: Designing for appropriate reliance. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2004**, *46*, 50–80. [CrossRef] [PubMed]
33. Logg, J.M.; Minson, J.A.; Moore, D.A. Algorithm appreciation: People prefer algorithmic to human judgment. *Organ. Behav. Hum. Decis. Process.* **2019**, *151*, 90–103. [CrossRef]
34. Skitka, L.J.; Mosier, K.L.; Burdick, M. Does automation bias decision-making? *Int. J. Hum.-Comput. Stud.* **1999**, *51*, 991–1006. [CrossRef]
35. Wyllie, S.; Shumailov, I.; Papernot, N. Feedback loop and bias amplification in recommender systems. In Proceedings of the CIKM '20: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, Virtual, 19–23 October 2020; pp. 2145–2148. [CrossRef]
36. Taori, R.; Hashimoto, T.B. Data feedback loops: Model-driven amplification of dataset biases. *arXiv* **2022**, arXiv:2209.03942. [CrossRef]
37. Wang, L.; Ma, C.; Feng, X.; Zhang, Z.; Yang, H.; Zhang, J.; Chen, Z.; Tang, J.; Chen, X.; Lin, Y.; et al. A survey on large language model based autonomous agents. *Front. Comput. Sci.* **2024**, *18*, 1–26. [CrossRef]
38. Gao, Y.; Xiong, Y.; Gao, X.; Jia, K.; Pan, J.; Bi, X.; Dai, Y.; Sun, J.; Wang, M.; Wang, H. Retrieval-Augmented Generation for Large Language Models: A survey. *arXiv* **2023**, arXiv:2312.10997. [CrossRef]
39. Shumailov, I.; Shumaylov, Z.; Zhao, Y.; Papernot, N.; Anderson, R.; Gal, Y. AI models collapse when trained on recursively generated data. *Nature* **2024**, *631*, 755–759. [CrossRef] [PubMed]
40. Wyllie, S.; Shumailov, I.; Papernot, N. Fairness Feedback loops: Training on synthetic data amplifies bias. In Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency, Rio de Janeiro, Brazil, 3–6 June 2024; pp. 2113–2147. [CrossRef]
41. Vidgof, M.; Bachhofner, S.; Mendling, J. Large Language Models for Business Process Management: Opportunities and Challenges. In *Business Process Management Forum. BPM 2023*; Lecture Notes in Business Information Processing; Springer: Cham, Switzerland, 2023; pp. 107–123. [CrossRef]
42. Grohs, M.; Abb, L.; Elsayed, N.; Rehse, J.-R. Large language models can accomplish business process management tasks. In *Business Process Management Workshops. BPM 2023*; Lecture Notes in Business Information Processing; Springer: Cham, Switzerland, 2024; pp. 453–465. [CrossRef]
43. Kourani, H.; Berti, A.; Schuster, D.; van der Aalst, W.M.P. Process Modeling with Large Language Models. In *Enterprise, Business-Process and Information Systems Modeling. BPMDS EMMSAD 2024*; Lecture Notes in Business Information Processing; Springer: Cham, Switzerland, 2024; pp. 229–244. [CrossRef]
44. Xi, Z.; Chen, W.; Guo, X.; He, W.; Ding, Y.; Hong, B.; Zhang, M.; Wang, J.; Jin, S.; Zhou, E.; et al. The rise and potential of large language model based agents: A survey. *Sci. China Inf. Sci.* **2025**, *68*, 121101. [CrossRef]

45. Piccialli, F.; Chiaro, D.; Sarwar, S.; Cerciello, D.; Qi, P.; Mele, V. AgentAI: A comprehensive survey on autonomous agents in distributed AI for industry 4.0. *Expert Syst. Appl.* **2025**, *291*, 128404. [CrossRef]
46. Yao, S.; Zhao, J.; Yu, D.; Du, N.; Shafran, I.; Narasimhan, K.; Cao, Y. REACT: Synergizing reasoning and acting in language models. *arXiv* **2022**, arXiv:2210.03629. [CrossRef]
47. Li, X. A review of prominent paradigms for LLM-Based agents: Tool use (Including RAG), planning, and feedback learning. *arXiv* **2024**, arXiv:2406.05804. [CrossRef]
48. Bai, X.; Wang, A.; Sucholutsky, I.; Griffiths, T.L. Explicitly unbiased large language models still form biased associations. *Proc. Natl. Acad. Sci. USA* **2025**, *122*, e2416228122. [CrossRef] [PubMed]
49. Zhao, J.; Wang, T.; Yatskar, M.; Ordonez, V.; Chang, K.-W. Men Also Like Shopping: Reducing Gender Bias Amplification using Corpus-level Constraints. *arXiv* **2017**, arXiv:1707.09457. [CrossRef]
50. Dietvorst, B.J.; Simmons, J.P.; Massey, C. Algorithm aversion: People erroneously avoid algorithms after seeing them err. *J. Exp. Psychol. Gen.* **2014**, *144*, 114–126. [CrossRef] [PubMed]
51. Vicente, L.; Matute, H. Humans inherit artificial intelligence biases. *Sci. Rep.* **2023**, *13*, 15737. [CrossRef] [PubMed]
52. Furnham, A.; Boo, H.C. A literature review of the anchoring effect. *J. Socio-Econ.* **2010**, *40*, 35–42. [CrossRef]
53. Johnson, E.J.; Goldstein, D. Medicine. Do defaults save lives? *PubMed* **2003**, *302*, 1338–1339. [CrossRef] [PubMed]
54. Van Der Aalst, W.M.P. *Process Mining: Data Science in Action*; CERN Document Server (European Organization for Nuclear Research): Geneva, Switzerland, 2016. [CrossRef]
55. Brynjolfsson, E.; Li, D.; Raymond, L. Generative AI at work. *Q. J. Econ.* **2025**, *140*, 889–942. [CrossRef]
56. Gans-Combe, C. Automated Justice: Issues, benefits and risks in the use of artificial intelligence and its algorithms in access to justice and law enforcement. In *Research Ethics Forum 2022*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 175–194. [CrossRef] [PubMed]
57. Yigit, Y.; Ferrag, M.A.; Ghanem, M.C.; Sarker, I.H.; Maglaras, L.A.; Chrysoulas, C.; Moradpoor, N.; Tihanyi, N.; Janicke, H. Generative AI and LLMs for Critical Infrastructure Protection: Evaluation Benchmarks, Agentic AI, Challenges, and Opportunities. *Sensors* **2025**, *25*, 1666. [CrossRef] [PubMed]
58. Wang, G.; Gunasekaran, A.; Ngai, E.W.; Papadopoulos, T. Big data analytics in logistics and supply chain management: Certain investigations for research and applications. *Int. J. Prod. Econ.* **2016**, *176*, 98–110. [CrossRef]
59. Lamine, E.; Thabet, R.; Sienou, A.; Bork, D.; Fontanili, F.; Pingaud, H. BPRIM: An integrated framework for business process management and risk management. *Comput. Ind.* **2020**, *117*, 103199. [CrossRef]
60. Humphreys, D.; Koay, A.; Desmond, D.; Mealy, E. AI hype as a cyber security risk: The moral responsibility of implementing generative AI in business. *AI Ethics* **2024**, *4*, 791–804. [CrossRef]
61. Acosta-Vargas, P.; Suarez, L.; Cuadrado, T.; Salvador-Ullauri, L. Mapping the Industry 5.0 landscape: Enabling Technologies, Human-Centered Systems, sectoral Applications, and SDG Alignment—A PRISMA-SCR review. *Technologies* **2026**, *14*, 268. [CrossRef]
62. Göndöcs, D.; Horváth, S.; Dörfler, V. Uncovering the dynamics of human-AI hybrid performance: A qualitative meta-analysis of empirical studies. *Int. J. Hum.-Comput. Stud.* **2025**, *205*, 103622. [CrossRef]
63. Kiss, A.N.; Libaers, D.; Barr, P.S.; Wang, T.; Zachary, M.A. CEO cognitive flexibility, information search, and organizational ambidexterity. *Strateg. Manag. J.* **2020**, *41*, 2200–2233. [CrossRef]
64. Sedera, D.; Gable, G.G. Knowledge management competence for enterprise system success. *J. Strateg. Inf. Syst.* **2010**, *19*, 296–306. [CrossRef]
65. Tabassi, E. *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2023. [CrossRef]
66. Regulation-EU-2024/1689-EN-EUR-LEX. 2024. Available online: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (accessed on 20 June 2026).
67. WITA. Recommendation of the Council on Artificial Intelligence—WITA. 2024. Available online: https://www.wita.org/atp-research/oecd-ai/ (accessed on 20 June 2026).
68. *ISO/IEC 42001:2023*; AI Management System. ISO: Geneva, Switzerland, 2023. Available online: https://www.iso.org/standard/42001 (accessed on 20 June 2026).
69. Mondal, S.; Das, S.; Vrana, V.G. How to Bell the Cat? A Theoretical Review of Generative Artificial Intelligence towards Digital Disruption in All Walks of Life. *Technologies* **2023**, *11*, 44. [CrossRef]
70. Mondal, S.; Uyen, N.C.T.; Das, S.; Vrana, V.G. Innovation Dynamics and Ethical Considerations of Agentic Artificial Intelligence in the Transition to a Net-Zero Carbon Economy. *Sustainability* **2025**, *17*, 8806. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.