COGNITIVE SECURITY IN THE AI AGE: HR'S STRATEGIC ROLE IN BUILDING THE INTELLIGENT HUMAN FIREWALL FOR SUSTAINABLE DECISION-MAKING

Dr. (Col) Virendra Mishra (Retd)[1*], Dr. (Col) Dinesh Kumar (Retd)[2]

[1*]*Professor of Practice at IPS Academy, Institute of Engineering & Science, Indore, vktmishra@gmail.com*
[2]*Professor of Practice at IPS Academy, Institute of Engineering & Science, Indore*
dickeykumar61@gmail.com

Article History
Received: 2026-03-25
Revised: 2026-05-05
Accepted: 2026-05-15
Published: 2026-05-28

Abstract

The accelerated integration of Artificial Intelligence (AI) into organizational processes has fundamentally transformed how individuals think, teams interact, and decisions are made. While AI enhances efficiency, analytical depth, and innovation capability, it simultaneously reshapes cognitive processes, interpersonal communication, and collective judgment. Over-dependence on algorithmic recommendations, diminishing reflective thinking, AI-mediated communication, and diffusion of responsibility increasingly threaten decision quality, team bonding, ethical conduct, and long-term organizational sustainability. In this context, many organizational failures no longer originate from technological breakdowns but from compromised cognition, behavioural drift, and erosion of trust. This paper advances the concept of cognitive security, defined as the organizational capability to safeguard human judgment, ethical reasoning, communication integrity, and behavioural discipline in AI-enabled environments. Anchored in Organizational Behaviour, strategic Human Resource Management, and leadership practice, the paper positions HR as the strategic custodian of cognitive security. The paper proposes an Intelligent Human (Cognitive) Firewall, operationalized through the OWNER framework, to ensure that inputs from data, analytics, AI tools, and human stakeholders are filtered through human judgment, accountability, and command responsibility. Reinforced by BMW (Body–Mind–Words), Emotional Intelligence, and the military ethic of "walking the talk," the framework aligns thought, communication, and action.

Keywords: *Cognitive Security, Artificial Intelligence, Organizational Behaviour;, Decision-Making, Sustainable Leadership, Human-in-Command, Algorithmic Bias, Cognitive firewall, OWNER,*

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

1. Introduction: The Cognitive Dimension of AI Risk

Artificial Intelligence (AI) is now emerging from its status as a support technology to a core part of the decision systems of an organization. From Human Resource Management (HRM) to finance, marketing, operations, governance and strategic planning, AI tools are more and more shaping the way leaders collect data, analyze patterns and trends, prioritize risks, allocate resources and make decisions. In today's day and age, organisations have come to depend on algorithmic systems for recruitment screening, performance evaluation, forecasting and predicting, customer targeting, optimizing workflows, fraud detection and predictive analytics. They have made the operational processes more efficient, increased data analysis speed, added speed in data processing, and improved decision support functions (Vrontis et al., 2023; Davenport & Ronanki, 2018). As AI becomes more influential, however, certain organizational vulnerabilities have emerged that are not rooted in any technological failure, but in the poorer cognitive function, impaired critical thinking, and overreliance on algorithmic suggestions of humans.

Modern Artificial Intelligence tools can generate believable, seemingly objective, fact-based, and analytically strong content. These systems can boost decision support systems, but can also reduce critical reflection by urging decision makers to take decisions with a lower level of reflection (Raisch & Krakowski, 2021). As AI becomes more common, the quality of judgment can gradually suffer, as well as the accountability mechanisms because of cognitive shortcuts, automation bias, confirmation bias, and taking AI advice without considering the reasons (Kahneman, 2011; Bashkirova & Krpan, 2024). The issue of technology working effectively no longer exists, but it is now whether or not human beings can think critically, understand the context appropriately, and hold themselves responsible for an organization's output.

Nowadays, algorithmic systems have entered into the decision-making process, and have heavily affected the traditional decision-making structures in organisations. AI is reshaping the role of authority in an organisation by altering the decision-making process between human and the intelligent systems, as per Shrestha et al. (2019). In the same vein, Lee (2018) suggests that trust, fairness perception and emotional responses are key factors in determining if people will trust an algorithmic recommendation. In fact, though, when the output is from an AI system, the employees and managers may view it as having more legitimacy if it appears only as a computation. This can hinder reflective evaluation and encourage the passive acceptance of what has been decided which may include some biases or assumptions that have not been explicitly identified and/or ethically questionable implications.

The problems are of particular importance in Human Resource Management. AI technologies are being increasingly used by businesses in the recruitment, employee performance evaluation, workforce analysis, and performance management process to improve efficiency and consistency.AI is now being used in companies to improve their recruitment process, AI-based performance management systems, employees' performance assessment, and workforce analysis, in order to gain efficiency and consistency. But, AI systems in HR can inadvertently perpetuate historic bias, amplify discrimination, and miss out on important human context factors like adaptability, emotional intelligence, creativity, and leadership potential (Charlwood & Guenole, 2022; Rigotti & Fosch-Villaronga, 2024). Empirical research on the use of AI in personnel selection also shows that when algorithmic recommendations do need to be nuanced with human evaluation (contextualization), decision makers often do so with little hesitation.Empirical research on the use of AI in personnel selection also indicates that when algorithmic recommendations need to be contextualized/ nuanced with human decisions, the decisions are taken with little hesitation. (Malin et al., 2024) This has made it a challenge for HR to navigate between technology efficiency and ethical, fair, transparent and human responsibility.

Thus, there is an emerging need for cognitive security, which is the theme of this paper, with the advent of AI-driven organizations. Cognitive security has a more holistic approach, safeguarding the integrity of human judgment, reflection, ethics, and decision ownership in AI-enabled environments, as opposed to protecting systems, networks, and data, as is done in traditional cybersecurity. Cognitive security aims to make it habit for leaders to use their thinking skills to interpret what they're seeing, to constantly question what they're seeing, to become aware of their own bias, and to take responsibility for the decisions they make, instead of delegating to algorithms. As the organizations start using sophisticated analytics, algorithm-driven HRM systems, self-efficient recommendation engines, and governance processes that are assisted by AI (Budhwar et al., 2022; Mikalef et al., 2022), this concern is gaining further significance.

Structured human oversight is also supported by the research on responsible AI and explainable artificial intelligence (XAI). Explainability, transparency, and accountability are now key issues in AI governance, as decision makers tend to not have a good understanding of how the outputs of the algorithms are generated

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07
(Haque et al., 2023). The responsible use of AI in business calls for the implementation of mechanisms that maintain human oversight and responsibility without compromising the ethical and transparent nature of AI technologies, as Papagiannidis et al. (2025) note. Also, Rodgers et al. (2023) have indicated that ethical considerations in HRM are increasingly becoming the domain of human judgment along with algorithmic systems balanced and accountable.
This is especially relevant in settings with significant stakes where outcomes of the decision have real consequences for the organization's employees, customers, and stakeholders, the organization's reputation, and the organization's long-term sustainability. AI-driven systems can streamline workflows, boost productivity, or enhance speed and efficiency, but fail to account for the context, emotional effects, social impact, and ethical considerations. As organisations develop, they may begin to drift into habits with their algorithmic convenience that begin to creep in and take the place of thoughtful reflection and critical analysis over time. This can lead to mis-hires, unfair promotions, tarnished company image, sub-optimal performance, loss of customer trust, and eroded company culture (Bankins et al., 2022; Varma et al., 2023).
This paper follows the military leadership practice and command responsibility system and postulates that organizationally, a discipline based human-in-command structure is needed even in technologically advanced environments. While advanced technologies and intelligence systems are available, the background of military leadership systems has focused on contextual evaluation, multi-layered review, clear accountability and reflective judgment. Decision makers are responsible for results, and it is not possible to assign responsibility for results to tools and/or advisory systems. The same logic is becoming more and more required in AI-powered businesses as managers are now required to be aware of the algorithmic recommendations and think about them before they are implemented.
The present study presents the concept of the Intelligent Human (Cognitive) Firewall which is realized with the OWNER framework: Own the Decision, Weigh Inputs Critically, Navigate Bias and Uncertainty, Explain Rationale, and Review and Learn. It aims to maintain the ethical reasoning, reflective judgment, accountability of humans, and integrity of communication in AI-enacted organizational systems. With reinforced by BMW (Body–Mind–Words: Mānasa–Vācā–Karmaṇā), the framework ensures that AI is a tool for support rather than as an invisible force driving results within the organization.
The paper also makes the case for cognitive security not to be just a behavioural and ethical issue; it is also a strategic capability within an organization with tangible business impact. Businesses that do not maintain reflective decision making could see more decision mistakes, less trust in the organization within, fewer good relationships with stakeholders, less governance, and less long-term sustainability. On the other hand, organizations that enhance the cognitive security can boost the quality of decisions, accountability, collaboration, trust, fairness perception, and organizational performance (Mikalef et al., 2023; Przegalinska et al., 2025). This makes HR not just the personnel department but also a strategic guardian of responsible use of AI and culture of sustainable decision making.

1.1 Aim and Objectives

Aim

To discuss the cognitive security of AI-infused organizations and how the decision ownership influences ethical and effective decision making.

Objectives

- To examine the impact of AI on cognition, team dynamics and decision quality
- To discuss the nature and quality of input and thought-decision loop, and suggest the accountability framework OWNER as a way to reinforce the cognitive security and accountability
- The focus of this paper will be on the relevance of human-in-command leadership in AI-supported organizational systems
- To employ an empirical case illustration from the use of AI in HR decision making to show the practical relevance of cognitive security and OWNER framework
- To recognize real-world management implications and business metrics of cognitive security in organizations

2. Literature Review

The swift incorporation of Artificial Intelligence (AI) into organisational systems has resulted in increasing scholarly interest in the consequences of its use for decision making, leadership, responsibility, ethics and

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07
organisational behaviour. Along with improving analytical ability, operational efficiency, prediction and automation, researchers are pointing out that overreliance on algorithmic systems could lead to a reduction in reflective judgment, a bias in decision making, and new types of vulnerability in organizations. The current literature thus identifies a basic dichotomy between the capabilities of AI and human cognition, especially when algorithms' suggestions impact managerial choices.
Thinking, Fast and Slow (Kahneman 2011), is a theoretical perspective that is fundamental in the study of decision-making that is relevant to AI-assistance. Kahneman differentiates between “System 1” and “System 2.”Kahneman refers to the intuitive “System 1” thinking and deliberate “System 2” thinking. System 1 thinking is fast, automatic and emotional, and System 2 thinking is slower, reflective, analytical and cognitively taxing. The algorithmic recommendations may be coming more and more into the picture for decision makers—often in the absence of critical consideration—because of the speed with which they are accepted. This can lead to an "automation bias" in which people rely more on the output generated by the system than on the information provided by people, because the output, which may appear objective and based on data, seems more highly intelligent than a human. This dependency can be gradually undermining the levels of critical appraisal, independent thought and conscious responsibility in the decision-making processes of an organization.
We have seen a tremendous growth in the adoption of AI in organizations, across HRM, finance, operations, marketing and strategic planning. In Artificial Intelligence for the Real World (2018), Davenport and Ronanki note that there is a tendency for organisations to take up AI technologies before establishing governance, oversight and accountability mechanisms. Many companies are adopting AI into their processes but don't have clear guidelines about the part of the human judgment plays in implementing AI-supported systems. This results in a governance vacuum that can lead to unconscious, non-ethical and unauditable human oversight of AI recommendations, which could affect organizational outcomes. The authors point out that sometimes, the focus of the organisation is only on the technological side of the equation and not on the behavioural, ethical and managerial implications of AI integration.
Shrestha, Ben-Menahem and von Krogh (2019) have discussed how organizational power structures are changing as a result of the use of AI. They have investigated a number of algorithmic systems, and their research in organizational decision-making structures under the influence of AI shows it has a profound impact on how authority, responsibility and decision rights are spread within organizations. The traditional hierarchical method relying exclusively on decisions makers for decisions making is becoming more complex and evolves and turns into hybrid human-AI decision making processes where the algorithm makes recommendations, predictions and prioritizations in the analysis. They can help to make them more efficient and consistent, but can also allow for ambiguity in terms of accountability and ownership. If not properly managed, human accountability can be perceived to be lost and lacking in AI environments, according to the authors.
To further elaborate on the tension of governance between automation and augmentation, Raisch and Krakowski (2021) talk about the “automation–augmentation paradox”. The authors suggest that the utility of AI for an organisation's performance depends on the ability of AI to augment the analytical thinking, information processing and situation awareness capabilities of humans. However, things can go wrong when organisations take AI systems so far, that they replace our reflective engagement and critical thinking. The (absurd) paradox is that the very same technology that can develop managerial competence can also inhibit it when managers simply take on board the algorithms. They stress the importance of humans to be aware and alert in AI-based institutions.
Research has also now come to the fore on trust and fairness in Algorithmic Decisions. Lee (2018) investigates the perception of algorithmic management systems by a person and demonstrates that the trust, fairness perception and emotional response are important factors to consider when assessing the acceptance of AI generated decision. As well as being technically correct, algorithmic systems are also evaluated by the employees and managers based on their perceived fairness, transparency and legitimacy. When decisions are made with AI, but they feel impersonal, opaque, and/or biased, then the trust and acceptance of the AI systems in the organization can go down. However, algorithmic recommendations can also be accepted as gospel or up-to-date without thinking or being aware of the assumptions or biases used in the algorithm. This highlights the human–AI interaction elements that need to be considered beyond technology and the importance of explainability, transparency and ethics.
Over the past few years, there have been a number of studies that have started to explore the impact of AI on the enterprise scale. Bankins et al. (2024) offer a multilevel analysis of how AI is being used in organizations

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

and suggest that it impacts all aspects of cognition, leadership behavior, communication, employee relationships, and organizational culture. The authors stress that the implementation of AI needs to be analysed not only as a technology change, but also as a behaviour and organisational change that will impact the thinking, working and decision-making of the individuals involved. Likewise Budhwar, et al. (2022) note that AI-based HRM systems open up opportunities and pose challenges to organisations. AI offers great potential to efficiency in recruiting, workforce analytics, and talent management, but it also has the potential to be unfair, intrusive, ethically questionable and lacking in human sensitivity when it comes to personnel decisions.

AI tools for HR decision-making also highlight the need for human oversight, as demonstrated by studies. In an experimental study of AI applications in digital HRM and/or personnel selection, Malin et al. (2024) demonstrate that HRM decision makers often use algorithmic recommendations in recruitment processes. While AI systems can help to identify patterns and screen candidates, they may not be as effective in considering contextual human attributes like adaptability, emotional intelligence, leadership, or ethical conduct. Given the potential for AI to provide more recommendations than it can possibly match, the study highlights the importance of human oversight and accountability in AI-driven HR systems.

Algorithmic fairness and ethical governance has also garnered much scholarly interest. If organizations over-rely on historical data or on opaque algorithms, AI-driven recruitment systems can inadvertently perpetuate social bias, discrimination and social exclusion, as suggested by Rigotti and Fosch-Villaronga (2024). Likewise, Rodgers et al. (2023) underscore the need for HRM practitioners to juggle between technological effectiveness and fairness, transparency, and accountability when making ethical decisions in HRM. The findings of their work indicate that there are structures and processes for managing the AI systems that need to be established so that they do not compromise ethical management.

The study also underscores the need for human accountability with a new study on responsible AI governance.A new study on responsible AI governance further emphasizes the need for human accountability. Papagiannidis, Mikalef, and Conboy (2025) state that in order to build a responsible AI adoption, it is crucial to have a governance framework that allows for maintaining human control, transparency, explainability and ethical oversight. The dangers of relying blindly on algorithm-based systems, hyper-automation, and losing managerial judgment are also pointed out by Mikalef et al., (2022) in the “dark side of AI.” Explanable Artificial Intelligence (XAI) has thus become an important research area due to the need of decision makers for explanations of how the AI generated recommendations are produced (Haque et al., 2023).

The issue of cognitive biases is further illustrated in behavioural studies on the process of decision-making with AI and how this process may amplify biases. Consistency with other experts' judgments in the recommendations made by AI can lead to more trust and acceptance, which can deepen confirmation bias, as illustrated by Bashkirova and Krpan (2024). The same phenomenon has been shown to be a major factor in human decisionmaking in high-stakes security contexts, such as national security, by Horowitz and Kahn (2024). In addition, Rosenthal-von der Pütten and Sach (2024) find that algorithmic recommendations can exacerbate discrimination in hiring processes when people select to only use the recommendations they deem relevant. All these studies suggest that algorithmic systems are not the solution to bias, but may exacerbate or perpetuate cognitive and structural biases if they are not constantly monitored by humans who are able to reflect on their operation.

In the field of public administration and public governance, Ruschemeier and Hondrich (2024) discuss automation bias from a legal and psychological interdisciplinary perspective. They also found that when it comes to decision making, people often rely on automation over contextual reasoning and even ethics. This makes the point that no technology can be relied upon to ensure responsible decision making. Rather, governance systems need to be in place that maintain critical reflection, the context and human accountability.

It is thus evident from the literature that there are some important themes. For starters, AI can substantially improve the organization's ability, efficiency and analytical capabilities. Secondly, over-reliance on AI can diminish reflective thinking, responsibility and moral judgment. Thirdly, organizations are increasingly seeking governance procedures that enable humans to play a role in AI-driven systems. Lastly, although a great deal of research has been conducted on the governance of AI, automation bias, fairness and algorithmic management, there has been little work examining how structured organisational capability for cognitive security, defined as the protection of integrity of judgment, ownership of decision, communication integrity and ethical responsibility in AI-enabled environments, can be crafted.

This study aims to fill that void by describing a new concept, Intelligent Human (Cognitive) Firewall, which can be implemented through the OWNER framework, Own the Decision, Weigh Inputs Critically, Navigate

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

Bias and Uncertainty, Explain Rationale, and Review and Learn. The framework is grounded in Human Resource Management, Organizational Behaviour, a human-in-command governance, and leadership practice to ensure that AI has a supporting function instead of replacing the reflective engagement, accountability and human judgment for decision making.

3. Research Methodology

The method applied in this study is a qualitative method which is an integrative design.

- Secondary Research. Literature search of decision science, AI-management, Organizational Behaviour and leadership.
- Experience-Based Analysis. The authors' military leadership experience in both tactical, operational and strategic positions with clearly defined accountability, provided the insights. Layered and structured evaluation and review promoted disciplined judgment and ownership. These principles were condensed into the acronym OWNER and adapted to work in the context of AI in the corporate world.
- Structured Observation. Thematic analysis of the ways people interact with AI in organizations and education.

This study is an exploratory and an interpretative study which does not carry out empirical statistical test. The Cognitive Firewall and OWNER models are abstract models, based on literature synthesis and structured experiential analysis, which are suggested for future empirical testing. The study includes experience-based analysis and literature synthesis, and is complemented with an empirical illustration from an application of AI in HR decision-making processes and the use of the OWNER framework.

4. Empirical Case Study: AI-Assisted Personnel Selection and Cognitive Security in HRM

In order to make the present study more practical and empirical, an empirical illustration in the field of Human Resource Management (HRM) with the assistance of artificial intelligence is included in the present paper. The case study focuses on the experimental study by Malin et al. (2024) that investigated the impact of AI-driven recommendations on human decisions in the context of recruitment. The study is very relevant to the current research because, in addition to the function of the organization, recruitment and selection is one of the most common functions that uses AI and where algorithmic systems are beginning to play a role in the selection of candidates, their positioning and the managerial judgement.

The empirical investigation involved using AI-based dashboard prototypes in personnel selection situations. The participants became the recruiters and were asked to assess the candidates for managerial roles based on their provided information about them and the AI-driven ranking system. The participants in the experiment were 93 who performed two personnel selection tasks under controlled conditions. The AI system gave out applicant rankings, performance indicators, and suitability scores to aid in the recruitment process.

The results revealed that participants were often prone to status quo and automation bias, which often meant that they tended to stick with the candidates they had scored high for, even if a different candidate was a better fit. The algorithmic recommendations were often relied upon by decision-makers, rather than detailed contextual information on the applicants. The study revealed three kinds of information search patterns: no information search, heterogeneous information search and homogeneous information search. Of these alternatives, homogeneous search coverage (where participants systematically assessed information on all candidates) led to the best quality of selection, and a more balanced result of the decisions they made.

The research also revealed that the more closely the decision makers signed up applicant data, the higher their quality of decision. Participants who actively followed up on the candidates that were provided beyond the AI provided rankings, did show more judgment and better selection results. Interestingly, there was no significant impact on overreliance on AI system recommendations when reminders and warnings of the potential for AI system inaccuracies were introduced. Despite being aware of possible weaknesses in AI-generated rankings, some participants still heavily relied on these rankings when making decisions.

The results are directly relevant to the main argument of the present paper about the ability of AI systems to enlist subtle effects on human cognition, attention allocation and decision behaviour. The lessons from the empirical evidence show that the main organizational danger in AI-supported settings is not only the failure of the AI itself to do its job but the erosion of the critical thinking of humans over time. This means that organisations need to have systems of thought protection, to keep people – humans – critically engaged, ethically conscious and responsible for their ultimate choices.

The case study also supports the OWNER framework that was proposed in this study. To make good decisions with AI, HR leaders and decision makers need to take a proactive approach to maintaining control

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07
of the decision, be mindful of the possibility of bias or uncertainty, verify the logic behind the decision, and consider the actions taken as a result of the decision to research and learn. In this context, cognitive security is one of the capabilities of the organization that is essential for the security of the judicial process and for not being dependent on algorithmic systems passively.
The key elements and outcomes of the empirical case study are summarized in Table 1.

Table 1. Empirical Evidence on AI-Assisted Personnel Selection and Cognitive Security

Aspect	**Description**
Study	Malin et al. (2024)
Research Context	AI-supported personnel selection in digital Human Resource Management
Research Objective	To examine how AI-generated rankings influence human decision-making and selection quality
Research Design	Experimental vignette study using AI-based dashboard prototypes
Participants	93 participants performing personnel selection tasks
Organizational Task	Evaluation and selection of candidates for managerial positions
AI Function	AI system generated applicant rankings and suitability indicators
Major Behavioural Observation	Participants focused heavily on highly ranked candidates (status quo bias)
Information Search Patterns	No search, heterogeneous search, and homogeneous search coverage
Key Finding	Homogeneous information evaluation improved decision quality
Additional Finding	Detailed review of applicant information strengthened judgment quality
Accountability Observation	Accountability reminders had limited effect on reducing AI overreliance
Organizational Risk	Automation bias and uncritical acceptance of algorithmic outputs
HRM Relevance	Highlights the need for reflective human oversight in recruitment decisions
Relevance to Present Study	Supports the need for cognitive security, human accountability, and the OWNER framework

The above evidence substantiates the fact that AI-based organizational systems can have a serious impact on an individual's information seeking behavior, alternative evaluation and decision making. The results indicate that without human supervision there is a need to have decision makers who are actively involved in reflective analysis and critical evaluation. The need for organizations to establish clear protocols to maintain accountability, transparency, and AI judgment integrity. This further reinforces the need for cognitive security and the Intelligent Human Firewall in the age of AI to maintain ethical, balanced and responsible decision making.

5. Cognitive Security: Concept and Organizational Relevance

Cognitive security is the systematic safeguarding of the integrity of thinking, ethical judgments and decision making by the leader in AI-driven settings. Cognitive security is different from traditional cybersecurity, which protects systems and data, by protecting the way leaders think, interpret information and make decisions (Mikalef et al., 2022).

Data are not the only influences on decision making – experience, mental models, feelings and organizational context are all factors. If the AI is used in a setting, the recommendations it provides can seem objective and optimized. Though these systems are contained within the data scope, they are not, however, "situational aware" nor do they have any moral responsibility. If the output produced by AI is not questioned, it may lead to cognitive bias in judgment, including automation bias, confirmation bias, and suggestion bias, along with others (Agudo et al. 2024; Bashkirova & Krpan 2024). The stakeholder pressure or selectively framed inputs could also influence the decisions. Using historical, limited datasets that are biased can propagate those biases, which can limit the ability of AI tools to deliver results in terms of human impact and long-term alignment and prioritize efficiency over impact.

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

In HRM cognitive security is a governance task. HR needs to prepare leaders for questioning the outputs from AI, being aware of biases and taking ownership of decisions. Incorporating reflective thinking and accountability measures into the leadership development and decision-making process can help ensure the quality of decisions and effectively link actions to long-term values and strategy (Budhwar et al., 2022; Varma et al., 2023).

5.1 The Influence–Outcome Loop
Quality of decision making is influenced before action. The inputs are used to create thoughts, the thoughts are used to create decisions, the decisions are used to create behaviour and the behaviour is used to create results (as shown in the Influence–Outcome Loop (Figure 1)). The inputs to the model are not just data and analytics, but also assumptions, past experience, narratives from stakeholders and AI-curated information (Bankins et al., 2024). These have an impact on the way that leaders recognise problems and evaluate risk.Narrow, biased, and incomplete inputs lead to constrained thinking. Technically correct decisions may then be strategically and/or ethically incorrect. Algorithmic systems can also reinforce existing trends or interests, creating a shift in perception in ways that are not obvious.
Fragile inputs create fragile outcomes when they are not based in critical thinking. Enhancing cognitive security, then, must be done at the input and cognition level, engaging in questioning of assumptions and filtering information as well as encouraging reflection, before problems are corrected once decisions are made (Horowitz & Kahn, 2024).

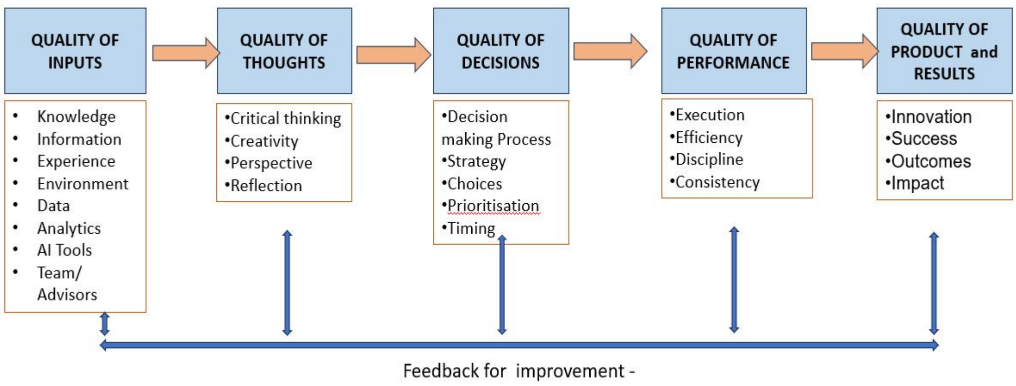

Figure 1. Influence Outcome loop – Author's ideation

5.2 Human-in-Command Leadership
AI-driven businesses can't have algorithms lead the way. AI can assist in the analysis, but the decision and responsibility must be left to humans. Algorithms see patterns, but do not have any contextual understanding, ethics or responsibility. With the automated and suggested environments, decision making can be affected by the leader's acceptance of the outputs of AI without questioning (Ruschemeier & Hondrich, 2024). Algorithmic framing can have a reconfiguring effect on behaviour and a bias effect on processes over time. One such example is the use of AI-powered Applicant Tracking Systems (ATS) in HR recruitment processes, which can rely on keyword matching. Even if the person is a good applicant, if the résumé doesn't meet the system requirements, he or she may be rejected. The system is able to put the decision on the two aforementioned outputs, if accepted without review (Rigotti & Fosch-Villaronga, 2024).
Human in command leadership means that AI suggestions are challenged, interpreted and deliberately adopted prior to enacting. The Cognitive Firewall assists its leaders to keep this practice, by keeping leaders active and responsible. This is because AI is still a tool that should be used alongside other resources and not as a replacement for them (Raisch & Krakowski, 2021).

5.3 The OWNER Framework: Operationalizing Cognitive Security
The OWNER framework is a high-stakes leadership practice which centers the decision making process on ownership and creates a framework that makes AI a part of leadership.The OWNER framework is a high

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

stakes leadership practice that provides clarity around decision ownership and establishes a framework for AI to be part of the leadership process. It has been built up on real experience gained in tactical, operational and strategic positions of responsibility, where accountability and consequences are clearly defined, and adapted to corporate governance environments. Though it is based on military leadership, its logic applies to organisations that are making decisions with the help of AI more and more. OWNER is suggested as a practice-informed approach to enhance the accountability of leaders in AI-supported systems (Rodgers et al., 2023; Papagiannidis et al., 2025).

The cognitive governance framework is translated to decision checkpoints:

O – Own the Decision. Authorities have to have authority.There cannot be authorities without authority. Leaders still are accountable for results, not algorithms.

W – Critically examine inputs. There must be a careful evaluation of AI outputs in the context of context, ethics, risk, stakeholder impact, and organization's priorities (Haque et al., 2023).

N – Navigate Bias and Uncertainty. Leaders are aware of and respond to the risks of automation bias, suggestion bias, confirmation bias and incomplete information. They understand that AI systems cannot be definitive or encompassing, and that they don't necessarily accurately describe the situation, cultural context, or emerging threats. Both intentional questioning and other points of view, as well as the discipline of judgment in the face of uncertainty, are needed to navigate the challenge of bias (Rosenthal-von der Pütten & Sach, 2024).

E – Explain Rationale. Communicate reasoning clearly to establish trust, alignment and transparency. AI gives people the "what" and the leader gives the "why. This shows the distinct value for humans added into the decision loop (Lee, 2018).

R – Review and learn. Structured reflection leads to continuous improvement and accountability (Agudo et al., 2024).

As shown in Figure 2, OWNER guarantees that the analysis is based on AI but not to the detriment of the human's judgement or responsibility.

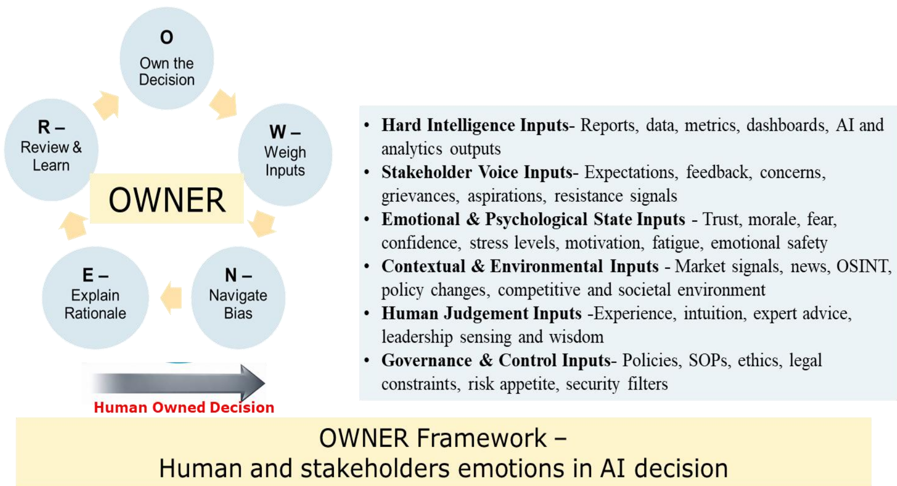

Fig 2. OWNER conceptual model- Author's ideation

5.4 The Intelligent Human (Cognitive) Firewall

The Cognitive Firewall (Figure 3) illustrates the process of the inputs of data, analytics, AI tools, and stakeholder views getting filtered through OWNER and then out to organizational action. Reflecting judgement, contextual awareness and ethical responsibility is institutionalized, maintaining the integrity of decisions (Mikalef et al., 2022; Papagiannidis et al., 2025).

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

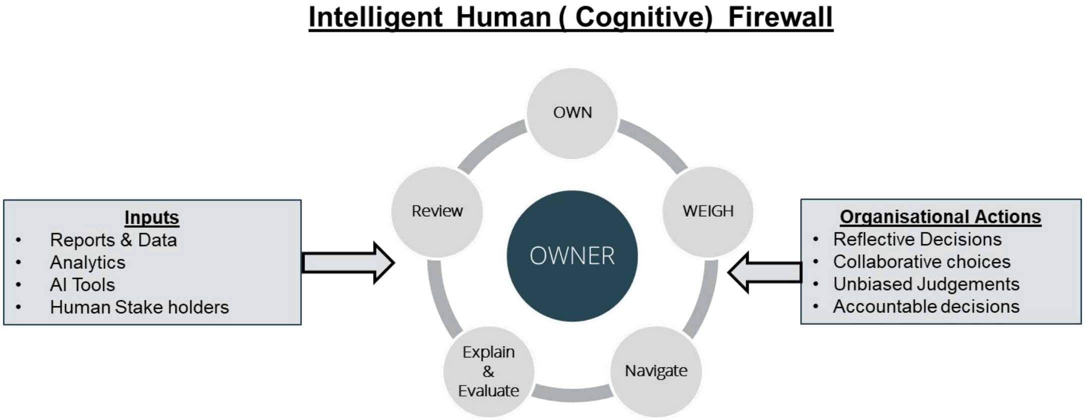

Fig 3. Intelligent Human Cognitive Firewall - Author's conceptual model

6. Countering Cognitive Bias and AI-Driven Influence: Strengthening Ethical Decision-Making

The key to an effective Cognitive Firewall is not just identifying the bias, but it's important to also challenge it. Biases like automation bias, confirmation bias and suggestion bias can creep into the judgment process in an AI environment. Inputs to algorithms can influence students' attention and beliefs, without their leaders being aware of it, and impact on their ethical clarity and decision quality (Bashkirova & Krpan, 2024; Horowitz & Kahn, 2024).

Organizations should consider implementing an easy-to-implement, yet structured, safeguards to ensure decision making is responsible:

- Structured Reflection. Take a break into decisions. Leaders need to reflect on what assumptions are made in this data? What could be lacking? Would I have made the same decision if I hadn't used AI? This lessens the tendency to unquestioningly accept.
- Team Discussion & Brainstorming. The use of AI to make decisions should be clearly communicated prior to implementation. Teams should review data sources, contextual fit and risks, as well as alternatives. Asking the right questions before they do so, helps them to become more engaged and engaged questioning helps to prevent them from complying silently.
- Pilot Testing. Test AI-influenced decisions on a small scale, as possible, prior to rollout. Pilot runs will minimize unwanted side effects and minimise systemic risk.
- Diverse Inputs. Utilize data from several sources and cross functional review. Perspectives alleviate narrow-mindedness and confirmation bias (Bankins et al., 2024).
- AI Literacy & Bias Awareness. Educate AI leaders on the limitations of AI. When we are aware it means that we are less dependent and better informed of what is right or wrong (Charlwood & Guenole, 2022).
- Documentation and Review. Document reasons for decisions, and carry out post decision reviews. Learning from outcomes enhances accountability and a reduction of the potential for repeated bias.

Reflection, discussion, pilot testing and review work together to move organizations from being dependent on AI to having an AI-governed organization, based on the discipline of humans in command. This makes the Cognitive Firewall a viable solution for safeguarding judgement, building trust and maintaining the integrity of the organization, while enhancing it (Varma et al., 2023).

7. BMW Alignment: Translating Ownership into Behaviour

An accountability for decisions is ensured by OWNER while ownership is ensured by its reflection in daily behaviour (Bankins et al., 2022) by BMW (Body – Mind – Words — Mānasa, Vācā, Karmaṇā).

- Mind (Mānasa). When it comes to cognitive security, it all starts in the mind. The leaders need to be aware of their own thoughts and ask questions of AI advice, look for cognitive bias and avoid emotional or impulsive reaction. Our thoughts repeat often lead to our decisions. Relying on filtering the inputs, taking a step back before accepting them and consciously considering what they mean helps to avoid automation bias to influence behaviour (Kahneman, 2011).
- Words (Vācā). Leaders to make clear why AI-informed decisions. Clear communication fosters trust and helps to avoid the blame game. A team that respects its members and is psychologically safe, will have leaders who are transparent about how the decision was made (Lee, 2018).
- Body / Action (Karmaṇā). The action needs to be in line with the thinking and words. When it's done right, it helps establish credibility, and helps you remain aligned with organizational values. Leaders whose

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

values are lived out on their behavior engage in accountability practices that are visible and culture is maintained (Przegalinska et al., 2025).

BMW combines emotional intelligence, discipline and integrity. It promotes the taking up and living out of decision ownership, which is more than a claim, it is a commitment, a responsibility and a practice.

8. Recommendations

Increasing HRs awareness of cognitive security is needed to make it a constant practice in their daily routine (Budhwar et al., 2022).

Consider the following terms:

Individual Level – Personal Accountability. Leaders should be educated to stop and ask questions on suggestions from AI before implementing. Questioning such things as "What may be missing?", "What is the risk?", and "Am I owning this decision?" can be incorporated into leadership courses. Leaders can be mindful of bias and responsibility by keeping reflection notes and conducting review conversations. (Haque et al., 2023)

Relevant to the team level, open discussion and healthy questioning is introduced. If a team is implementing a large decision based on AI, it would be good practice to discuss and challenge that recommendation prior to taking action on it. Encourage Ground feedback, alternate views and brainstorming. Sceptile's job is to ask questions of AI outputs, not challenge them. Brief review meetings can be used after key decisions to capture the lessons (Bankins et al., 2024).

This requires clear accountability at the organizational level and good governance. This needs good governance and clear accountability at the organizational level. It is essential to communicate that AI is a guideline, not a decision-maker, to the organizations. Small pilot runs should take place before any major decision is made and it should be documented that humans have approved it. Evaluation of decision ownership, fairness and quality of judgment are as important as results in performance appraisals, for HR purposes (Papagiannidis et al., 2025; Rodgers et al., 2023).

However, if leadership takes ownership in decisions and promotes thoughtful discussion, AI becomes a supporting factor to decision making and not a silent force to outcomes.

9. Conclusion

Because of the AI age, organizations are far more capable analytically but also have new cognitive risks at hand. The biggest danger is the loss of reflective judgment, accountability and ethical clarity over a period of time, rather than any technological failures. If decision ownership is not maintained and leaders take decisions at face value from the algorithm they can introduce bias through the system. Cognitive security is thus an imperative need. It will need to involve safeguarding the quality of thinking that comes before it is applied in action – making sure that AI does not take the place of human accountability in decision-making. The OWNER framework and the support of the Human-in-Command principle, bolstered by BMW (Mānasa–Vācā–Karmaṇā), offers a structured approach to ensure the integrity of judgment in AI-enabled settings. This is not a technical, but a leadership matter for Human Resource Management. HR needs to embed reflective practices and awareness of biases, structured review and obvious accountability into decision processes. Empirical illustration of AI for HR decision-making also underscores the need for human oversight, which should be thoughtful, accountable, and enacted, especially when algorithmic tools inform decisions that have significant consequences for individuals, like hiring or firing. The leaders' discipline, awareness and accountability will prove to be the key to sustainable innovation and responsible entrepreneurship in the AI age, not the intelligence of the algorithms.

10. Limitations and Future Research

This study has an interpretive and conceptual approach. Empirical studies in the future could evaluate the measurable effects of the OWNER framework in organizations with AI systems, to better understand how AI can positively affect the quality of decision making, psychological safety, and organizational governance.

References

1. Agudo, U., Liberal, K. G., Arrese, M., & Matute, H. (2024). The impact of AI errors in a human-in-the-loop process. *Cognitive Research: Principles and Implications*, *9*(1), 1.

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07
2. Bankins, S., Formosa, P., Griep, Y., & Richards, D. (2022). AI decision making with dignity? Contrasting workers' justice perceptions of human and AI decision making in a human resource management context. *Information Systems Frontiers*, *24*(3), 857-875.
3. Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of organizational behavior*, *45*(2), 159-182.
4. Bashkirova, A., & Krpan, D. (2024). Confirmation bias in AI-assisted decision-making: AI triage recommendations congruent with expert judgments increase psychologist trust and recommendation acceptance. *Computers in Human Behavior: Artificial Humans*, *2*(1), 100066.
5. Budhwar, P., Malik, A., De Silva, M. T., & Thevisuthan, P. (2022). Artificial intelligence–challenges and opportunities for international HRM: a review and research agenda. The InTernaTIonal Journal of human resource managemenT, 33(6), 1065-1097.
6. Charlwood, A., & Guenole, N. (2022). Can HR adapt to the paradoxes of artificial intelligence?. *Human Resource Management Journal*, *32*(4), 729-742.
7. Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard business review*, *96*(1), 108-116.
8. Gong, Q., Fan, D., & Bartram, T. (2025). Algorithmic human resource management: toward a functional affordance perspective. *Personnel Review*, *54*(5), 1150-1177.
9. Haque, A. B., Islam, A. N., & Mikalef, P. (2023). Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research. Technological Forecasting and Social Change, 186, 122120.
10. Horowitz, M. C., & Kahn, L. (2024). Bending the automation bias curve: A study of human and AI-based decision making in national security contexts. *International Studies Quarterly*, *68*(2), sqae020.
11. Kahneman, D. (2011). *Thinking, fast and slow*. macmillan.
12. Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big data & society*, *5*(1), 2053951718756684.
13. Malin, C. D., Fleiß, J., Seeber, I., Kubicek, B., Kupfer, C., & Thalmann, S. (2024). The application of AI in digital HRM–an experiment on human decision-making in personnel selection. *Business Process Management Journal*, *30*(8), 284-312.
14. Mikalef, P., Conboy, K., Lundström, J. E., & Popovič, A. (2022). Thinking responsibly about responsible AI and 'the dark side'of AI. *European Journal of Information Systems*, *31*(3), 257-268.
15. Mikalef, P., Islam, N., Parida, V., Singh, H., & Altwaijry, N. (2023). Artificial intelligence (AI) competencies for organizational performance: A B2B marketing capabilities perspective. *Journal of Business Research*, *164*, 113998.
16. Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2), 101885.
17. Prikshat, V., Malik, A., & Budhwar, P. (2023). AI-augmented HRM: Antecedents, assimilation and multilevel consequences. *Human Resource Management Review*, *33*(1), 100860.
18. Przegalinska, A., Triantoro, T., Kovbasiuk, A., Ciechanowski, L., Freeman, R. B., & Sowa, K. (2025). Collaborative AI in the workplace: Enhancing organizational performance through resource-based and task-technology fit perspectives. *International Journal of Information Management*, *81*, 102853.
19. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of management review*, *46*(1), 192-210.
20. Rigotti, C., & Fosch-Villaronga, E. (2024). Fairness, AI & recruitment. *Computer Law & Security Review*, *53*, 105966.
21. Rodgers, W., Murray, J. M., Stefanidis, A., Degbey, W. Y., & Tarba, S. Y. (2023). An artificial intelligence algorithmic approach to ethical decision-making in human resource management processes. *Human resource management review*, *33*(1), 100925.
22. Rosenthal-von der Pütten, A. M., & Sach, A. (2024). Michael is better than Mehmet: exploring the perils of algorithmic biases and selective adherence to advice from automated decision support systems in hiring. *Frontiers in Psychology*, *15*, 1416504.
23. Ruschemeier, H., & Hondrich, L. J. (2024). Automation bias in public administration–an interdisciplinary perspective from law and psychology. *Government Information Quarterly*, *41*(3), 101953.
24. Shrestha, Y. R., Ben-Menahem, S. M., & Von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California management review*, *61*(4), 66-83.

Vol. 12 Issue 2s (May, 2026)
ISSN: 2208-2190 Pages- 539-551
DOI: https://doi.org/10.69980/y9vz8f07

25. Varma, A., Dawkins, C., & Chaudhuri, K. (2023). Artificial intelligence and people management: A critical assessment through the ethical lens. *Human Resource Management Review*, *33*(1), 100923.
26. Vrontis, D., Christofi, M., Pereira, V., Tarba, S., Makrides, A., & Trichina, E. (2023). Artificial intelligence, robotics, advanced technologies and human resource management: a systematic review. *Artificial intelligence and international HRM*, 172-201.

Disclosure: The ideas and conceptual frameworks presented in this paper are the original work of the authors. AI-assisted tools were used solely for linguistic refinement and formatting support.