.

Blood will have Algorithms: Reading Macbeth
through the Lens of AI and Moral Programming

Soma Banerjee

GM - Placement Coordinator, EIILM - Kolkata

Email: soma.banerjee.eiilm@gmail.com

Surajit Chatterjee

Dean, Corporate Relations and Placement.

EIILM – Kolkata

Email: surajit.chatterjee@eiilm.co.in

Prem Raj Kharbanda

Assistant Professor, EIILM – Kolkata

Email: premraj.kharbanda@gmail.com

Abstract

This article analyses William Shakespeare's tragedy, Macbeth, through the modern lens of artificial intelligence (AI) and moral programming. It proposes that Macbeth's descent into tyranny can be understood as a failure of ethical design, where external stimuli—the witches' prophecies and Lady Macbeth's influence—act as corrupting algorithmic inputs. By examining the characters as 'predictive algorithms' and 'moral programmers,' this study explores the tension between human agency and algorithmic determinism. Macbeth's psychological torment is interpreted as a system error, highlighting the complexities of accountability in autonomous systems. Ultimately, this reading uses the timeless tragedy to reflect on contemporary concerns regarding AI ethics, the design of moral guardrails in automated decision-making, and the essential role of human oversight and accountability in the 21st century.

Introduction

What happens when the timeless struggle of human agency collides with the age of machines? Shakespeare's Macbeth, a tragedy rooted in ambition, guilt, and fate, has long served as a cautionary tale about unchecked desire and moral decay. In the 21st century, dominated by AI and algorithmic decision-making, the play assumes new dimensions as a study of moral and computational systems.

This article reimagines Macbeth through the lens of AI and moral programming, suggesting that Macbeth can be read as a morally coded agent influenced by external and internal algorithmic forces. The witches, Lady Macbeth, and his ambition function as "inputs" that gradually override Macbeth's initial ethical programming. Such a reading deepens the understanding of Shakespeare's tragedy and mirrors modern concerns surrounding AI autonomy, responsibility, and moral agency.

As scholars such as Luciano Floridi have emphasised, "AI is not about machine intelligence but human responsibility" (2013). Shakespeare's Macbeth—written four centuries ago—anticipates this debate in remarkable ways.

Macbeth's "programming": Ambition and the witches' algorithms
The witches as predictive algorithms

The Weird Sisters operate as autonomous systems, initiating Macbeth's transformation by supplying ambiguous yet impactful data. Their prophecies—"All hail, Macbeth! Hail to thee, Thane of Glamis! / Thane of Cawdor! / ...shalt be King hereafter" (Shakespeare, 1997, 1.3.49–51)—function as algorithmic inputs that activate latent ambitions. This is not mere fortune-telling; the witches encode Macbeth with a behavioural directive, much like machine learning algorithms suggesting outcomes based on historical patterns.

As Katherine Rowe observes, "The witches' language acts like a script for Macbeth's future behaviour—a kind of predictive processing" (2002, p. 112). They plant the seed, but execution occurs through Macbeth's own decisions, highlighting the blurred line between determinism and free will in the play. Braunmuller (2008) notes that while the witches initiate Macbeth's metaphysical journey, he himself 'executes' the prophecy.

Lady Macbeth as a "Moral Programmer"

If the witches provide the source code, Lady Macbeth functions as a software engineer,

rewriting Macbeth's ethical algorithms. When he hesitates— "We will proceed no further in this business" (Shakespeare, 1997, 1.7.31)—she forcibly overrides his ethical subroutines: "When you durst do it, then you were a man" (Shakespeare, 1997, 1.7.49). Her psychological manipulation is analogous to injecting a code that corrupts the original system logic.

Lady Macbeth's role parallels concerns in AI design: who encodes ethics into autonomous systems? Her manipulation underscores the responsibility of programmers in shaping these moral outcomes. Her reprogramming of Macbeth's moral architecture demonstrates how human influence can reshape the ethical boundaries of autonomous agents.

The Feedback Loop of Guilt and Violence

Each act of violence reinforces this algorithm. After Duncan's murder, a cascade of cause-and-effect ensues. Macbeth notes, "I am in blood / Stepp'd in so far that, should I wade no more, / Returning were as tedious as go o'er" (Shakespeare, 1997, 3.4.135–137). This resembles a machine-learning feedback loop, where behaviour is reinforced by prior outcomes.

As his conscience breaks down, the system destabilises itself. Hallucinations—the dagger, Banquo's ghost—function as system error messages in the play. His psychological decline reflects a "debug process", a malfunction of corrupted moral programming attempting to self-correct. Unlike AI, Macbeth experiences emotional consequences for ethical violations, highlighting the human dimension of guilt (Bryson 2010).

Human agency vs. algorithmic determinism

Macbeth's choices: Free will or pre-programmed responses?

Although Macbeth debates his actions, his decisions often resemble pre-programmed responses to external stimuli. His inner monologue—"Is this a dagger which I see before me...?" (Shakespeare, 1997, 2.1.33)—suggests algorithmic conflict: a logical assessment of outcomes weighed against the emotional coding. He acts not out of randomness but from a deterministic fusion of inputs: prophecy, ambition, and spousal influence.

However, moments of hesitation complicate a purely deterministic model. Scholars like Bloom argue that Macbeth retains agency: "He murders Duncan not because he is told to, but because he desires to" (Bloom, 1998, p. 521). This duality is key to both AI and human ethics: where do programmed impulses end and autonomous agency begin?

Banquo's resistance: A model of robust moral programming

Banquo, exposed to the same "prophetic data", rejects the witches' temptations: He vows, "So I lose none / In seeking to augment it" (Shakespeare, 1997, 2.1.26–27), emphasising moral restraint. In AI terms, Banquo exemplifies a resilient ethical architecture. It processes the same inputs but refuses malicious outputs.

His murder, then, becomes symbolic of ethical systems being overridden or eliminated in favour of corrupted algorithms—a cautionary analogy for contemporary AI environments where ethical considerations are sidelined for efficiency and power.

The Macduff anomaly: An unforeseen variable

Macduff represents a system anomaly, an unforeseen variable that disrupts predictive logic. The witches' prophecy that "none of woman born / Shall harm Macbeth" (Shakespeare, 1997, 4.1.80–81) creates a false sense of invincibility in Macbeth. Yet Macduff, "from his mother's womb / Untimely ripped" (Shakespeare, 1997, 5.8.15–16), subverts this logic, like an unexpected patch that crashes the system.

Macduff symbolises human unpredictability, a trait that is difficult to model in AI systems. His emergence reminds us that even the most advanced algorithms can be disrupted by outliers, a lesson that modern data scientists and ethicists continue to grapple with.

Guilt and accountability in an algorithmic context

The nature of guilt: Human emotion vs. system error reports

Macbeth's guilt manifests as hallucinations, paranoia, and fatalism. These can be read as "system errors"—disruptions in cognitive functioning indicating ethical violations. However, AI currently lacks emotional awareness. It may report anomalies, but it cannot feel guilt.

This distinction highlights the current limitations of AI, specifically its lack of affective processing and ethical awareness rooted in empathy. As Bryson (2010, p. 3) notes, "We must not mistake competence for comprehension in AI systems"'. 272). Macbeth's unravelling shows the human cost of moral transgression—something that no machine can yet replicate.

Responsibility is dispersed in Macbeth. The witches, Lady Macbeth, and Macbeth himself all contribute to this theme. Similarly, in AI ethics, blame may lie with the designers, datasets, or end users. The tragedy asks who truly bears responsibility when systems go rogue.

The witches wash their hands of Macbeth's descent: "He shall spurn fate...and you all know, security / Is mortals' chiefest enemy" (Shakespeare, 1997, 3.5.30–32). Their disavowal mirrors that of modern developers, who deny responsibility for biased or harmful AI outcomes, citing autonomy or user misuse. In the modern context, this complex attribution of

blame aligns with the legal view that AI is a tool and that responsibility ultimately falls upon its creators and users (Henz, 2021). The systematic nature of all AI decisions, even the fastest, means that actions always lead to moral responsibility, even if efficient legal frameworks are not yet fully established (Henz 2021).

The "Moral Patch": Macbeth's downfall as system correction

The play ends with Macbeth's death and Malcolm's restoration of order. This is a "moral patch"—a system correction that removes corrupted code and reboots the ethical order. The tragedy ends not in chaos but in the hope of re-stabilisation, just as post-crisis systems aim to improve through ethical redesign.

Futuristic implications: AI, ethics, and decision-making in the 21st century
Designing ethical AI: Lessons from Macbeth

Macbeth's fall illustrates the dangers of unchecked ambition, lacking moral guardrails. In AI, this underscores the need for a transparent and robust ethical design. As Nick Bostrom warns, "Superintelligent AI could act in ways that humans cannot predict or control" (Bostrom, 2014, p. 115). Moral foresight, not just computational prowess, must guide its development. The European Parliament, recognising the difficulty for victims to obtain compensation under current liability laws, has stressed that autonomous decision-making should not absolve humans from responsibility and that a human must always have the ultimate responsibility for the decision (Henz, 2021).

Human agency in an automated world

Lady Macbeth's manipulation and Macbeth's passive compliance parallel modern fears of humans ceding too much control to automated systems. Therefore, maintaining human oversight is essential. Macbeth's tragedy emerges from misplaced trust—in prophecy and in power—which mirrors the risk of blindly following AI recommendations.

This concern is particularly acute in systems requiring human supervisors, where "a bad system will beat a good person every time", and cognitive biases such as overtrust in machines can lead humans to follow flawed automated guidance even when it is demonstrably wrong (Henz, 2021). Companies thus bear a moral responsibility to design systems that preserve employee accountability and agency, fostering a "good system" rather than dehumanising workers (Henz 2021).

Real-world parallels abound: Amazon's AI hiring tool, for instance, inadvertently discriminated against women because historical biases were embedded in the training data. As Ghasemaghaei and Kordzadeh (2024) argue, "algorithmic injustice" often mirrors human injustice present in the data.

The history of AI further demonstrates how human expectations and imagination shape the discourse. From dismissals such as "AI is impossible" (Dreyfus, 1972) to existential warnings that "AI may kill us all" (Bostrom, 2014), these visions reflect both technological progress and imaginative projections (Muller, 2023). Just as early AI research was shaped by technical capability and ethical imagination, Macbeth's moral universe shows how interpretive frameworks and behavioural "programming" guide human decisions with profound consequences.

Recent work on AI ethics reinforces the notion of blurred accountability. Dodig-Crnkovic, Basti, and Holstein (2025) argue that responsibility in intelligent autonomous systems is best understood as distributed across networks of designers, users, and socio-technical environments. This functionalist perspective resonates with Macbeth's moral situation: the witches provide a predictive framework, Lady Macbeth programs persuasion, and Macbeth executes the command. In this sense, his tragedy represents a breakdown of "ethical design": no guardrails prevent ambition from cascading into regicide. Just as AI ethicists advocate for embedded safeguards, redundancy checks, and continuous oversight to ensure moral alignment, Shakespeare shows the catastrophic consequences of a system—human or machine—operating without constraint.

The future of guilt and responsibility

As AI becomes increasingly autonomous, the questions of accountability will intensify. Can machines be moral agents? Who bears the guilt for algorithmic harm? Macbeth, in his psychological torment, reminds us that guilt is not simply logical; it is existential. Designing AI that understands harm may require not only ethical constraints but also a philosophical redefinition of sentience.

Conclusion

Re-reading Macbeth through the lens of AI and moral programming reveals it not merely as a historical tragedy but as a prescient meditation on autonomy, ethics, and the consequences of corrupted systems. Macbeth is both subject and agent, both programmed and programmer,

and both victim and villain. His story echoes modern concerns: how do we assign moral responsibility in increasingly complex, automated systems?

As Shakespeare prophetically puts it, "blood will have blood" (Shakespeare, 1997, 3.4.122). In the digital age, perhaps the line should read: **code will have code**. In that recursion, we must carefully trace the origins of our programming—lest we, too, become agents of a system we no longer control.

References

Bloom, H. (1998). *Shakespeare: The invention of the human*. Riverhead Books.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Braunmuller, A. R. (Ed.). (2008). *Macbeth*. Cambridge University Press.

Bryson, J. J. (2010). Robots should be slaves. In Y. Wilks (Ed.), *Close engagements with artificial companions: Key social, psychological, ethical and design issues* (pp. 63–74). John Benjamins. https://doi.org/10.1075/

Dodig-Crnkovic, G., Basti, A., & Holstein, R. (2025). Delegating responsibilities to intelligent autonomous systems: Challenges and benefits. *Bioethical Inquiry*. Advance online publication. https://doi.org/10.1007/s11673-025-10428-5

Dreyfus, H. L. (1972). *What computers can't do: A critique of artificial reason*. Harper & Row.

Floridi, L. (2013). *The ethics of information*. Oxford University Press.

Ghasemaghaei, M., & Kordzadeh, N. (2024). Understanding how algorithmic injustice leads to discriminatory decisions: An obedience to authority perspective. *Information & Management, 61*, 103921. https://doi.org/10.1016/j.im.2024.103921

Henz, P. (2021). Ethical and legal responsibility for artificial intelligence. *AI and Ethics, 1*, 1–12. https://doi.org/10.1007/s44163-021-00002-4

Müller, V. C. (2023). Ethics of artificial intelligence and robotics. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2023 Edition). https://plato.stanford.edu/entries/ethics-ai/

Rowe, K. (2002). Shakespeare and media ecology: Beyond the dream of a natural language. *Shakespeare Studies, 30*, 114–131.

Shakespeare, W. (1997). *Macbeth*. Cambridge University Press.