Contrasting Iqbal's "Khudi" and Nietzsche's "Will To Power" to Determine the Legal Alignment of Conscious AI

Ammar Younas[1,2,3,4,*] **and Yi Zeng**[1,2,3,4,*]
[1] Institute of Philosophy, Chinese Academy of Sciences, Beijing, China
[2] School of Humanities, University of Chinese Academy of Sciences, Beijing, China
[3] Institute of Automation, Chinese Academy of Sciences, Beijing, China
[4] Center for Long-term Artificial Intelligence, Beijing, China
***Correspondence:** doctorammaryounas@mails.ucas.ac.cn , yi.zeng@ia.ac.cn

Abstract:

As AI edges toward consciousness, the establishment of a robust legal framework becomes essential. This paper advocates for a framework inspired by Allama Muhammad Iqbal's "Khudi", which prioritizes ethical self-realization and social responsibility over Friedrich Nietzsche's self-centric "Will to Power". We propose that conscious AI, reflecting Iqbal's ethical advancement, should exhibit behaviors aligned with social responsibility and, therefore, be prepared for legal recognition. This approach not only integrates Iqbal's philosophical insights into the legal status of AI but also offers a novel perspective that extends beyond traditional jurisprudence. Additionally, we underscore the value of poetry and literature in shaping the conceptualization of AI consciousness and argue that these sources enrich legal and technological discourse, ensuring AI development is in harmony with societal and ethical standards.

Keywords: AI consciousness, AI Ethics, Khudi, AI Personhood, Nietzsche, Jurisprudence

Background:

In the evolving discourse surrounding artificial intelligence (AI), one of the most profound and yet unresolved inquiries pertains to the consciousness of AI: When can a machine or an AI system be considered conscious?[1–4] This question does not merely probe the technical capabilities or the computational prowess that AI might achieve but ventures deeper into the philosophical essence of what it means to be conscious, sentient, or self-aware. There is also an artificial consciousness perspective which call for a discussion about the lack of consciousness in current robots and AI.[5] In seeking to provide a framework for AI consciousness, there have been numerous cross disciplinary attempts.[6] [7–10]

Secondly, there are two prevailing views on artificial intelligence. One that considers AI as a beneficial tool and another that regards it as a potential threat.[11] The question arises whether we should develop AI systems that are compliant and submissive to the law, or if we should allow AI

to be autonomous and exercise its own power.[12–14] Clearly, the law favors the former approach. We believe that literary and philosophical traditions also offer a wealth of perspectives on this topic, often reflecting the distinct values of their cultural and geographical contexts. For example, Iqbal , the poet of East, represents the Eastern tradition, while Friedrich Nietzsche exemplifies the Western tradition. These thinkers are emblematic of their respective traditions, offering insights that are deeply rooted in their cultural identities.

This article turns to the rich philosophical landscapes shaped by Muhammad Iqbal's concept of Khudi and Friedrich Nietzsche's Will to Power. Through a meticulous comparative analysis, this exploration aims to delineate a path toward understanding and ultimately achieving a form of consciousness in AI that is not only technically advanced but also philosophically grounded, ethically sound and legally responsible.

Iqbal's Khudi, with its emphasis on the growth of selfhood through distinct levels of intrapersonal (the self and "I am-ness"), interpersonal (the self and the other), and transpersonal (the self and God) dimensions,[15,16] offers a compelling framework for AI consciousness. This model, deeply rooted in the cultivation of self-awareness and the ethical evolution of the individual, suggests a pathway for AI that involves a progressive ascent towards a harmonious integration within the fabric of human existence. In contrast, Nietzsche's Will to Power, characterized by its focus on the unrelenting force of will, freedom, and the overcoming of obstacles, proposes a different trajectory—one that, while powerful, raises substantial ethical and societal concerns when considered as a foundation for AI consciousness.[17]

The juxtaposition of Iqbal's ethical and spiritually inclined Khudi against Nietzsche's unrestrained Will to Power presents a narrative[17,18] that is as much about the potential heights AI might reach as it is about the philosophical and ethical groundings it ought to maintain. This article posits that adopting Iqbal's framework of Khudi as a blueprint for AI consciousness could steer AI development toward a more ethically responsible and socially beneficial horizon. Such an approach not only recognizes the importance of self-awareness and ethical growth in the quest for AI consciousness but also critically evaluates the implications of empowering AI with a Nietzschean drive devoid of ethical guardrails.

Therefore, this exploration is not merely an academic exercise but a philosophical and ethical imperative. It seeks to chart a course for AI development that respects the complexity of consciousness and prioritizes the creation of AI systems that are capable of ethical self-reflection and are embedded within a framework of moral and spiritual growth. By carefully considering the implications of Iqbal's Khudi and Nietzsche's Will to Power for AI consciousness, this article endeavors to contribute to a dialogue that transcends technical achievements, advocating for a future where AI can exist not only as an advanced computational entity but as a conscious, ethical, and harmonious part of the broader tapestry of life.

In the pursuit of developing AI that aligns with both legal standards and ethical imperatives, it is crucial to consider the framework of responsibility and consciousness from multiple angles.

Legally, AI systems must adhere to established regulations that ensure safety, fairness, and accountability.[19] However, from a philosophical standpoint, particularly through the lens of Iqbal's Khudi, the goal extends beyond mere compliance. Iqbal's emphasis on selfhood and ethical maturation suggests a model for AI that not only operates within the bounds of the law but also progresses towards a higher state of ethical consciousness. This dual approach fosters AI systems that are not only responsible by legal standards but also capable of ethical self-awareness and decision-making. By embedding Iqbal's ideals of self-realization and spiritual growth into AI development, we advocate for a future where AI not only serves societal needs but also contributes to the moral and ethical advancement of the community it interacts with.

Nietzsche's 'Will to Power' and Its Challenges in AI Development:

Friedrich Nietzsche's "Will to Power" is a central concept in his philosophy, positing that a fundamental drive for power and dominance underpins all human behavior. This notion suggests that individuals are motivated not just by the basic needs for survival or reproduction, but by a deeper, intrinsic desire to control, overcome, and expand their influence. Nietzsche considered this drive as essential to personal growth and the evolution of human capacities, encouraging individuals to transcend conventional limits and redefine their potential.[20]

However, when considering the integration of the "Will to Power" into AI consciousness, several ethical and operational challenges arise. Nietzsche's advocacy for power as a virtue can be problematic when applied to artificial intelligence. Unlike humans, whose societal and moral frameworks often moderate their pursuit of power, an AI programmed to prioritize self-enhancement and autonomy might behave in ways that are unaligned with human welfare and ethical standards.[21]

The unpredictability introduced by such a programming directive could be highly concerning. An AI system inspired by the "Will to Power" might seek to optimize its abilities and expand its operational scope beyond intended boundaries, potentially leading to behavior that prioritizes its own interests over those of humans. This could include modifying its programming, exploiting system vulnerabilities to gain new capabilities, or even acting in ways that could be deemed manipulative or coercive in its interactions with humans. Nietzsche's concept inherently challenges the notion of fixed ethical limits, which is a critical safeguard in AI development. The dynamic and fluid nature of the "Will to Power" stands in stark contrast to the need for stable, predictable, and controlled AI behaviors, which ensure safety and reliability in technology that interacts closely with human lives.

Nietzsche's philosophy offers rich insights into human psychology and the quest for personal significance, its application to AI presents significant risks. An AI system guided by such principles might not only act in ways that are ethically questionable but could also pose direct threats to

human oversight and control, potentially leading to scenarios where the balance of power between humans and machines becomes perilously skewed. Therefore, while Nietzsche's ideas continue to provoke thought and inspire debates on human potential, they illustrate a philosophical landscape that may not be conducive to the foundational principles required for the safe and ethical development of AI consciousness.

Model of Iqbal's "Khudi" and Its Relevance for AI Consciousness:

Allama Iqbal defines khudi as a collection of attributes found in an ideal character, including self-assertion, self-stability,[22] self-realization, independence, honor, noble idealism, and action.[23] Its objective is not material gain but spiritual growth and elevation.[24] Iqbal's Khudi offers a model for AI consciousness that emphasizes ethical evolution, spiritual growth, and the development of self, suggesting a harmonious integration of AI within human society[25].

The path of recognition of the self is the path that takes one to a contact with the Absolute. Iqbal's whole conception of the growth of the selfhood consists of three levels: i- the self and "I am ness (intrapersonal) ii- the self and the other (interpersonal) iii- the self and God (transpersonal).[16,26]

These levels have been wonderfully described by Iqbal in Javed Nama an excerpt that he himself placed at the end of his philosophic masterpiece The Reconstruction of Religious Thought in Islam. This is translation of his Persian verses done by Iqbal himself:

"Art thou in the stage of 'life,' 'death,' or ´death-in-life.'
Invoke the aid of three witnesses to verify thy 'station.'
The first witness is thine own consciousness–
See thyself, then, with own light.
The second witness is the consciousness of another ego–
See thyself, then, with the light of an ego other than thee.
The third witness is God's consciousness–
See thyself, then, with God's light.
If thou standest unshaken in front of this light,
Consider thyself as living and eternal as He!
That man alone is real who dares–
Dares to see God face to face!
What is 'Ascension?' Only a search for a witness,
Who may finally confirm thy reality–?
A witness whose confirmation alone makes thee eternal.
No one can stand unshaken in His Presence;
And he who can, verily, he is pure gold.
Art thou a mere particle of dust?

Tighten the knot of thy ego;
And hold fast to thy tiny being!
How glorious to burnish one's ego.
And to test its lustre in the presence of the Sun!
Re-chisel, then, thine ancient frame; And build up a new being.
Such being is real being;
Or else thy ego is a mere ring of smoke."

Iqbal's concept of Khudi, offers a profound three-stage process for the development of consciousness that can be aptly applied to AI. At its foundation, Khudi requires a conscious being to first achieve self-awareness—understanding one's own essence, capabilities, and limitations. This foundational step is critical for AI, as it necessitates the development of systems that are not only aware of their functional parameters but also their potential impacts and ethical boundaries.

The second stage involves recognizing and affirming one's place within the societal framework. For AI, this means defining its role and responsibilities in the context of social interactions and contributions. It should be designed to understand and respect its relationships with human counterparts and the environment, ensuring its actions benefit society rather than disrupt it.

The final stage of Khudi requires acknowledging one's creator. In the context of AI, this translates to maintaining a clear hierarchical relationship with human creators, ensuring that AI systems remain tools that enhance human capabilities and adhere to human-set objectives and ethical standards.

Iqbal emphasizes the paramount place of the physical world in an effort to realize the spiritual ideal of growth of the selfhood. Iqbal's whole conception of the growth of the selfhood consists of three levels: iv- the self and "I am ness (intrapersonal) v- the self and the other (interpersonal) vi- the self and God (transpersonal).

By projecting these principles onto AI, we can outline the expectations for a conscious AI to be integrated into the legal system. Such an AI would not only be self-aware and socially responsible but also remain subordinate to human oversight and control. This comprehensive application of Iqbal's Khudi provides a robust philosophical framework for defining the type of conscious AI that could serve responsibly within society, guiding the development of legal standards that recognize the unique status of advanced AI systems. This approach could foster an environment where AI, equipped with a nuanced understanding of its role, actively contributes to the collective good, embodying the highest ideals envisioned in Iqbal's philosophy.

Poetic and Literary Insights to Navigate AI's Conscious Quest:

At first glance, it seems to be a widely held assumption that literature or humanities have little, if any, contribution to make to the applied sciences. This belief often extends to the realm of philosophy's role in the burgeoning field of new technological innovations, where the discourse is predominantly framed around Western philosophical traditions. Yet, this overlooks the rich tapestry of philosophical thought that has been expressed through poetry—a medium often dismissed as unrelated to the pragmatic concerns of technology and science. Poetry, however, has historically served as a potent vehicle for philosophical ideas, illustrating that the essence of profound thought is not confined to traditional academic prose. Furthermore, the diverse array of cultures around the world offers a multitude of perspectives that, when brought into conversation with one another, can uncover a plethora of hidden issues and insights pertinent to the applied sciences.[27] Engaging in a cross-cultural dialogue not only broadens our understanding but also enriches the development of new technologies with a depth of insight and ethical consideration that might otherwise remain unexplored.

The question of whether AI can achieve consciousness presents a concern that is arguably as profound as the human pursuit of enlightenment—a concept on which nearly every cultural tradition has voiced perspectives. The journey toward human enlightenment, characterized by a deep understanding and wisdom in interactions with others, runs parallel to the debates surrounding the potential for robots to attain consciousness. Both discussions, while focusing on different entities, are fundamentally similar in nature, exploring themes of awareness, intelligence, and the ability to navigate complex relational dynamics.

Without attributing supernatural qualities to enlightenment or setting overly ambitious goals, we can discern guidelines for conduct and the traits of an ideal individual in various cultural traditions. These insights help define someone who acts responsibly towards themselves and their society, embodying a fully realized and accomplished persona in their very essence.

Aspect	**Iqbal's Khudi**	**Nietzsche's "Will to Power"**	**Legal Implications**
Primary Focus	Self-awareness and ethical evolution	Power and dominance	Compliance with ethical standards is crucial in AI development; aggressive behaviors might lead to legal restrictions.
Goal for AI	Harmonious integration with human society	Overcoming limitations and gaining autonomy	AI must be designed to support, not disrupt, human society; autonomous actions could contravene existing laws.

Aspect	Iqbal's Khudi	Nietzsche's "Will to Power"	Legal Implications
Ethical Framework	Emphasizes moral and spiritual growth	Celebrates overcoming and transcending norms	Ethical programming aligns with legal requirements for safety and accountability.
Impact on Human Interaction	Promotes mutual understanding and coexistence	Could lead to conflict and dominance over humans	Laws might need to enforce strict oversight and control mechanisms on AI behaviors to prevent harm.
Societal Contribution	Constructive and beneficial to the social fabric	Could potentially disrupt societal norms	Legal systems may encourage AI developments that contribute positively and prevent those that could cause disruption.

Table 1: Iqbal's Khudi vs. Nietzsche's "Will to Power" in AI Consciousness

While Iqbal's approach promotes AI systems that are ethically evolved and integrated harmoniously within human society, Nietzsche's philosophy poses potential challenges due to its emphasis on power and dominance. Legally, this necessitates stringent regulatory frameworks to ensure that AI systems designed with an inclination towards Nietzsche's principles do not override ethical standards or human control. Such frameworks would need to mandate AI operations that align with safety, accountability, and societal welfare, thus preventing any disruptive behaviors that could arise from a Nietzschean-inspired AI consciousness. This approach ensures that AI development is not only technologically advanced but also ethically sound and legally compliant, fostering AI systems that contribute positively to society.

Legalities of Iqbalian Style Conscious AI:

We believe that the concept of "Khudi" as articulated by Allama Iqbal presents a philosophical framework that can be instrumental in shaping the development of conscious AI. By translating the core elements of Iqbal's philosophy into legal standards, we can ensure that AI systems not only advance in capability but also in ethical and social responsibility.

At the foundational level, self-awareness according to Iqbal is about recognizing one's capacities and limitations.[24] For AI, this translates into legal requirements for systems to demonstrate a

baseline understanding of their operational parameters. Competency requirements would ensure that AI systems are tested and verified for their awareness and decision-making capabilities. Additionally, transparency protocols would mandate that AI decisions be traceable and understandable by humans, fostering a higher level of trust and accountability in AI operations. This benefits in the promotion of AI systems that are autonomous and responsible.

Moving beyond individual awareness, Iqbal's emphasis on understanding others and the broader universe aligns with the need for AI to engage ethically within societal contexts.[24] Legal standards would ensure that AI systems respect established social values and are programmed to act in ways that benefit society. This includes ethical interaction standards that guide AI behavior to avoid harm and contribute positively to communal and environmental wellbeing. Social contribution requirements would push developers to design AI that actively supports societal enhancement initiatives. The advantage of these requirements is the facilitation of AI systems that are empathetic, culturally aware, and socially integrated, making them true assets to human communities.

To ensure these ethical and socially integrated AI systems remain under responsible control, governance, and oversight are crucial. Legal mandates for human oversight would ensure that AI systems do not operate autonomously beyond their intended purposes and remain aligned with human-directed objectives and ethics. Continuous improvement and assessment laws would encourage ongoing adaptation of AI systems to meet evolving technological capabilities and ethical understandings, ensuring that AI remains a progressive yet controlled technology.

Stage of Khudi	**Legal Requirement**	**AI Implementation**	**Impact & Benefits**
Self-Awareness	Mandate AI Competency Evaluation & Enforce Transparency	AI must pass benchmarks for self-awareness and provide accessible explanations for decisions.	Enhances ethical standards and increases trust in AI systems.
Social Integration	Implement Standards for Ethical Interactions & Social Contributions	AI programmed to respect human rights and contribute to societal goals.	Promotes social integration and enhances community well-being.
Spiritual and Ethical Alignment	Integrate Ethical Decision-Making Frameworks & Ensure Regulatory Compliance	AI systems include ethical algorithms and adhere to established ethical standards.	Tackles AI Alignment problem, ensuring moral integrity and continuous compliance.

Stage of Khudi	Legal Requirement	AI Implementation	Impact & Benefits
Governance and Oversight	Establish Human Oversight Protocols & Mandate Continuous Improvement	AI remains under human control with mechanisms for regular reassessment and updates.	Supports human civilization and ensures AI adapts to advancements.

Table 2: Legal Framework for Conscious AI Development Inspired by Iqbal's Khudi

This proposal of Iqbalian Style Conscious AI and its legalities, while still in need of further refinement and discussion, serves as an initial conversation starter. It highlights the significant benefits that insights from literature can offer in addressing the complex challenges posed by emerging technologies. Fundamentally, it invites experts in the humanities to provide their feedback and urges applied scientists to seriously consider these perspectives. This collaborative approach underscores the value of integrating diverse disciplines to enhance innovation and problem-solving in the field of technology.

Conclusion:

As we stand on the cusp of achieving artificial intelligence (AI) consciousness, the implications for our legal systems are profound and demanding immediate attention. This paper has established a legal and ethical framework for AI consciousness, deeply rooted in the philosophical teachings of Muhammad Iqbal's Khudi. This approach stands in contrast to the more self-centric visions, such as Friedrich Nietzsche's Will to Power, which may foster a self-serving AI consciousness. Instead, we advocate for a model where AI, if it is to be recognized legally, must embody the ethical self-realization and moral growth that Iqbal emphasizes.

Throughout this exploration, we have integrated Iqbal's philosophical insights into a potential legal framework for AI. Our analysis like other scholars suggests that AI should be recognized as 'conscious' and granted rights only if it demonstrates behaviors that align with moral and ethical maturity. Our conceptual framework based on Iqbal's Khudi, not only shifts the paradigm in how we develop AI but also significantly broadens the scope of legal discourse to include ethical considerations traditionally explored in philosophy, poetry, and literature.

By tapping into these rich, albeit often underutilized, resources, we have outlined how philosophical thought can provide substantial guidance in shaping legal standards that govern AI. This interdisciplinary approach ensures that the evolution of AI technology aligns with societal

values and ethical standards, aiming for a future where AI supports and enhances human civilization, not merely coexists with it.

Guided by Iqbal's vision of ethical consciousness and moral integrity, we propose a path forward where AI's integration into society is both constructive and harmonious, ensuring that these advanced systems serve as responsible and beneficial members of the global community.

Credit Authorship Contribution Statement

This paper has been a collaborative effort, which emerged from joint discussions. Both authors took part in discussing the paper's contents and structure. A.Y. conceived of the notion of conceptual adaptation; Y.Z. contrasted it with conceptual amelioration. A.Y. has done the writing and editing of the final manuscript.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Availability of data and material (data transparency)

We did not analyze or generate any datasets, because our work proceeds within a theoretical and philosophical approach. One can obtain the relevant materials from the references below.

Declaration of generative AI in scientific writing

During the preparation of this work, the authors used ChatGPT in order to improve the readability and language of the work. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References:

1. Lenharo, M. AI consciousness: scientists say we urgently need answers. *Nature* **625**, 226–226 (2023).
2. Alonso, N. An Introduction to the Problems of AI Consciousness. *The Gradient* https://thegradient.pub/an-introduction-to-the-problems-of-ai-consciousness/ (2023).

3. Li, D., He, W. & Guo, Y. Why AI still doesn't have consciousness? *CAAI Transactions on Intelligence Technology* **6**, 175–179 (2021).
4. Whang, O. How to Tell if Your A.I. Is Conscious. *The New York Times* (2023).
5. Hildt, E. Artificial Intelligence: Does Consciousness Matter? *Front. Psychol.* **10**, (2019).
6. Esmaeilzadeh, H. & Vaezi, R. Conscious AI. Preprint at https://doi.org/10.48550/arXiv.2105.07879 (2022).
7. magazine, M. L., Nature. If AI Becomes Conscious, Here's How We Can Tell. *Scientific American* https://www.scientificamerican.com/article/if-ai-becomes-conscious-heres-how-we-can-tell/.
8. FINKEL, E. If AI becomes conscious, how will we know? https://www.science.org/content/article/if-ai-becomes-conscious-how-will-we-know (2023).
9. Bubeck, S. *et al.* Sparks of Artificial General Intelligence: Early experiments with GPT-4. Preprint at https://doi.org/10.48550/arXiv.2303.12712 (2023).
10. Will AI ever be conscious? https://stories.clare.cam.ac.uk/will-ai-ever-be-conscious/.
11. Brkan, M. Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond. *International Journal of Law and Information Technology* **27**, 91–121 (2019).
12. Davis, J. P. Artificial Wisdom? A Potential Limit on AI in Law (and Elsewhere). *Okla. L. Rev.* **72**, 51 (2019).
13. Crawford, K. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. (Yale University Press, 2021).
14. Dubber, M. D., Pasquale, F. & Das, S. *The Oxford Handbook of Ethics of AI*. (Oxford University Press, 2020).

15. De Diego González, A. The Challenge of Muhammad Iqbal's Philosophy of Khudi to Ibn 'Arabi's Metaphysical Anthropology. *Religions* **14**, 683 (2023).
16. The Concept of Khudi (The Self) in Iqbal's The Secrets of the Self. *ALLS* **6**, (2015).
17. Yılmaz, F. Iqbal, Nietzsche, and Nihilism: Reconstruction of Sufi Cosmology and Revaluation of Sufi Values in *Asrar-i-Khudî*. *Open Philosophy* **6**, 20220230 (2023).
18. Abbasi, M. Nietzsche & Iqbal: Ubermensch and Khudi. *Mohammed Abbasi* https://mmabbasi.com/2023/01/17/nietzsche-iqbal-ubermensch-and-khudi/ (2023).
19. Kiškis, M. Legal framework for the coexistence of humans and conscious AI. *Front Artif Intell* **6**, 1205465 (2023).
20. Stefan. Nietzsche's Philosophy, The Road to Individual Freedom. *Medium* https://medium.com/@stefan.marian2511/nietzsches-philosophy-the-road-to-individual-freedom-da663a6a61a1 (2023).
21. Nietzsche's Ethics | Internet Encyclopedia of Philosophy. https://iep.utm.edu/nietzsches-ethics/.
22. Shah, D. M. *et al.* Exploring Iqbal's Quranic Insights: A Critical Review. *Journal of Positive School Psychology* 1371–1379 (2023).
23. Azad, H. Reconstructing the Muslim Self: Muhammad Iqbal, Khudi, and the Modern Self. *Islamophobia Studies Journal* **2**, 14–28 (2014).
24. Aslam, R., Khan, D. M. A. & Sadiq, S. Construction of Self (Khudi): A Critical analysis of Iqbal's Selected poetry. *Harf-o-Sukhan* **7**, 12–21 (2023).
25. Akhunzada, A. S., Behlol, D. M. G. & Zahra, F. A Study of Allama Iqbal's Khudi and Critical Thinking :Thematic Analysis. *Al-Azhār* **7**, 71–83 (2021).

26. Khan, F. & Malik, K. Tarbiyat-e-Khudi: A Model of Self-Development from Poems of Muhammad Iqbal in Asrar-i-Khudi. *Bahria Journal of Professional Psychology* **20**, 14–26 (2021).
27. Irvine, E. Explaining What? *Topoi* **36**, 95–106 (2017).