SIR Publishers
CENTER FOR INTERNATIONAL SCIENTIFIC PUBLICATIONS
JOURNAL OF MULTIDISCIPLINARY BULLETIN
ISSN(Online): 2984-6722
SJIF Impact Factor |(2024): 6.752|
Volume-9, Issue-4, Published |20-11-2026|

AYBSHINOSI AS A PHILOSOPHY OF MORAL SELF-KNOWLEDGE IN THE AGE OF ARTIFICIAL INTELLIGENCE

Komilov D.R

Abstract: *This article offers a philosophical and hermeneutic conceptualization of aybshinosi as a normatively oriented form of moral self-knowledge in which a person's recognition of his or her own morally significant imperfection does not end with the mere identification of a fault or a negative self-assessment, but is progressively transformed into a contextual interpretation of the act, conscientious self-evaluation, the determination of differentiated responsibility, the practice of correction, and the subsequent moral development of the person. The historical-philosophical reconstruction draws on the works of Jalal al-Din Muhammad Balkhi, Farid al-Din Attar, Saadi Shirazi, and Abu Hamid al-Ghazali, whose writings are considered not as direct anticipations of contemporary digital ethics but as sources of anthropological models concerning the human attitude toward one's own and others' imperfections. The central thesis of the study is that algorithmically mediated evaluation may increase the accuracy of detecting particular facts, reveal patterns, and ensure a more consistent application of formalized criteria; however, the moral meaning of a detected fault cannot be automatically derived from technical classification, since it requires the interpretation of intention and context, the determination of the appropriate degree of responsibility, respect for human dignity, and recognition of the individual's capacity for change over time. A six-stage normative-analytical model of aybshinosi is proposed: recognition, contextual interpretation, conscientious self-assessment, differentiated responsibility, correction, and moral development. It is argued that meaningful human oversight is not equivalent to the formal presence of an operator, while distributed responsibility must not degenerate into its institutional diffusion. The article substantiates the normative preference for human-centered sociotechnical integration, in which the analytical capacities of artificial intelligence are combined with contestability, contextual interpretation, moral repair, and addressable human responsibility.*

Keywords: *aybshinosi, moral self-knowledge, philosophical anthropology, Persian-Tajik ethics, conscience, human dignity, responsibility, moral repair, artificial intelligence, algorithmically mediated evaluation, human oversight.*

Author Information: *Daler R. Komilov – Candidate of Philosophical Sciences, Associate Professor, Head of the Department of History and Intercultural Relations, State Educational Institution "Tajik International University of Foreign Languages named after Sotim Ulughzoda". 17/6 Muhammadiyev Street, Dushanbe 734019, Republic of Tajikistan. E-mail: daler.komilzoda@gmail.com*

INTRODUCTION

The problem of human imperfection belongs to the fundamental questions of philosophical anthropology, since the human being is not only a being capable of error in

JOURNAL OF MULTIDISCIPLINARY BULLETIN
ISSN(Online): 2984-6722
SJIF Impact Factor |(2024): 6.752|
Volume-9, Issue-4, Published |20-11-2026|

cognition, moral evaluation, and practical action, but also a subject capable of returning to an already committed act, subjecting it to critical interpretation, determining his or her own degree of responsibility, and transforming an experienced error from a mere biographical fact into a foundation for subsequent moral transformation. Consequently, a fault may become not only an object of external condemnation but also a starting point for internal moral work.

Within the Persian-Tajik philosophical and ethical tradition, this problematic is concentrated around the category of ayb, whose semantic field encompasses deficiency, flaw, wrongdoing, and a morally blameworthy quality, whereas aybshinosi shifts the research emphasis from the mere existence of imperfection to the process through which the moral subject recognizes and interprets it.

In the present study, aybshinosi is understood as a normatively oriented process through which a person recognizes his or her own morally significant imperfection, interprets it contextually, subjects it to conscientious self-assessment, accepts an appropriately justified degree of responsibility, and transforms the resulting knowledge into a practice of correction and moral development.

The central thesis of the article is formulated as follows: algorithmically mediated detection of human error may expand knowledge about behavior and increase the consistency of certain formalized evaluations; however, the moral meaning of the detected fault emerges not from technical classification itself, but from the human process of interpretation, attribution of responsibility, respect for personal dignity, and recognition of the possibility of moral change.

In an earlier study, the author examined artificial intelligence primarily through the categories of the meaning of human existence, self-consciousness, freedom, and general moral responsibility [Komilov 2025, 138–147]. The present study substantially specifies the subject: rather than analyzing the opposition between the human being and technology as such, it focuses on the structure of aybshinosi, the transition from knowledge of one’s own imperfection to correction, and the conditions for the normatively justified use of algorithmically mediated evaluation.

Accordingly, the present argument deliberately excludes previously used general claims concerning the allegedly unconditional superiority of human inner experience over technical information processing. Instead, it addresses a more precise question: how can informationally accurate detection of a fact be incorporated into a morally legitimate evaluative procedure without eliminating responsibility, context, and the possibility of personal transformation?

1. Methodology and the Boundaries of the Source Corpus

The methodology of the study combines philosophical-anthropological, hermeneutic, comparative, and scenario-futurological approaches. The unit of historical-philosophical analysis is not the mechanical occurrence of the word ayb in a text, but the functional relation between human imperfection and self-knowledge, moral evaluation, relations with others, responsibility, and the possibility of correction.

The study does not directly compare classical Persian-Tajik texts and artificial intelligence as historically commensurable phenomena, which would result in anachronism.

Instead, it compares two different regimes of evaluation: a conscientious-interpretive regime, which presupposes the person's inner relationship to his or her own act, and an algorithmically mediated regime, in which the detection, classification, and presentation of significant information are partly delegated to a technical system embedded within a particular institution.

The source corpus is deliberately limited to the works of Balkhi, Attar, Saadi, and al-Ghazali. Balkhi represents the return of the critical gaze from the other person to the evaluating subject himself; Attar represents the refusal to reduce the person to a single external form; Saadi represents the moral limits governing the treatment of information about another person's faults; and al-Ghazali represents the transition from the recognition of one's own ayb to internal correction.

Such a limitation does not imply an underestimation of the legacy of Firdawsi, Abd al-Rahman Jami, or other thinkers. Their systematic inclusion constitutes a prospect for a separate historical-philosophical study.

2. Classical Foundations of Aybshinosi

Jalal al-Din Muhammad Balkhi reveals the moral and epistemic asymmetry of a person who is capable of carefully detecting the imperfections of others while failing to apply an analogous measure of criticism to himself:

Ҳирс нобиност, бинад мӯ ба мӯ,
Айби халқону бигӯяд кӯ ба кӯ.
Айби худ як зарра чашми кӯри ӯ
Менабинад, гарчи ҳаст ӯ айбҷӯ [Balkhi 2013, 134].

Author's translation: "Blinded desire discerns the faults of people hair by hair and spreads them everywhere, yet its blind eye fails to notice even the slightest flaw of its own."

What is philosophically significant here is not an absolute prohibition against criticizing another person, since critical judgment is necessary where harm is caused and justice is violated, but rather the identification of an asymmetry in moral cognition: the other becomes the object of rigorous analysis, while the position of the evaluator himself is placed beyond analogous scrutiny.

Farid al-Din Attar formulates another aspect of the problem:

Чанд гардӣ гирди сурат айбҷӯй,
Ҳусн дар ғайб аст, он дар ғайб ҷӯй [Attar 2013, 134].

Author's translation: "How long will you circle around outward form, searching for faults? Beauty lies in the hidden depth; seek it there."

This statement does not imply moral relativism and does not eliminate the objective significance of harm. Its philosophical function consists in refusing to transform a single act, an external characteristic, or one episode from the past into an exhaustive definition of the entire person.

Saadi Shirazi transfers the problem into the sphere of moral communication:

"Мардумонро айби ниҳонӣ пайдо макун, ки мар эшонро расво кунию худро беэътимод" [Saadi 2008, 194].

Author's translation: "Do not disclose people's hidden faults, for you will disgrace them and make yourself unworthy of trust."

In this case, moral evaluation concerns not only the truthfulness of information about another person's fault but also the purpose, form, and consequences of its dissemination. Therefore, aybpushi—the concealment of another person's private fault—cannot be extended to crime, violence, corruption, or systematic violations of rights, since preserving the dignity of one person cannot justify the continuation of substantial harm to another.

Particularly important for the proposed model is the position of Abu Hamid al-Ghazali:

"...талаби айби худ ҳамекунад, ки чун иллат надонад, илоҷ натавонад кард" [Ghazali 1999, 414].

Author's translation: "One must seek out one's own fault, for without knowing the cause of the illness, one cannot undertake its treatment."

Here, knowledge of one's own fault acquires a practical purpose: it is needed not for self-humiliation, but for identifying the cause of wrongful behavior and for its subsequent correction.

3. Aybshinosi, Self-Criticism, Guilt, Shame, and Conscience

Aybshinosi must be distinguished from a number of closely related but non-identical phenomena. Self-criticism denotes the capacity to evaluate one's own act or quality negatively, yet it may remain merely a rhetorical procedure that produces no real change in behavior. Guilt is primarily related to the recognition of one's involvement in a wrongful act and, in its constructive form, may motivate restitution and correction. Shame may affect not only a particular act but the whole image of the self, transforming the judgment "I acted wrongly" into the destructive generalization "I am completely unworthy."

Conscience constitutes a broader internal normative horizon within which the subject relates his or her intentions and actions to what is recognized as morally required. Contemporary philosophy likewise emphasizes the complexity of conscience and the impossibility of reducing it exclusively to a single rational or emotional component [Coady 2023, 2497–2516].

Thus, self-criticism creates the possibility of negative self-evaluation; guilt and shame express different forms of morally significant experience; conscience establishes the normative horizon; whereas aybshinosi is a dynamic process through which a recognized imperfection passes through interpretation and responsibility toward practical correction and moral development.

4. Human Dignity and the Possibility of Change

In this article, human dignity is understood as the fundamental normative status of the person that prevents his or her exhaustive reduction to a single wrongdoing, rating, statistical characteristic, or digital profile, even where a particular act deserves serious moral or legal condemnation [Debes 2026].

Such an understanding does not eliminate responsibility. On the contrary, responsibility is meaningful precisely because the human being is treated as a subject capable of understanding a normative demand, answering for his or her action, and changing subsequent behavior.

Therefore, the judgment "this person committed a morally unacceptable act" is fundamentally different from the judgment "this person is completely and permanently identical with his or her worst act." The former concerns the evaluation of an action, whereas the latter turns a single biographical fact into a final definition of the person and thereby eliminates the meaning of moral development.

5. The Six-Stage Normative-Analytical Model of Aybshinosi

The proposed model is not an empirically established universal psychological sequence that every person necessarily undergoes. Rather, it is a normative-analytical construction that identifies the essential semantic components of mature moral self-knowledge.

Stage	Content	Moral Function	Principal Risk
Recognition	Acknowledgement of morally significant imperfection	Overcoming denial	Self-justification
Contextual interpretation	Analysis of intention, knowledge, freedom, circumstances, and consequences	Determining moral meaning	Formalism or relativization of harm
Conscientious self-assessment	Relating the act to what is recognized as morally required	Internal appropriation of the problem	Indifference or self-humiliation
Differentiated responsibility	Determining responsibility according to control, knowledge, and the ability to prevent harm	Fair attribution of responsibility	Shifting responsibility or hypertrophy of guilt
Correction	Changing behavior and, where possible, eliminating or compensating for consequences	Transition from recognition to action	Repentance without practical change
Moral development	Integrating the experience of error into lasting personal change	Preventing repetition	Lifelong stigmatization by the past

Source: developed by the author.

In condensed form, the model is expressed as follows:

recognition → interpretation → conscientious self-assessment → differentiated responsibility → correction → moral development.

At the same time, moral development cannot be understood as a fully isolated "self-production" of the person. Research on moral self-cultivation demonstrates the importance of interpersonal feedback, example, and the social environment [Slote 2020, 27–48].

The stage of correction may also include moral repair – the recognition of harm caused, restitution, apology, or an attempt to restore a damaged relationship [Norlock 2022, 929–946]. However, the wrongdoer's obligation to undertake corrective action does not create a symmetrical obligation on the part of the victim to forgive.

6. Algorithmically Mediated Evaluation as a Sociotechnical Process

It is methodologically incorrect to describe algorithmically mediated evaluation as an autonomous confrontation between "the human being" and "the algorithm," since an actual decision emerges within a sociotechnical structure:

human being – institution – technical system.

JOURNAL OF MULTIDISCIPLINARY BULLETIN
ISSN(Online): 2984-6722
SJIF Impact Factor |(2024): 6.752|
Volume-9, Issue-4, Published |20-11-2026|

The technical system processes data and produces an output; the institution defines the purpose, criteria, thresholds, user authority, and real-world consequences; concrete individuals design the system, construct the data environment, implement it, interpret its output, and make or execute decisions.

Therefore, the fairness of algorithmically mediated evaluation depends not only on mathematical accuracy but also on data quality, the normative justification of selected criteria, the organization of institutional procedures, and the distribution of authority [Mittelstadt et al. 2016; Selbst et al. 2019].

The strongest counterargument against a human-centered position is that, in certain formalized tasks, an algorithmic procedure may indeed be more consistent than an individual human decision-maker and less dependent on fatigue, personal sympathy, or situational impulsiveness. This argument should be accepted: the philosophy of aybshinosi must not rest on a romantic idealization of the human evaluator.

However, statistical accuracy does not automatically generate the normative legitimacy of a final decision, since before any calculation is performed it is necessary to determine which features are relevant, which differences are permissible, what level of error is acceptable, and what consequences may justifiably follow from a classification [Coeckelbergh 2020; Dignum 2019].

Thus, human participation is required not because human intuition always epistemically surpasses computation, but because morally and socially significant decisions must have an addressable bearer of responsibility, a possibility of justification, and a procedure for review.

7. Human Oversight, Automation Bias, and Distributed Responsibility

Meaningful human oversight is not equivalent to the formal presence of an operator within a technological procedure. Article 14 of Regulation (EU) 2024/1689 requires high-risk artificial intelligence systems to be capable of effective oversight by natural persons [European Union 2024, art. 14]. UNESCO likewise proceeds from the principle that the use of artificial intelligence should not displace ultimate human responsibility and accountability [UNESCO 2022].

However, the formal presence of human oversight does not in itself guarantee its practical effectiveness. Automation bias denotes the user's tendency to rely excessively on an automated recommendation. Contemporary reviews of human–AI interaction demonstrate the persistence of this risk and its dependence on the nature of the task, user competence, and the manner in which machine-generated outputs are presented [Romeo, Conti 2026, 259–278].

Consequently, meaningful human oversight presupposes the ability to understand the significance of the output, obtain essential information concerning its basis, take account of context absent from the digital profile, challenge the classification, reject the recommendation, and initiate review.

It is equally important to distinguish distributed from diffused responsibility. To determine individual and institutional responsibility, the following four factors are proposed:

actual control + accessible knowledge + institutional authority + real capacity to prevent or reduce harm.

This framework allows one to avoid both the excuse that "the system decided" and the unjust transformation of the final operator into the sole bearer of blame for a decision architecture over which he or she had no effective control [Elish 2019, 40–60].

8. Discussion and the Independence of the Proposed Concept

The proposed model does not replace existing artificial intelligence ethics; rather, it adds a philosophical-anthropological dimension to it.

It shares with Paul Ricoeur's hermeneutics an understanding of the person as a temporally unfolding structure of self-interpretation, action, and responsibility [Ricoeur 1992]. In the present study, however, this perspective is applied to the narrower problem of transforming an acknowledged moral error into correction and subsequent change.

The concept also shares with Floridi and colleagues an emphasis on autonomy, justice, and explainability [Floridi et al. 2018], yet aybshinosi poses an additional question: what happens to the moral subject after information concerning his or her fault has already been established?

With the approaches of Dignum and Coeckelbergh, the present position shares attention to responsible design, application, and institutional organization of artificial intelligence [Dignum 2019; Coeckelbergh 2020], but it shifts the focus from technological governance to the structure of the human being's moral relation to self.

The author's previous publication addressed the broader relationship among artificial intelligence, self-consciousness, freedom, and moral responsibility [Komilov 2025, 138–147]. The present article does not reproduce that ontological and existential argument. Instead, the concepts of responsibility and dignity are incorporated into a different research construction:

aybshinosi → interpretation → differentiated responsibility → correction → moral development.

It is precisely this structure, together with its application to algorithmically mediated evaluation, that constitutes the independent result of the present study.

9. Futurological Scenarios

The first scenario – the intensification of algorithmic control – is characterized by a high degree of automation, weak possibilities of contestation, limited transparency, and a tendency to transform a person's digital past into a stable characterization of his or her present.

The second scenario – responsible technological assistanc – presupposes the use of artificial intelligence as an analytical instrument whose result can genuinely be verified, supplemented by contextual information, and revised by a competent human decision-maker.

The third scenario – human-centered sociotechnical integration – combines the analytical advantages of technical systems with a clear distribution of responsibility, procedures of appeal, correction of erroneous data, contextual interpretation, and recognition of the person's capacity to change over time.

Within the framework of the proposed concept, the third scenario is normatively preferable, because it does not reject computational efficiency while at the same time preventing the digital profile from becoming a final definition of the person.

JOURNAL OF MULTIDISCIPLINARY BULLETIN
ISSN(Online): 2984-6722
SJIF Impact Factor |(2024): 6.752|
Volume-9, Issue-4, Published |20-11-2026|

10. Main Results, Limitations, and Prospects

The study reveals that the classical Persian-Tajik ethical tradition associates a mature attitude toward human imperfection with the return of the critical gaze to the evaluating subject himself; establishes a conceptual distinction between aybshinosi, self-criticism, guilt, shame, and conscience; proposes a six-stage normative-analytical model of the transition from recognition of one's own imperfection to moral development; clarifies that correction may include not only self-correction but also moral repair; and demonstrates that algorithmically mediated evaluation is sociotechnical in nature and does not acquire full moral legitimacy solely through technical accuracy.

The study is theoretical and philosophical in character; therefore, the proposed model does not claim the status of an empirically proven universal psychological sequence. The source corpus is deliberately limited, while the futurological scenarios possess an analytical rather than probabilistic-predictive status.

Future research may expand the historical-philosophical corpus through a special examination of Firdawsi, Abd al-Rahman Jami, and other thinkers, as well as empirically investigate the impact of algorithmically mediated evaluations on human self-reflection, the sense of responsibility, and practices of moral correction.

Conclusion

Aybshinosi makes it possible to regard human imperfection not as the final essence of the person and not as material for endless external condemnation, but as a possible starting point of moral self-knowledge in which a detected error acquires transformative significance when the person is capable of acknowledging it, understanding its circumstances and causes, determining his or her own degree of responsibility, changing behavior, and, insofar as possible, repairing the consequences of the harm caused.

The classical Persian-Tajik tradition reveals different dimensions of this movement: Balkhi redirects the critical gaze from the other person to the evaluating subject himself; Attar warns against reducing the person to outward appearance; Saadi connects truth about another person's fault with responsibility for the form and purpose of its disclosure; while al-Ghazali directly links knowledge of one's own ayb with the necessity of correction.

The digital environment complicates this structure, since a technical system may indeed be accurate and consistent in detecting formalizable features; however, accuracy of detection is not identical with completeness of moral judgment.

There is a fundamental philosophical gap between the statements "the system has reliably detected a fact" and "the system has exhaustively determined the moral status of the person." This gap cannot be eliminated merely by increasing the volume of data.

The philosophical significance of aybshinosi lies precisely in the fact that moral knowledge of a fault achieves its purpose not at the moment of technical detection or public condemnation, but when truth about one's own error becomes responsibility, responsibility becomes correction, correction becomes moral repair, and lived experience becomes the foundation of a more mature human existence.

Primary Sources

Attar 2013 – Attar, Farid al-Din. Mantiq al-Tayr and Asrarnama. Dushanbe: Adib, 2013. 480 pp. (In Tajik).

Balkhi 2013 – Balkhi, Jalal al-Din Muhammad. Masnavi-yi Ma'navi: Books Five and Six. Dushanbe: Adib, 2013. 480 pp. (In Tajik).

Ghazali 1999 – Ghazali, Abu Hamid Muhammad. Kimiyoi Saodat [The Alchemy of Happiness], Vols. I–II. Tehran: Soroush, 1999. 827 pp. (In Tajik).

Saadi 2008 – Saadi Shirazi. Gulistan. Dushanbe: Maorif va Farhang, 2008. 232 pp. (In Tajik).

REFERENCES:

Coady, C.A.J. (2023). "The Significance and Complexity of Conscience." Philosophia, Vol. 51, pp. 2497–2516. DOI: 10.1007/s11406-023-00692-3.

Coeckelbergh, Mark (2020). AI Ethics. Cambridge, MA: MIT Press.

Debes, Remy (2026). "Dignity." The Stanford Encyclopedia of Philosophy. Spring 2026 Edition.

Dignum, Virginia (2019). Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Cham: Springer.

Elish, Madeleine Clare (2019). "Moral Crumple Zones: Cautionary Tales in Human–Robot Interaction." Engaging Science, Technology, and Society, Vol. 5, pp. 40–60. DOI: 10.17351/ests2019.260.

European Union (2024). "Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence." Official Journal of the European Union.

Floridi, Luciano, Cowls, Josh, Beltrametti, Monica et al. (2018). "AI4People – An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." Minds and Machines, Vol. 28, pp. 689–707. DOI: 10.1007/s11023-018-9482-5.

Komilov, Daler R. (2025). "Artificial Intelligence and the Philosophy of the Meaning of Human Existence: Moral Responsibility." Worldview Horizons of Artificial Intelligence Development, No. 5(24), pp. 138–147. DOI: 10.24412/2181-1784-2025-24-138-147. (In Russian).

Mittelstadt, Brent Daniel, Allo, Patrick, Taddeo, Mariarosaria, Wachter, Sandra, and Floridi, Luciano (2016). "The Ethics of Algorithms: Mapping the Debate." Big Data & Society, Vol. 3, No. 2. DOI: 10.1177/2053951716679679.

Norlock, Kathryn J. (2022). "Forgiveness and Moral Repair." In Manuel Vargas and John M. Doris (eds.), The Oxford Handbook of Moral Psychology. Oxford: Oxford University Press, pp. 929–946.

Ricoeur, Paul (1992). Oneself as Another. Chicago: University of Chicago Press.

Romeo, Giuseppe and Conti, Daniela (2026). "Exploring Automation Bias in Human–AI Collaboration: A Review and Implications for Explainable AI." AI & Society, Vol. 41, pp. 259–278. DOI: 10.1007/s00146-025-02422-7.

JOURNAL OF MULTIDISCIPLINARY BULLETIN
ISSN(Online): 2984-6722
SJIF Impact Factor |(2024): 6.752|
Volume-9, Issue-4, Published |20-11-2026|

Selbst, Andrew D., Boyd, Danah, Friedler, Sorelle A., Venkatasubramanian, Suresh, and Vertesi, Janet (2019). "Fairness and Abstraction in Sociotechnical Systems." Proceedings of the Conference on Fairness, Accountability, and Transparency. New York: ACM, pp. 59–68. DOI: 10.1145/3287560.3287598.

Slote, Michael (2020). "Moral Self-Cultivation East and West: A Critique." In Between Psychology and Philosophy: East-West Themes and Beyond. Cham: Palgrave Macmillan, pp. 27–48.

UNESCO (2022). Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO.

Full Author Information:

Daler Rustamovich Komilov – Candidate of Philosophical Sciences, Associate Professor, Head of the Department of History and Intercultural Relations, State Educational Institution "Tajik International University of Foreign Languages named after Sotim Ulughzoda", Dushanbe, Republic of Tajikistan.

17/6 Muhammadiyev Street, Dushanbe 734019, Republic of Tajikistan.

E-mail: daler.komilzoda@gmail.com