UDC 33 **DOI: 10.34670/AR.2026.87.14.073**

Sociotechnical Systems Theory and Human-AI Collaboration in Engineering Management: An Analytical Framework for Adaptive Governance

Samantsi Success Eyram

Master's Student, Higher School of Management,
Peoples' Friendship University of Russia,
117198, 3, Miklukho-Maklaya str., Moscow, Russian Federation;
e-mail: 1032245633@rudn.ru

Abstract

Artificial intelligence (AI) is transforming the way engineering management makes decisions, manages resources, and governs risks. Although discussions of technology often focus on automation efficiency, engineering systems remain fundamentally sociotechnical in nature, requiring alignment among human expertise, algorithmic logic, and organizational structures. The article uses the theory of Sociotechnical Systems (STS) as a way to create an analytical framework on human-AI collaboration in Engineering Management. The article explores how joint optimization principles, boundary conditions and governance mechanisms can be deconstructed to arrive at three dimensions of alignment: cognitively complementary, procedurally interoperable and institutionally legitimate. Theoretical propositions are proposed to explain the impact of misalignment in these dimensions on performance degradation and adaptive friction and ethical risk. The analysis challenges techno-deterministic premises, challenges STS theory and proposes a research agenda for actionable and empirical validation of adaptive governance models in algorithmic governance contexts. The authors conclude that sociotechnical design is essential for effective engineering management of human-AI collaboration rather than simply introducing technology.

For citation

Samantsi Success Eyram (2026) Sociotechnical Systems Theory and Human-AI Collaboration in Engineering Management: An Analytical Framework for Adaptive Governance. *Ekonomika: vchera, segodnya, zavtra* [Economics: Yesterday, Today and Tomorrow], 16 (5A), pp. 682-689. DOI: 10.34670/AR.2026.87.14.073

Keywords

Sociotechnical systems theory, human-AI collaboration, engineering management, algorithmic governance, adaptive leadership, joint optimization.

Introduction

Engineering management has historically balanced technical optimization with human organizational dynamics. For decades, computational tools have been used to supplement project scheduling, risk assessment, resource leveling and quality control. But today's AI systems are also complex, with machine learning and predictive capabilities, self-optimizing systems, and generative models adding new layers of complexity to engineering decision ecosystems. AI systems differ from traditional software because they demonstrate some concepts which have to be understood through probabilistic reasoning, adaptive learning, and opaque decision-making processes, all of which go against the traditional management paradigm with deterministic control and explicit accountability.

The prevailing discourse in engineering management about integrating AI has tended to be linear: algorithms take the place of human decisions, latency is cut, and resources are allocated more efficiently. This techno-deterministic story may ignore the sociotechnical character of engineering systems, in which tech and human actors co-create organizational outcomes. The theory of Sociotechnical Systems (STS) was developed in mid-20th century organizational studies and has been advocated that it is better to optimize both social and technical subsystems than using only technology optimization [Cherns, 1976; Trist, Bamforth, 1951].

When applied to engineering management with the aid of artificial intelligence, STS theory offers a strong analytical framework for identifying and comprehending the causes of collaboration friction, developing adaptive governance, and avoiding capability erosion. This article proposes a theoretical model which incorporates the concepts of STS and the latest algorithmic governance research to interpret the dynamics of collaboration between humans and AI in engineering management. The analysis is carried out by looking at the foundations and adaptation of STS to the context of AI, articulating dimensions of alignment, developing theoretical propositions, and finding research gaps. The framework offers a counter-narrative to the reductionist approach of automation, and suggests that adaptive governance is essential for sustainable efforts to integrate AI into engineering systems.

Theoretical Framework: Extending Sociotechnical Systems Theory to Algorithmic Contexts

The theory of classical STS arose from the coal mining research that showed that innovation in technology alone failed to produce optimal results [Trist, Bamforth, 1951]. The principles that are emphasized are joint optimization, responsible autonomy and boundary management. Joint optimization claims that technical efficiency and social satisfaction have to be designed jointly; either cannot be sacrificed without creating friction in the system. Responsible autonomy advocates decentralized decision-making in line with local knowledge and constraints of the system. The emphasis of boundary management is on the interface design between subsystems that allows for the exchange of information without causing the system to collapse.

The traditional STS assumptions are challenged by modern AI systems. Algorithms are now adapting to learn, offering probabilistic results and non-deterministic behavior, making accountability, predictability and calibration of trust more challenging. These challenges are exacerbated in engineering management contexts: explainability, certification and human override, for example in safety critical systems (such as aerospace, civil infrastructure, medical devices) are at odds with black-box AI architectures. Cyber-physical systems, algorithmic governance and human-machine teaming literature, all of which are part of contemporary STS extensions, are included [De Winter et al., 2024;

Rahwan et al., 2019]. These extensions envision AI as a quasi-agent in sociotechnical networks and necessitate a new form of governance, one that addresses epistemic uncertainty, redistributes accountability, and retains human oversight.

The theoretical synthesis results in three dimensions of alignment that are important for human-AI collaboration in engineering management:

– Cognitive Complementarity: Extent of human expertise and algorithmic processing being complementary to solve complementary problem space without redundancy and conflict.
– Procedural Interoperability: The structure and the workflow of human decision processes and integration of algorithmic output.
– Institutional Legitimacy: Organizational, regulatory and cultural recognition of AI mediated decision in engineering governance.

Any misalignment creates adaptive friction, or performance loss or endangers systemic risk. The next section analytically unpacks these dimensions, and draws out theoretical propositions.

Analytical Discussion: Alignment Dimensions, Governance Mechanisms, and Propositions

Cognitive Complementarity and Epistemic Division of Labor

Human engineers are good at cross domain synthesis, anomaly detection, ethical judgment and contextual reasoning. AI systems are excellent at pattern recognition, multi-dimensional optimization, probabilistic forecasting and speedy simulation. Specific problems have to be solved in a way that achieves cognitive complementarity, which involves an epistemic division of labor between subsystems, so that each one deals with a problem for which it has a comparative advantage. In theory, this is where human context conflicts with AI, as, for example, in scenarios where AI is used to perform tasks that do not lend themselves to algorithmic processing, such as negotiations with stakeholders, regulatory interpretation, safety trade-off evaluation, etc., or where humans are assigned to AI oversight roles but lack real power. This produces capability atrophy, decision paralysis or overreliance on opaque outputs.

A key aspect of STS theory is responsible autonomy, or in the AI context, calibrated delegation: where humans take responsibility for factors that are outside of the scope of algorithms, such as boundary conditions, value trade-offs and exception handling, while algorithms are responsible for routine optimization within a known set of parameters [5, 6]. Problem complexity and uncertainty are additional moderators of complementarity. AI support is beneficial in well-organized engineering challenges, like routing in supply chains, or finite element analysis. In multi-stakeholder, novel or ill-structured contexts, human leadership must prevail. Cognitive overload, algorithmic brittleness or decision displacement occurs when misalignment occurs.

Proposition 1: Human-AI collaboration performance is highest when the division of labor is epistemic, where AI takes over well-structured optimization problems and humans take over ill-structured value-laden decisions.

Proposition 2: The positive effect of epistemic division of labor on collaboration performance is moderated by task uncertainty, with greater benefits occurring under conditions of moderate uncertainty.

Procedural Interoperability and Workflow Integration

Procedural interoperability has to do with the way the output of the algorithms is utilized within engineering management processes. Traditional stage-gate processes, risk registers and resource

allocation models are based on linear and traceable decision-making. AI systems provide recommendations in a probabilistic way, ensembles of scenarios, and adaptive updates that challenge the traditional governance cycles.

Some challenges involve temporality (AI evolves too quickly for review cycles), format (algorithmic results do not have a standard engineering documentation format), and authority (who should have final say in cases of disagreement between human and AI). The principles of the STS boundary management suggest that interoperability needs interface standardization, decision latency calibration and explicit handoff protocol.

Theoretical considerations suggest that hybrid workflow architectures that incorporate AI outputs as decision support, rather than decision replacements, incorporate review thresholds based on confidence intervals, and retain human veto capacity for high-consequence decisions are more likely to promote interoperability. Too much rigidity, though, prevents AI systems from being adaptable; too much flexibility, on the other hand, compromises auditability and compliance.

Contingency alignment must involve workflow design matching the level of stringency of the regulations and criticality of the project. To ensure procedural interoperability, in highly regulated engineering fields, the focus needs to be on traceability, explainability, rather than algorithmic autonomy, and on human certification. Flexible and iterative integration can be a more rapid path to learning and adaptation in innovation contexts.

Proposition 3: When the procedure is interoperable, the collaboration effectiveness is positively related to the workflow design in the context of high stringencies, where traceability and human certification are emphasized, and low stringencies, where iterative changes are emphasized.

Proposition 4: When there are no escalation protocols, AI-human disagreement is negatively associated with collaboration performance, while, when structured conflict resolution mechanisms are institutionalized, it is positively associated with collaboration performance.

Institutional Legitimacy and Governance Acceptance

Institutional legitimacy refers to the elements of the organization's culture, professional norms, adherence to regulations, and stakeholder trust in AI-driven decisions. Accountability, peer review and ethical responsibility have been central to the engineering professions from their inception. AI systems violate these principles by adding a layer of opacity and dependency on data, as well as a layer of distributed accountability. Legitimacy deficits include inaction and refusal to adopt, the existence of shadow systems (humans frustrating AI), regulatory non-compliance or liability fragmentation.

In STS theory, one important aspect is that technical systems need to be consistent with social norms if they are to be integrated successfully. In engineering applications, it entails algorithm transparency, professional training, ethical guidelines, and liability structures defining the responsibilities of humans and AI. Theoretical analysis indicates the following ways to build legitimacy: participative design (when all stakeholders are involved in co-developing AI systems), explainability standards (such as SHAP, LIME, counterfactual reasoning), and governance certification (such as AI audit trails, compliance dashboards). But legitimacy depends on professional culture: Learning oriented cultures change more easily, compliance driven cultures and hierarchical cultures have legitimacy friction.

Legitimacy is further reduced in the face of the regulatory environments. Legitimacy develops more rapidly in domains governed by established AI regulations (e.g., ISO/IEC 42001, EU AI Act classification), whereas it is delayed in unregulated or fragmented domains, with unregulated domains showing parallel decision architectures, and fragmented domains showing adoption hesitancy [Shneiderman, 2020; ISO/IEC 42001, 2023; European Commission, 2021].

Proposition 5: The relation between the levels of integration of AI and the portfolio performance is positively moderated by institutional legitimacy, especially in learning-oriented cultures and regulated domains

Proposition 6: Engineering organizations that implement AI accountability frameworks (like human override policies, audit trails, and liability policies) are more likely to have sustainable collaboration than those where the policy is implicit or the deployment is technocratic.

Critical Synthesis and Theoretical Extensions

The proposed framework brings STS theory into the realm of algorithmic governance, yet it has some theoretical shortcomings that need to be recognized.

First, in classical STS, the technical and social subsystems take for granted a more or less stable environment, while the growth of AI systems forces a continuous realignment of these subsystems. Recent theoretical models should be extended with recursive adaptation loops, which involve the fact that social and technological parameters change in time depending on human-AI interactions.

Second, AI is conceptualized as a single monolithic domain, but it can be used in a variety of engineering systems, like predictive maintenance algorithms, generative design systems, autonomous scheduling agents and risk simulation models. The governance needs, epistemic profiles and integration issues for each exhibit are different. Theoretical granularity should be used to differentiate the categories of AI and match them to engineering management functions.

Third, institutional legitimacy relies on rational acceptance, but engineering acceptance is often more of an externally imposed process of competitive pressure, vendor lock-in or executive mandate. Critical theory perspectives emphasize how power differentials, ownership struggles over data, and potential job losses result from STS frameworks that fail to adequately account for these issues [Mumford, 1983; Bendersky, Hays, 2012; Cresswell, Sheikh, 2013]. The scope of the analysis can be enriched by incorporating the concepts of political economy and institutional theory.

In terms of methods, it is critical to validate these propositions by using mixed-methods research, including quantitative performance measurement, qualitative ethnography of the human-AI workflows and experimental simulation of governance procedures. Adaptation and evolution of legitimacy are not possible to be captured with cross-sectional surveys. There are more comprehensive validation pathways, such as longitudinal case studies, digital trace analytics, and agent-based modeling [Tavani, 2016].

Conclusion

Theoretically, this article has expanded the scope of the STS theory to investigate human-AI collaboration in engineering management. The framework highlights three key dimensions of alignment: cognitive, procedural, and institutional, shifting away from techno-deterministic narratives and underscoring the need for adaptive governance to ensure sustainable integration of AI. The six propositions provide empirically testable foundations to validate the approaches, whereas the critical synthesis indicates theory, method and normative gaps for scholarly attention.

In engineering management practice, this means that while implementing AI, it is not enough to retrofit organizations; instead, the AI project needs to be designed in a sociotechnical fashion. To scale an AI integration, engineering leaders need to create an epistemic map, fine-tune processes and interfaces, define escalation procedures, and institutionalize accountability mechanisms. To build theoretical rigor in this area, continual discussion and engagement with STS literature, algorithmic governance research and engineering ethics frameworks are essential [Dignum, 2019].

Recursive adaptation dynamics, AI typologies as regards to engineering functions, and political economy aspects of power, data ownership, and labor transformation, serve as avenues for future research. An AI-integration approach based on sociotechnical theory—rather than technological imperative—can help engineering leaders to navigate the complexity of a human-algorithmic collaboration with analytical clarity, ethical sensitivity, and adaptive resilience.

References

1. Bendersky, C., & Hays, N.A. (2012). Status conflict in groups. Organization Science, 23(2), 323–340.
2. Cherns, A. (1976). The principles of sociotechnical design. Human Relations, 29(8), 783–792.
3. Cresswell, K., & Sheikh, A. (2013). Organizational issues in the implementation and adoption of health information technology innovations: An interpretative review. International Journal of Medical Informatics, 82(5), e73–e86.
4. De Winter, J.C.F., Dodou, D., & Hancock, P.A. (2024). Human–robot interaction at work: A sociotechnical systems perspective. Human Factors.
5. Dignum, V. (2019). Responsible artificial intelligence: How to develop and use AI in a responsible way. Springer.
6. European Commission. (2021). Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (COM(2021) 206 final). Brussels.
7. ISO/IEC 42001:2023. (2023). Information technology – Artificial intelligence – Management system. International Organization for Standardization.
8. Lee, M.K., & See, K.A. (2004). Trust in automation: Designing for appropriate reliance. Human Factors, 46(1), 50–80.
9. Mumford, E. (1983). Designing human systems for new technology: The ETHICS method. Manchester Business School.
10. Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230–253.
11. Rahwan, I., Cebrian, M., Couzin, I., et al. (2019). Machine behaviour. Nature, 568(7753), 477–486.
12. Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction, 36(6), 495–504.
13. Tavani, H.T. (2016). Ethics and technology: Controversies, questions, and strategies for ethical computing (5th ed.). Wiley.
14. Trist, E.L., & Bamforth, K.W. (1951). Some social and psychological consequences of the longwall method of coal-getting. Human Relations, 4(1), 3–38.

Теория социотехнических систем и взаимодействие человека и ИИ в инженерном менеджменте: аналитическая структура адаптивного управления

Самантси Саксесс Эйрам

Магистрант,
Высшая школа управления,
Российский университет дружбы народов,
117198, Российская Федерация, Москва, ул. Миклухо-Маклая, 3;
e-mail: 1032245633@rudn.ru

Аннотация

Искусственный интеллект (ИИ) трансформирует процессы принятия решений, управления ресурсами и контроля рисков в сфере инженерного менеджмента. Несмотря на то, что дискуссии вокруг технологий зачастую сосредоточены на эффективности автоматизации, инженерные системы по своей сути остаются социотехническими. Они

требуют гармоничного выравнивания между человеческим опытом, алгоритмической логикой и организационными структурами. В данной статье теория социотехнических систем (СТС) используется для формирования аналитической основы взаимодействия человека и ИИ в инженерном менеджменте. В работе исследуется, как принципы совместной оптимизации, граничных условий и механизмов управления могут быть деконструированы для выделения трех измерений согласованности: когнитивной взаимодополняемости, процедурной совместимости и институциональной легитимности. Предложены теоретические положения, объясняющие влияние рассогласованности в этих измерениях на снижение производительности, адаптивное трение и этические риски. Проведенный анализ ставит под сомнение технодетерминистские предпосылки, расширяет классическую теорию СТС и предлагает программу исследований для прикладной и эмпирической валидации моделей адаптивного управления в контексте алгоритмического менеджмента. Авторы приходят к выводу, что для эффективного взаимодействия человека и ИИ в инженерном менеджменте первостепенное значение имеет социотехническое проектирование, а не простое внедрение технологий.

Для цитирования в научных исследованиях

Самантси Саксесс Эйрам. Sociotechnical Systems Theory and Human-AI Collaboration in Engineering Management: An Analytical Framework for Adaptive Governance // Экономика: вчера, сегодня, завтра. 2026. Том 16. № 5А. С. 682-689. DOI: 10.34670/AR.2026.87.14.073

Ключевые слова

Теория социотехнических систем, взаимодействие человека и ИИ, инженерный менеджмент, алгоритмическое управление, адаптивное руководство, совместная оптимизация.

Библиография

1. Bendersky C., Hays N.A. Status Conflict in Groups // Organization Science. – 2012. – Vol. 23, no. 2. – P. 323–340.
2. Cherns A. The Principles of Sociotechnical Design // Human Relations. – 1976. – Vol. 29, no. 8. – P. 783–792.
3. Cresswell K., Sheikh A. Organizational Issues in the Implementation and Adoption of Health Information Technology Innovations: An Interpretative Review // International Journal of Medical Informatics. – 2013. – Vol. 82, no. 5. – P. e73–e86.
4. De Winter J.C.F., Dodou D., Hancock P.A. Human–Robot Interaction at Work: A Sociotechnical Systems Perspective // Human Factors. – 2024.
5. Dignum V. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. – Cham: Springer, 2019. – 148 p.
6. European Commission. Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). COM(2021) 206 final. – Brussels: European Commission, 2021. – 108 p.
7. ISO/IEC 42001:2023. Information Technology – Artificial Intelligence – Management System. – Geneva: International Organization for Standardization, 2023. – 46 p.
8. Lee M.K., See K.A. Trust in Automation: Designing for Appropriate Reliance // Human Factors. – 2004. – Vol. 46, no. 1. – P. 50–80.
9. Mumford E. Designing Human Systems for New Technology: The ETHICS Method. – Manchester: Manchester Business School, 1983. – 112 p.
10. Parasuraman R., Riley V. Humans and Automation: Use, Misuse, Disuse, Abuse // Human Factors. – 1997. – Vol. 39, no. 2. – P. 230–253.
11. Rahwan I., Cebrian M., Couzin I. et al. Machine Behaviour // Nature. – 2019. – Vol. 568, no. 7753. – P. 477–486.
12. Shneiderman B. Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy // International Journal of Human–Computer Interaction. – 2020. – Vol. 36, no. 6. – P. 495–504.
13. Tavani H.T. Ethics and Technology: Controversies, Questions, and Strategies for Ethical Computing. – 5th ed. –

Hoboken: Wiley, 2016. – 448 p.
14. Trist E.L., Bamforth K.W. Some Social and Psychological Consequences of the Longwall Method of Coal-Getting // Human Relations. – 1951. – Vol. 4, no. 1. – P. 3–38.