https://doi.org/10.48417/technolang.2026.03.03
Research article

Cognitive Capabilities of the Interface in Machine Learning: A Philosophical Perspective

Anastasia A. Lisenkova and Olga D. Shipunova (✉)
Peter the Great St.Petersburg Polytechnic University, St.Petersburg, Polytechnicheskaya, 29, 195251, Russia
shipunova_od@spbstu.ru

Abstract

The actual problems of digital technology in the machine learning field are related to the expansion of autonomous intelligent systems in respect to cognitive activity. The growth in the volume of information in modern networks requires new formats for working with knowledge, based on hybrid systems and neural models of artificial intelligence (AI). Neural network training involves the formation of experience of semantic orientations in multi-valued contexts using self-correction mechanisms. In the modeling of virtual machine learning environments, the role of the interface increases. The article examines the cognitive potential of the interface as a factor shaping the boundaries and modes of functional autonomy of artificial intelligence systems in machine learning. The interface is treated not as a neutral technical channel of information transfer but as a semantic matrix that programs a sequence of cognitive operations – selection, context-binding, inference, revision, and action. Particular attention is given to the temporal dimension of the interface's cognitive potential. The authors introduce the concept of chronotope (Mikhail Bakhtin) as a model for organizing the interplay of space and time within the virtual learning environment, and analyze temporal event models grounded in metaphors, image schemas, and grammatical constructions. The communicative design of the interface is discussed through the lens of cultural archetypes (Carl Gustav Jung), which shape "trusted agent" models in human–AI dialogue. Factors determining the interface's cognitive function – structuring the space of relevant information, temporal dynamics, inference formation, optimization regimes, and actor motivation – are systematized. The authors conclude that the interface's cognitive potential establishes the conditions for the subject's semantic orientation and the digital agent's goal-directed behavior, ensuring the coupling of human and artificial intelligence in the expanded production of knowledge.

Keywords: Interface; Machine learning; Artificial intelligence; Chronotope; Semantic matrix; Digital epistemic agent; Temporal models

Citation: Lisenkova, A., & Shipunova, O. (2026). Cognitive Capabilities of the Interface in Machine Learning: A Philosophical Perspective. *Technology and Language, 7*(3), 35-47. https://doi.org/10.48417/technolang.2026.03.03

© Lisenkova, A., Shipunova, O. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

УДК 1:004.5
https://doi.org/10.48417/technolang.2026.03.03
Научная статья

Когнитивные возможности интерфейса в машинном обучении: Философский взгляд

Анастасия Алексеевна Лисенкова и Ольга Дмитриевна Шипунова (✉)
Санкт-Петербургский политехнический университет Петра Великого, ул. Политехническая, д. 29, Санкт-Петербург, 195251, Россия
shipunova_od@spbstu.ru

Аннотация
Актуальные проблемы цифровых технологий в области машинного обучения связаны с расширением автономности интеллектуальных систем в когнитивной деятельности. Темпы роста объемов информации в современных сетях требуют новых форматов работы с системами знаний, основанных на гибридных системах и нейронных моделях искусственного интеллекта (ИИ). Обучение нейронных сетей включает в себя формирование опыта семантических ориентаций в многозначных контекстах с использованием механизмов самокоррекции. В моделировании виртуальной среды машинного обучения возрастает роль интерфейса. Статья посвящена анализу когнитивного потенциала интерфейса как фактора, определяющего границы и режимы функциональной автономии систем искусственного интеллекта в процессе машинного обучения. Авторы рассматривают интерфейс не как нейтральный технический канал передачи информации, а как семантическую матрицу, программирующую последовательность когнитивных операций – отбор, контекст-связывание, вывод, пересмотр и действие. Особое внимание уделено темпоральному измерению когнитивного потенциала интерфейса: вводится понятие хронотопа (М.М. Бахтин) как модели, организующей взаимодействие пространства и времени в виртуальной среде обучения, и анализируются темпоральные модели событий на основе метафор, образ-схем и грамматических конструкций. Рассматривается коммуникативный дизайн интерфейса через призму культурных архетипов (К. Юнг), формирующих модели “доверенного агента” в диалоге человека и ИИ. Систематизация факторов детерминации когнитивной функции интерфейса (структурирование пространства информации, темпоральная динамика, формирование вывода, режимы оптимизации, мотивация действий актора) представлена в виде таблицы. Делается вывод о том, что когнитивный потенциал интерфейса задаёт условия смысловой ориентации субъекта и целенаправленного поведения цифрового агента, обеспечивая сопряжение человека и искусственного интеллекта в процессах расширенного производства знания.

Ключевые слова: Интерфейс; Машинное обучение; Искусственный интеллект; Хронотоп; Темпоральные модели; Семантическая матрица; Цифровой эпистемический агент

Для цитирования: Lisenkova, A., Shipunova, O. Cognitive Capabilities of the Interface in Machine Learning: A Philosophical Perspective // Technology and Language. 2026. № 7(3). P. 35-47. https://doi.org/10.48417/technolang.2026.03.03

Лисенкова, А.А., Шипунова О.Д. This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

INTRODUCTION

Current problems in machine learning engineering are stem from the expansion of the autonomy of intelligent systems in cognitive and professional activities. The exponential growth of information generated by modern networks encourages the development of new methods and technologies aimed at actively producing new knowledge and creating conceptual structures based on hybrid systems and AI learning models. The independent orientation of modern AI neural network models in the semantic domain of scientific and professional knowledge correlates with the activity principle, which defines a smart system's new status as a digital epistemic agent (Internet Encyclopedia of Philosophy, n. d.). The cognitive resources of such agents are created through machine learning, and they imply various scenarios for their orientation within information bases and the semantic fields of global knowledge networks. In this context, the authors highlight new aspects of expanded knowledge production engineering based on the functional architecture of AI systems.

In the development of new forms for constructing contextual structures using neural network models based on machine learning, the dependence of the interpretation of linguistic and symbolic structures on the knowledge base becomes paramount. The cognitive potential of the interface is associated with the window of opportunity for the transfer of knowledge and cognitive functions during machine learning. In this case, the translation of subject-specific knowledge – that is, reading and interpreting conceptual content – is a necessary condition for machine learning, which is associated not only with memory mechanisms but, above all, with the reproduction of the original information and the technological implementation of the cognitive process under changing conditions.

Moreover, the presence of fields of meaning and a message encoding language are a necessary condition not only for the transmission of information but also for the activation of a temporal sequence of operations to identify the initial context that provides the linkages of the knowledge elements in a learning AI system. The prospects for the development of intelligent machine learning technology are linked to the reconstruction of cognitive functions at the level of perception, transmission, and production of semantic content – as a necessary condition for an artificial agent to navigate the knowledge network.

The authors examine the role of the interface and its cognitive potential in machine learning with respect to modeling semantic horizons within network interactions that determine the ability to generate and move meanings in autonomous AI systems.

COGNITIVE POTENTIAL OF INTERFACE IN EXPANDING THE AI FUNCTIONAL AUTONOMY

In machine learning technology, models of intelligent system behavior are based on the concept of a global neural workspace, defined by the parameters of AI access to a conceptual resource (Mashour et al, 2020). Information integration in the process of extended knowledge reproduction is implemented using computational algorithms based on nonlinear dynamics and recurrent processing tools (Tononi, 2004). In training expert

systems for extracting arguments from text documents, new opportunities are associated with the modeling of contextual generalization functions, which make it possible to identify the relationship between events and reduce the effect of forgetting key attitudes by eliminating redundant information (Shuang et al, 2025). To solve the problem of ambiguity in modeling the virtual environment of machine learning, a content-oriented model based on the method of generating a chain of questions with contextual binding is proposed (Wang et al, 2026, Najafabadi et al, 2026).

Reasoning models rely on a functional architecture that includes attention layers, transformers, and learning databases, defining the contours of the GAI cognitive potential, which is distinct from rule-based computational operations, including principles of computational reasoning (Teovanović et al, 2026) and the formal theory of belief change (Alchourrón et al, 1985)].

The cognitive potential of the virtual machine learning environment interface is particularly acute in the STEM education system, where artificial intelligence generates a subject field and provides dialogue tools and real-time feedback (Liu et al., 2026; Wang, B. et al., 2025). The prospect of contextual modeling of the virtual environment for machine learning is associated with the introduction of algorithmic affordances that compensate for the effect of inflating the semantic bubble in the targeted correction of semantic horizons (Jiang, 2025) and the formation of a normative framework for trust in AI agents, including requirements for consistency with human cognitive goals (Marchal et al., 2026). When modeling a machine learning environment, issues related to the semantic boundaries of the virtual environment, as well as critical reflection on the meanings and principles on which digital entities are built, become relevant.

The challenges of expanding the smart systems' behavioral autonomy in an information network highlight the key role of the interface, which defines boundaries for modeling the information environment in machine learning. In this aspect, emphasizing the ambiguity of the interface's influence on the nature of contextual information in the global network of digital society become relevant. The material aspect of the interface, represented in its technical parameters, reinforces the autonomy of the sign system for coding information processes in interactions. For its part, the matrix of the interface defines the boundaries of the virtual environment in which the processes of constructing metaverses unfold, generating new semantic and temporal horizons for interactions involving digital agents.

The interface's cognitive potential is correlated with a contextual model of semantic space, which presupposes not only the presence of a field of potential meanings and a message encoding language, but also serves as a necessary condition for activating the temporal dynamics of the interactive environment through the discovery and transmission of meanings by actors. The specifics of machine learning technology in each case are determined by the differences in the formation of the initial dataset, which defines the semantic boundaries of the learning algorithm. Feedback blocks ensure optimization of the training data used to find a solution in accordance with the specified criterion. The semantic matrix of the interactive environment forms a new continuum of human actions within networked cyberspace. In the technology of constructing modern neural network models with a flexible attention system, the principle of topology plays a key role,

ensuring global connectivity of the semantic field of the virtual learning environment. The knowledge transfer is modeled by a system of feedback loops, which allows for filtering meaningful messages.

The time factor in modeling a digital network is linked to its internal dynamics. Within the boundaries of the temporal unfolding of the digital environment, defined by the semantics of the sign system, a new language of interaction is created. The combination of the principle of topology in organizing the semantic space of interactions based on a sign system with the principle of temporality in describing the autonomous movement of meaning constitutes the specificity of contextual modeling of the learning environment. From this perspective, the interface matrix contains possible scenarios for organizing actions, assuming a potential actor capable of understanding signs and applying algorithms in interactions.

In scenarios of contextual modeling of the semantic horizons of an interactive network, the interface matrix defines its own chronotope. This implies a certain virtual space in which the relationship between time and action is governed by specific laws (for example, the *Minecraft* or *Legend of Zelda* game interface). Algorithms trained on big data begin to interpret the world through the lens of their own models, which leads to self-reference. For example, text-generating neural networks (such as GPT) can produce seemingly meaningful responses based on statistical patterns (Floridi, 2020; Petrunin, 2025).

ENCODING POTENTIAL EVENTS IN THE INTERFACE: TEMPORAL MODELS IN MACHINE LEARNING

The temporal dynamics of the processes of analyzing and generalizing knowledge structures are realized in various machine learning modes of AI agent systems with internal world models and complex self-updating circuits. From this perspective, the cognitive potential of an interface can be correlated with the boundaries of background knowledge, captured in specific technological formats. At the same time, the intelligent architecture embedded in the interface design for using specific algorithms makes it possible to implement the optimal mode of transition from one given state to another in accordance with the established criterion for achieving the goal. In deep learning of neural networks, most action parameters are generated based on the semantic relationship of contexts using self-correction mechanisms for intermediate results. (Burkov, 2020, pp. 18-20, 41, 59).

The interface's potential for creating potential meanings and characters in a virtual environment is revealed through the chronotope of the virtual environment. It defines the point of meaning movement through the correlation of value dominants of the past and present, shapes the perception of events in a specific historical context, and allows for the interpretation of events occurring in a "new space." Since the chronotope functions as a stable narrative framework for conveying meaning, the expansion of this function amounts to the challenge of modeling the semantic space of interactions in virtual environments.

In the digital age, the interface becomes a tool that shapes a new chronotope as a spatiotemporal continuum of human experience in the cyberspace of media-environments and metaverses. In machine learning, the context of the chronotope allows us to connect the spatial structure of the interface with the principle of its existence in time. That essentially is modeling the vectors of potential interactions in a new locale. Every entry of an actor into the sphere of meaning occurs through the "chronotope gate" (Bakhtin, 1975, p. 406).

In machine learning, "gate of the chronotope" correlations can be temporal models based on grammatical structures that represent conceptual schemas reflecting the stereotype of human perception of time and space (Skvortsova, 2023). Event temporal models correspond to cognitive scripts that are based on the current interpretations of the semantic field interactions (Shipunova & Mureyko, 2016). In some cases, when describing the state of complex systems, computational temporalities are represented by fuzzy cognitive models that predict events in chronological order, taking into account corresponding components of multidimensional time-series (Borisov & Luferov, 2020). Modern models of temporal dynamics for training smart systems are based on recurrent neural networks that allow processing information in an orderly sequence of data. They also allow for different memory modes and the generation of texts, taking into account context orientation.

Temporal models based on metaphors, image schemes, and frames are formally presented in grammatical structures and phraseological expressions. Deep schemes for conceptualizing time provide the context of the chronotope at the level of an individual's subconscious perception within the space and time continuum and can be triggered in various combinations, depending on personal experience and cultural background. Thus, Russian and English narratives of the chronotope are presented in various temporal models of the world: cyclic, linear, reversible, spatial, dynamical, and propositional (Bronnikova 2017). In verbal temporal models, orientation and motion-images are presented which on the one hand allow for structuring a specific linguistic segment, and which on the other hand organize temporal and spatial perceptions of events.

In Lakoff's theory of conceptual metaphors the temporal perspective of interaction between knowledge systems is represented as "metaphorical projection" which creates stable forms for transferring features from one concept domain to another (Lakoff & Johnson, 1980). Conceptual metaphors also produce a chronotope model that reflects a coordinate system for imagined events and the sequence of changes in body positions. The interface matrix is modeling the temporal dynamics in the virtual environment, producing potential meanings and actions, and defining the physical and semantic limits of these actions.

COMMUNICATIVE DESIGN OF THE INTERFACE: ACTOR MODELS FOR MACHINE LEARNING

The digital world, where a person exists in a network of artifacts, rules, protocols and interfaces, expands not only the capabilities of the subject but also structures its actions. The interface defines the boundaries of multi-user interactions, with their effect

of full immersion in the network's semantics. The multilayered meanings within the semantic field are produced by the multitude of informational and temporal connections within the social network, creating the effect of a “deferred being” (Baudrillard, 2016). The automatism of actions based on algorithms and patterns simulates the circular movement of thoughts, in which the stable stereotypes of interpreting reality represent the actor's models that are embedded in this semantic context. The interface does not act as a neutral medium, but rather as an active contributor to the formation of collective attitudes and spaces for the creation of new meanings (Piccardi et al., 2025). Virtual media environments shape the habits, perceptions and collective behaviors of users through the transmission of meanings in contexts established by human culture, while simultaneously restructuring cognitive and emotional structures.

In the modern world, digital platform interfaces become the primary intermediaries between people and reality. Cultural archetypes become manipulation tools that serve as actor models in online interactions. Carl Gustav Jung argued that archetypes not only reflect but also actively shape human cultural experience, transforming the chaos of perception into ordered, symbolic forms through universal models of unconscious perception. These become manifest in myths, religion, and everyday practice, defining boundaries of individual and collective experience. Voice assistants and neural networks exploit the archetypal “wise mentor” to create the illusion of a personal, almost sacral interaction, making users dependent on algorithmically constructed authorities.

Research into the communication design of intelligent systems shows how the archetypical model is becoming a tool for creating digital actors. Designers of interfaces are consciously adopting archetypes such as the sage, the hero, and the wonderful assistant to create voice assistants and chatbots that are recognizable, reducing the user's cognitive burden and accelerating the formation of trust in the machine (Karimova & Goby, 2021). Moreover, the structure of archetypes is not fixed, but dynamically adjusted by algorithms to meet user expectations. Therefore, archetypes function not as static symbols of the collective unconscious, in the Jungian sense, but rather as programmable templates that engender trust and recognition in the asymmetric relationship between humans and artificial intelligence (Karimova, 2025).

An interface modeling AI as a “wise mentor” suggests the possibility of restructuring the process of belief revision by shifting the subject's autonomy as a source of knowledge towards the system that forms the context for their reasoning (Coeckelbergh, 2025). However, the role of AI systems in generating knowledge remains fundamentally uncertain because the criteria for competence and integrity that apply to human experts cannot be transferred to algorithmic architectures without losing meaning (Hauswald, 2025, Marchal et al., 2026).

Recent empirical studies have highlighted the effect of catastrophic forgetting caused by increasing contextual noise when a choice has to be made between speeding up information processing and maintaining long-term attention to a target setting. Methods for blocking information flows in education are implemented through strict program structuring (module hierarchy, prerequisites, content isolation), which creates a static dependency graph where access to the next element is impossible without mastering the previous one. This is also achieved through contextual inference — dynamic rules

that analyze the learner's current state and, in real time, open or close the flow of knowledge based on semantic relationships. Taken together, these methods form a two-level access control system that prevents cognitive overload, knowledge leakage, and unauthorized access to the full course overview, ensuring the sequential acquisition of material. (Shivkumar et al., 2025).

Extended knowledge retrieval in machine learning is associated with task complexity, generalization issues, and the utilization of knowledge in a given field. Research into interactive machine learning emphasizes that system performance crucially depends on how feedback from humans is collected, organized, and maintained (Amershi et al, 2014, 2019). In this respect, the interface acts as a mediator between information databases, human users, and knowledge-based systems. The interface's cognitive potential determines the criteria for consistency in knowledge representation based on shared formal structures and concepts in the cognitive activities of agents.

, Human norms and judgments are crucial for the creation and training of large language models. Current literature emphasizes that reinforcement learning, based on human feedback, is the dominant method for integrating generative AI systems with human preferences and values (Ouyang et al., 2022). In human-AI alignment, interfaces function as tools for resolving uncertainty in AI actions, transforming the process of formulating assumptions into informed decision-making. The effectiveness of feedback depends not only on the model's functionality but also on the interface which acts as a cognitive aid for identifying more informed assessments in the decision-making process (Du et al., 2026).

CHARACTERIZATION OF THE COGNITIVE POTENTIAL OF INTERFACES IN MACHINE LEARNING

The most significant features of cognitive potential in the interface for machine learning systems include factors that determine potential actions in standby mode and during activation, within the boundaries set by the sequence of functions defined by the actor model (Table 1).

Thus, it can be argued that the interface is a part of the cognitive organization of machine learning systems, since the semantic matrix links to the mechanisms of attention, information retrieval and memory, as well as feedback. The cognitive potential of the interface programs a sequence of actions with contextual structures based on the cognitive cycle: selection, binding, inference, revising, and acting. The chronotopic principle, as an essential condition for machine learning technologies, allows for predicting the temporal nature of events in different interactions, enabling events to occur at a new local level and in new semantic contexts.

Table 1. Features of the interface cognitive potential in machine learning

Determining factor for cognitive function in machine learning	Structure status in the interface	Potential cognitive function in machine learning
Structuring factor for the relevant information space	External architecture of cognitive process	Defining the boundaries of cognitive action: analyzing the initial conditions and the database – what is available, what is considered relevant, in what order it is considered, what connections become visible, what actions are suggested, when a return to the previous context occurs, when a check is performed, and when the system takes action.
Temporal dynamics factor of context formation	Computational Tools for the Temporal Dynamics of Context Temporal Model Chronotope Model	Activation of the temporal model at the moment of transition to new information, changing the space of possible conclusions and actions.
Factor activating the output procedure	Functional architecture of cognitive procedure Reasoning Model	The possibility of implementing a cognitive procedure for generating an inference in which the model architecture, the context structure, the user's actions, the information extraction mechanisms, and the principle of feedback organization jointly determine the sequence of operations.
Tuning modes factor for optimized cognitive function	External parameter of the accessible context Frame Semantic matrix	Activation of modes for optimizing the search for context available to the system, operations of selection, structuring, ranking and linking of relevant information objects, inclusion in the current context, formation of a semantic field of reasoning.
Motivation factor for cognitive actions	Goal Setting Actor Model	Activation of the cognitive role that the interface offers to the user and the system

Digital technologies are changing the way we consume information and thus they are restructuring the foundations of human experience. Within the knowledge network any object acquires a virtual character. It is imbued with temporal content that synchronizes the order of semantic orientation during its description. A large language model – a neural network that simulates reasoning processes – functions as a navigator

in semantic spaces, relying on the knowledge encoded in its parameters. When using the RAG architecture, it additionally accesses an external database, which allows it to refine and verify the generated content. The neural network's temporal dynamics are characterized by the production of relevant content, which is constructed as a multitude of increasingly connected contexts, constantly controlled by the subject of the discussion (question, task). In reinforcement learning, the interface not only defines the boundaries of meaning but also organizes the available context, determining the sequence of cognitive operations in computational correspondences, thereby shaping the conjugation of humans and AI. At the same time, machine learning requires a clearer distinction between informational, cognitive, and agent-based designs of the interface.

CONCLUSION

Modern machine learning technologies fit an artificial agent into the semantic matrix of an interface that is tied to a basic knowledge system and natural language of communication. From this point of view, the interface models a potential actor capable of navigating the contexts defined in the basic matrix and of adequately applyinging algorithms for actions within the boundaries of the semantics of the virtual environment.

The temporal horizon of cognitive activity in a virtual learning environment is expanded due to the impartial vector of past-present-future. The time factor in constructing a semantic horizon at the interface is connected to the principle of continuity of knowledge, ensuring continuity of cognitive space in society. Investigating the role of time in creating cognitive scenarios for actors' actions allows us to integrate potential boundaries and semantic horizons into factors that motivate cognitive actions in virtual environments. In this respect, the cognitive capacity of the interface involves creating the necessary circumstances for the subject to orient itself semantically, enabling goal-oriented behavior in machine-learning processes and practical tasks.

The cognitive potential of the interface in modeling a machine learning environment is determined by the semantics of the sign system and the laws of information technology. The contextual structure of the interface connects the construction of the semantic horizons of the interactive network with the principle of its existence in time, directing the vectors of potential interaction, defining the boundaries of a new locality of the digital environment in accordance with the semantics of the symbolism.

REFERENCES

Alchourrón C. E., Gärdenfors P., & Makinson D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic, 50*(2), 510–530.

Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *AI Magazine, 35*(4), 105–120. https://doi.org/10.1609/aimag.v35i4.2513

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., and Horvitz, E. (2019). Guidelines for human-AI interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13. ACM. https://doi.org/10.1145/3290605.3300233

Bakhtin, M.M. (1975). *Voprosy literatury i estetiki: Issledovaniya raznykh let* [Issues of literature and aesthetics]. Khudozhestvennaya Literatura Publishing House. (in Rus.)

Baudrillard, J. (2016). *Symbolic Exchange and Death.* SAGE Publications

Borisov V. V., & Luferov V. S. (2020). Metod mnogomernogo analiza i prognozirovaniya sostoyaniya slozhnykh sistem i protsessov na osnove nechetkikh kognitivnykh temporal'nykh modeley [The method of multidimensional analysis and forecasting states of complex systems and processes based on Fuzzy Cognitive Temporal Models]. *Systems of Control, Communication and Security, 2,* 1-23. https://doi.org/10.24411/2410-9916-2020-10201

Bronnikova, O. V. (2017). *Modusnyy metakontsept vremya v russkom i angliyskom yazykakh: na materiale dannykh natsional'nykh korpusov* [The Modal Metaconcept Time in Russian and English (Based on National Corpora]. [Ph.d. dissertation]. Tyumen State University.

Burkov, A. (2020). *Mashinnoye obucheniye bez lishnikh slov* [Machine learning without further ado]. Peter Publishing.

Coeckelbergh, M. (2025). AI and Epistemic Agency: How AI Influences Belief Revision and Its Normative Implications. *Social Epistemology, 40*(1), 59–71. https://doi.org/10.1080/02691728.2025.2466164

Du, Y., Chen, Y., & Xue, C. (2026). Shared mental models as cognitive scaffolds: Optimizing interface design for human-in-the-loop alignment tasks, *International Journal of Industrial Ergonomics, 114,* 103955. https://doi.org/10.1016/j.ergon.2026.103955

Floridi, L. & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences, *Minds & Machines, 30,* 681–694. https://doi.org/10.1007/s11023-020-09548-1

Hauswald, R. (2025). Artificial Epistemic Authorities. *Social Epistemology, 39*(6), 716–725. https://doi.org/10.1080/02691728.2025.2449602

Internet Encyclopedia of Philosophy. (n. d.). *Artificial Intelligence*. https://iep.utm.edu/artificial-intelligence/

Jiang, S. (2025). Restraining the formation of filter bubbles with algorithmic affordances: Toward more balanced information consumption and decreased attitude extremity. *Journal of the Association for Information Science and Technology, 76*(7), 989–1005. https://doi.org/10.1002/asi.24988

Karimova, G.Z., Goby, V.P. (2021). The adaptation of anthropomorphism and archetypes for marketing artificial intelligence. *Journal of Consumer Marketing, 38(2), 229–238.* https://doi.org/10.1108/JCM-04-2020-3785

Karimova, G.Z. (2025). Developing a Framework for Integrating Chatbots with Personality in Organizations. In: Humanizing AI with Personality. *SpringerBriefs in Computer Science* (pp 7–17). Springer, Cham. https://doi.org/10.1007/978-3-031-82327-5_2

Lakoff, G., & Johnson, M. (1980). *Metaphors We Live By*. University of Chicago Press.

Liu, Z., Cheon, S., Stanbury, A., Jiao, X., Xing, W., & Kang, H. (2026). (2026). Towards contextual-based AI: A scoping review of artificial intelligence in X reality for personalized learning, *Computers and Education: Artificial Intelligence, 10,* 100523. https://doi.org/10.1016/j.caeai.2025.100523

Marchal, N., Chan, S. C. Y., Franklin, M., Revel, M., Keeling, G., Fischli, R., Chandra, B., & Gabriel, I. (2026). Architecting Trust in Artificial Epistemic Agents. arXiv. https://doi.org/10.48550/arXiv.2603.02960

Mashour G. A., Roelfsema, P., Changeux, J.-P., & Dehaene, S. (2020). Conscious Processing and the Global Neuronal Workspace Hypothesis. *Neuron, 105*(5), 776–798. https://doi.org/10.1016/j.neuron.2020.01.026

Najafabadi, M. K., Chen, J., Saman Shojae Chaeikar, S.Sh., & Javad Rezazadeh, J. (2026). A survey and illustrative example of hybrid semantic and contextual models for improving diversity and accuracy in personalized tag recommendation, *Computer Science Review, 61,* 100934. https://doi.org/10.1016/j.cosrev.2026.100934

Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Simens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems, 35*, 27730–27744.

Petrunin, Yu. Yu. (2025). Generativnyy iskusstvennyy intellekt i problema soznaniya [Generative artificial intelligence and the problem of consciousness]. *Public administration. Electronic bulletin, 112, 78-92.* https://doi.org/10.55959/MSU2070-1381-112-2025-78-92

Piccardi, T., Saveski, M., Jia, C., Hancock, J.T., Tsai, J.L., & Bernstein, M.S. (2025). Reranking partisan animosity in algorithmic social media feeds alters affective polarization. *Science, 390*(6776), eadu5584. https://doi.org/10.1126/science.adu5584

Shipunova O.D., & Mureyko L.V. (2016). Cognitive scenarios in the design of mass consciousness, *Terra Linguistica, 3*(250), 93-101. https://doi.org/10.5862/JHSS.250.11

Shivkumar, S., Lengyel, M., & Wolpert, D.M. (2025). Curriculum effects in multitask learning through the lens of contextual inference, *Current Opinion in Neurobiology, 95,* 103123. https://doi.org/10.1016/j.conb.2025.103123

Shuang, K., Zhou, J., Wang, Q., Qian, B., Wei, Y., & Guo, J. (2025). Utilizing contextual summarizing and reasoning for enhancing document-level event argument extraction, *Expert Systems with Applications, 287,* 128075. https://doi.org/10.1016/j.eswa.2025.128075

Skvortsova, N.S. (2023). Mechanisms of conceptualization of aspectual-temporal meanings of grammatical constructions. Philology. *Theory & Practice, 16(5), 1505-1514.* https://doi.org/10.30853/phil20230238 (in Rus.).

Teovanović, T., Vesic, S., & Yun, B. (2026). Human compliance with computational argumentation principles, *Artificial Intelligence, 351, 104457,* https://doi.org/10.1016/j.artint.2025.104457

Tononi G. (2004). An Information Integration Theory of Consciousness. *BMC Neuroscience, 5, 42,* https://doi.org/10.1186/1471-2202-5-42 .

Wang, B., Ji, R., Zhang, L., Wu, Y., Liu, J. (2025). Learning to zoom: Exploiting mixed-scale contextual information for object detection, *Expert Systems with Applications, 264,* 125871. https://doi.org/10.1016/j.eswa.2024.125871 .

Wang, R., Sun, B., He, J., Qi, T., Zheng, G., Hoel, T., & Yuan, L. (2026). A hippocampal-PFC inspired neuro-symbolic architecture for contextually anchored question chain generation, *Neurocomputing, 686,* 133743. https://doi.org/10.1016/j.neucom.2026.133743

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

Лисенкова Анастасия Алексеевна, oskar46@mail.ru,
ORCID: 0000-0002-8825-3760

Anastasia A. Lisenkova, oskar46@mail.ru,
ORCID: 0000-0002-8825-3760

Шипунова Ольга Дмитриевна, o_shipunova@mail.ru,
ORCID: 0000-0001-8953-7434

Olga D. Shipunova, o_shipunova@mail.ru,
ORCID: 0000-0001-8953-7434

Статья поступила 4 июля 2026
одобрена после рецензирования 17 августа 2026
принята к публикации 21 сентября 2026

Received: 4 July 2026
Revised: 17 August 2026
Accepted: 21 September 2026