

DOI: <https://doi.org/10.17816/DD625319>

Эпистемический статус искусственного интеллекта в медицинских практиках: этические вызовы

А.В. Баева

Московский государственный университет имени М.В. Ломоносова, Москва, Россия

АННОТАЦИЯ

В современных научных исследованиях в последнее время всё чаще появляются дискуссии о том, что в связи с развитием технологий искусственного интеллекта встают вопросы об объективности, правдоподобности и достоверности знания, а также о том, не заменят ли эти технологии фигуру эксперта как ту инстанцию, которая до сих пор выступала гарантом объективности и центром принятия решений. Современные историки науки Л. Дастон и П. Галисон в своей книге, посвящённой истории научной объективности, говорят о сменяемости «эпистемических добродетелей», в качестве одной из которых с определённого момента утвердилась и объективность. При этом выдвигание той или иной добродетели, регулирующей научную самость, то есть выступающей нормативным принципом для учёного при выборе способа видения и научной практики, зависит от принятия решений в трудных случаях, требующих воли и ограничения самости. В этом смысле эпистемология соединяется с этикой: учёный, руководствуясь определёнными моральными принципами, отдаёт предпочтение тому или иному способу поведения, выбирая, например, не более точное изображение, сделанное от руки, а неретушированную фотографию, возможно, нечёткую, но полученную механически, а значит — более объективную и свободную от какой-либо примеси субъективности. В этой связи небезыntenесным представляется эпистемический статус современных технологий на основе искусственного интеллекта, которые всё больше берут на себя функции научной самости, в том числе и в части оказания влияния на принятие конечных решений и получение объективного знания. Так, например, в области медицины роботизированные аппараты уже оказывают существенную поддержку: им передаётся часть функций, например, врача первого звена для сбора и анализа стандартизированных данных о пациенте и диагностики. Есть предположение, что в ближайшее время всё больше обязанностей будет передаваться искусственному интеллекту: обработка данных, разработка новых лекарств и способов лечения, налаживание дистанционного взаимодействия с пациентом и др. Значит ли это, что научная самость может быть заменена алгоритмами на основе искусственного интеллекта, а на смену объективности придёт другая эпистемическая добродетель, окончательно разрывающая связь этики и эпистемологии, — этот вопрос нуждается в исследовании.

Ключевые слова: современные научные практики; объективность; эпистемическая добродетель; научная самость; технологии на основе искусственного интеллекта.

Как цитировать:

Баева А.В. Эпистемический статус искусственного интеллекта в медицинских практиках: этические вызовы // Digital Diagnostics. 2024. Т. 5, № 1. С. 120–132. DOI: <https://doi.org/10.17816/DD625319>

DOI: <https://doi.org/10.17816/DD625319>

Epistemic status of artificial intelligence in medical practice: Ethical challenges

Angelina V. Baeva

Lomonosov Moscow State University, Moscow, Russia

ABSTRACT

Advances in artificial intelligence have raised controversy in modern scientific research regarding the objectivity, plausibility, and reliability of knowledge, and whether these technologies will replace the expert figure as the authority that has so far served as a guarantor of objectivity and the center of decision-making. In their book on the history of scientific objectivity, modern historians of science L. Duston and P. Galison discuss the interchangeability of “epistemic virtues,” which now include objectivity. Moreover, selecting one or another virtue governing the scientific self, i.e., serving as a normative principle for a scientist when adopting a perspective or scientific practice, depends on making decisions in difficult cases that require will and self-restriction. In this sense, epistemology and ethics are intertwined: a scientist, guided by certain moral principles, prefers one or another course of action, such as choosing not a more accurate hand-drawn image but an unretouched photograph, perhaps fuzzy, but obtained mechanically, which means it is more objective and free of subjectivity. In this regard, the epistemic standing of modern artificial intelligence technologies, which increasingly perform the functions of the scientific self, including influencing ultimate decision-making and obtaining objective knowledge, is intriguing. For example, in medicine, robotic devices considerable support and are assigned some of the responsibilities of a primary care physician, such as collecting and analyzing standardized patient data and diagnosis. It is expected that artificial intelligence will take on more tasks such as data processing, development of new drugs and treatment methods, and remote interaction with patients. It remains to be seen whether this implies that the scientific self can be replaced by artificial intelligence algorithms and another epistemic virtue will replace objectivity, thus breaking the link between ethics and epistemology.

Keywords: modern scientific practices; objectivity; epistemic virtue; scientific self; artificial intelligence technologies.

To cite this article:

Baeva AV. Epistemic status of artificial intelligence in medical practice: Ethical challenges. *Digital Diagnostics*. 2024;5(1):120–132.

DOI: <https://doi.org/10.17816/DD625319>

Submitted: 26.12.2023

Accepted: 13.02.2024

Published online: 11.03.2024

DOI: <https://doi.org/10.17816/DD625319>

人工智能在医疗实践中的认识论地位：伦理挑战

Angelina V. Baeva

Lomonosov Moscow State University, Moscow, Russia

摘要

最近，在现代科学研究中越来越多地讨论认为，随着人工智能技术的发展，知识的客观性、可信度和可靠性出现了问题，以及这些技术是否会取代专家的问题。迄今为止，专家一直是客观性的保证和决策中心。现代科学史学家L. Daston和P. Galison在他们关于科学客观性历史的著作中谈到了“认识论美德”的更迭。客观性从某个时刻就确立为其中之一。美德是科学家选择理解方式和科学实践的规范性原则。管理科学信念的特定美德的弘扬过程取决于科学家在需要意志和信念限制的困难情况下做出的决策。从这个意义上说，认识论与伦理学是相通的。科学家在某些道德原则的指导下，倾向于这样或那样的行为方式。例如，科学家选择一张未经修饰的照片，而不是更精确的手绘图像。照片可能模糊不清，但它是通过机械方式获得的，这意味着这样的照片更加客观，未受到任何主观因素的影响。在这方面，以人工智能为基础的现代技术的认识论地位很有意思。这些人工智能技术越来越多地承担科学信念的功能，包括在影响最终决策和获取客观知识方面。例如，在医学领域，机器人设备已经开始提供重要的支持。一些功能被转移到这些机器上，例如一线医生收集和分析病人标准数据的功能和诊断功能。有一种假设认为，在不久的将来，越来越多的职责将移交给人工智能：数据处理、新药物和新疗法的开发、与病人建立远程互动等等。这是否意味着科学可能会被基于人工智能的算法所取代，客观性将被另一种最终打破伦理学与认识论之间联系的认识论美德所取代。这是一个需要探讨的问题。

关键词：现代科学实践；客观性；认识论美德；科学信念；基于人工智能的技术。

引用本文：

Baeva AV. 人工智能在医疗实践中的认识论地位：伦理挑战. *Digital Diagnostics*. 2024;5(1):120–132. DOI: <https://doi.org/10.17816/DD625319>

收到: 26.12.2023

接受: 13.02.2024

发布日期: 11.03.2024

INTRODUCTION

The integration of artificial intelligence (AI) technologies into modern scientific practices, especially within the medical field¹, poses significant questions for researchers. Among these questions are the epistemic nature of AI and the ethical challenges it presents. Clarifying the epistemic nature of AI becomes significant as its widespread adoption in scientific practice raises concerns about how technological agency may threaten the decision-making authority of healthcare professionals (HCPs) and challenges the traditional notion of objectivity as a fundamental virtue in scientific inquiry.²

In a notable study exploring the historical concept of objectivity, L. Daston and P. Galison [1] used specific material scientific practices, particularly the creation of visual images for scientific atlases. Their study showed that throughout history, objectivity as an epistemic virtue has rested on two main aspects: epistemic virtues (especially objectivity) and visuality.

“Delving into distinct forms of scientific perception places two crucial questions at the forefront: What practices are needed to produce this kind of image? What practices foster the development of a scientific persona capable of such a perception? The history of scientific vision consistently demands this double motion, toward the evolution of an epistemology centered on imagery, on one hand, and toward the ethical refinement of the scientific individual, on the other. Fidelity to nature has always borne a triple responsibility: visual, epistemological, and ethical. However, what unfolds when fidelity itself is abandoned and nature blends with the artifact? We concluded by glimpsing into contemporary scientific atlases: depictions in which creation is synonymous with observation” [1]. These two aspects are intertwined through particular methodologies of visualizing the functional elements of science.³

Through alterations in imagery and practices, various epistemic virtues are exemplified. In this respect, the challenge posed by the evolution of visualization

through digital and AI technologies impacts both the epistemological virtue of objectivity and the scientific persona. In current scientific discourse, various methods of visualization (including diagrams, maps, photographs, and the creation of atlases) have become dispensable, forming an integral part of argumentation. Concurrently, visualization transcends illustration; it becomes a form of research facilitated by the capabilities of modern digital technologies. A paradigm shift is underway in how science is perceived and practiced, characterized by a transition from representation to presentation. Manipulating the observed object or phenomena now equates to manipulating a visual representation. Computer modeling and imaging represent the subsequent revolutionary frontiers in science following observation and experimentation. In this context, a pressing question emerges regarding the present and prospective state of objectivity as an epistemic virtue in the era of digitization and scientific innovation, where technology and engineering play a significant role in knowledge production, shifting from discovering facts to inventing them. What are the emergent scientific methodologies? Does the evolution of virtues necessitate a re-evaluation of the underlying aims and goods associated with specific practices? Can we define the scientific persona in terms beyond virtues, and how might the incorporation of AI technology reshape it?

MODERN SCIENTIFIC PRACTICES: MATERIALITY AND EPISTEMIC STATUS OF ARTIFACTS

Contemporary scientific practices blur traditional boundaries between objectivity and subjectivity, the abstract and concrete, and the discovered and constructed. Artifacts have played a significant role in shaping scientific knowledge and influencing its main characteristics, notably objectivity, which is now viewed not as an abstract scientific quality detached from the observer but as intimately intertwined with subjective engagement. Advancements in technology have

¹ When we speak of modern scientific practices, we are referring to a fundamentally complicated and empirically diverse scientific space that includes not only propositional knowledge production modes, but also various non-propositional forms using graphs, diagrams, visual images, etc. A research project by A. Mol is one of the most striking examples of how scientific practice not only recognizes its object, but also creates it in practice. This study is dedicated to the multiplicity of medical practice ontology, using the example of the implementation of diseases such as atherosclerosis in branched practices [2].

² Objectivity, viewed as an epistemic virtue, emerges during a particular historical phases characterized by intricate coordination between the observer and the practice of observation. It manifests through distinct visual practices and visualization technologies, as epistemic virtues are cultivated as stable traits in specific research methodologies, thereby shaping a unique scientific identity. Each manifestation of the scientific self-pursues a particular good, implying that sustainable practices are those capable of fostering the evolution of epistemic virtues.

³ In the second half of the 19th century, the concept of “speaking for nature itself” emerged as a fundamental principle driving a new form of scientific objectivity. Concerned about human interference between nature and science, French physiologist Etienne-Jules Marais and his contemporaries, who studied many visual methods of science, turned to mechanical image reproduction to eliminate potential biases. Employing polygraphs, photographs, and other technologies, they attempt to create atlases that served as the definitive guides of observable science, similar to a scientific Bible. These atlases revolutionized discussions surrounding scientific objectivity [3]. Atlases serves as functional tool for visual sciences by training the observer to recognize certain objects as exemplary (commonly referred to as typical) and to perceive them in a certain manner. In instances where atlases present images captured through new instruments (such as X-ray atlases from the early 20th century), the entire field associated with atlas necessitates a fresh interpretive approach. Even in disciplines where other senses are essential, atlases rely on visuals, as they play a crucial role in refining observational skills [3].

led to the emergence of technical artifacts,⁴ constructed in scientific laboratories, and endowed with novel properties crucial for knowledge creation. Beyond their utilitarian function, the materiality of artifacts emerges as a critical aspect. Unlike idealized entities, artifacts are rooted in the real world, embedded within cultural environments, historical periods, and social practices, rendering them intentionally connected (ontologically, rather than causally) to the processes involved in their interaction. Through the act of creation, humans reshape and refine nature, generating new objects that serve as tools for conceptualizing idealizations and understanding the world. Technical artifacts employed in scientific endeavors possess distinctive characteristics and functionalities, constituting integral elements essential for maintaining scientific knowledge stability. Consequently, artifacts can be understood as components within a system imbued with “material,” capable of embodying culture-specific meanings, mechanisms for production, processes of learning and interpretation, and catalysts for cultural evolution. In essence, artifacts hold ontological significance, embodying the essence of culture in ways that transcend mere representation and contribute profoundly to the dynamics of knowledge creation and transformation [4].

The artificial nature of an object cannot be fully understood in isolation from contextual elements such as other objects, relationships, and operational frameworks; these elements collectively unveil the object’s essence as an artifact. As M. Polanyi suggests, the presence of a tool, or the tool itself, transcends mere mechanical adequacy; it assumes a role similar to an extension of the human body, seamlessly integrated into our physicality or expanding our bodily capabilities through incorporation. M. Lynch further elaborates on this notion, referring to Polanyi’s concept as “interiorization” [6], the process through which a physical tool becomes an integral part of our embodied experience. The dichotomy between “objective” and “constructed” realities emerges within laboratory-driven scientific practices, where the distinction between artifact and natural objectivity is a subject of intense scrutiny and debate.⁵

Lynch’s analysis of artifacts in scientific practices distinguishes between positive and negative artifacts. Positive artifacts are characterized by their external manifestation, such as a blemish on a microscope slide. They form part of the subjective conditions of observation, relying on

the instrumental conditions of perception. Lynch outlines several key characteristics of these artifacts. First, they are tangible and visible, making them accessible for examination. Second, they are prevalent and reproducible, often presenting routine challenges encountered in creating technical images. Furthermore, once recognized as artifacts, they can be effectively separated from appropriately constructed image features like those in an electron micrograph. However, the detection of such artifacts prompts consideration of whether they should be acknowledged and integrated into subsequent analyses and research endeavors.

However, Lynch highlights certain challenges associated with adopting an “ethnographic” focus on artifacts within laboratory settings. First, he observes that the artifacts presented as examples in reports exclude all potential artificial elements within the studies. This selective presentation may distort the understanding of what constitutes an artifact. Second, what is artificial is often determined by how it is presented in the reporting records. For example, a neural ultrastructure can be represented as an analytical dataset. This dataset assumes the status of an artifact within the field of neural entities accessible through analysis. The presentation format of the artifact may be characterized by the two-dimensionality of the photograph; black and white textural variations that delineate the forms of the photographed phenomenon; and sequencing photographic series to depict a continuous sequence of events [6]. Recording can be viewed as a means of visualizing an otherwise imperceptible phenomenon. However, in numerous instances, artifacts have been discovered in laboratory reports detailing previously unanticipated phenomena. These artifacts emerged as discoveries, representing new phenomena in previously unexplored areas. These are what Lynch referred to as situational or negative artifacts. In any case, Lynch notes that instances of such artifacts evolving into discoveries imply that the outcome of an observation or experiment is greatly influenced by the conditions under which it is conducted.

Lynch provides an illustrative example of the discovery of the microglia phenomenon as an artifact, elucidating the interplay between positive and negative artifacts within scientific practices [6]. Its “incredibility” stemmed not from empirical impossibility but from being an isolated occurrence among a more credible alternative narrative. The theoretical

⁴ “Artifacts are objects intentionally made to serve a given purpose; natural objects come into being without intervention of any agents. Artifacts inherently possess intended functions, while natural objects do not” [5]. On the one hand, artifacts are commonly understood as objects created for specific purpose, distinct from natural objects. On the other hand, modern epistemology studies rightly highlight that a technical artifact can be not only artificially designed but also a completely natural, living organism used to address certain challenges. In such instances, we must acknowledge that an artifact’s defining characteristic is not solely its artificiality but rather its use in human cognitive endeavors. Functionality stands as one of its core properties. “An artifact includes a vaccine, hadron collider, and poking stick. These objects are all connected to the human life-world, defining them as technical artifacts” [4].

⁵ Karl Popper observes that “objectivity is closely linked with the social aspect of the scientific method, emphasizing that science and scientific objectivity emerge not solely from an individual scientist’s effort to remain “objective” but from the collaborative yet adverse cooperation of numerous scientists. Scientific objectivity can be understood as the intersubjectivity of the scientific method. However, this social aspect of science is often neglected by those identifying as sociologists of knowledge” [6].

framework that gave particular significance to the capillary microglia (via a series of close-up images) resulted from a deeper investigation into the phenomenon than would otherwise have occurred. The microglia photograph posed challenges not only because it contradicted a laboratory assumption regarding microglia in brain physiology but also due to the presence of a competing assumption documented within the photo. Not only did the appearance of microglial cells in the laboratory version of brain physiology lack a clear explanation, but it also served as potential evidence for an alternative explanation. The depiction of microglia within capillaries reflects an unconscious construct more than an accurate representation of reality [6].

In this instance, the example of an artifact did not manifest in a tangible, positive sense. The actual appearance of the phenomenon was not the primary issue; rather, the challenge stemmed from a different interpretation of materiality. Axon sprouting showed a material extension in a direction contrary to what had been widely accepted as indisputable among laboratory researchers, sparking a challenge based on an alternative material argument [6]. This artifact became evident within a specific discourse regarding sprouting axons, leading to controversy. In scenarios like this, where conflicting viewpoints clash, the artifact assumes a role that is more similar to an “antithing” rather than a concrete object. Artifacts transcend being mere “things;” they can also represent opportunities that emerge in contrast to established expectations. These characteristics were often noted by their absence rather than presence observations (such as spots, stains, and blurring in photographs, which could be interpreted as “intrusions”).⁶ In this context, the artifacts arose within the realm of uncertainty.

Negative artifacts are not viewed as intrusions, distortions, or specific defects in the observed field but rather as the absence of the expected results or effects. In the context of negative artifacts, the lack of a positive result from an experiment or observation implies the adequacy of the laboratory procedures undertaken, allowing for an examination of any factors that may have contributed to the achieved result. Failure caused by uncertainty prompts an investigation into why the desired result was not achieved. Lynch considered the uncertainty associated with such negative results as an essential addition to the technical framework necessary for achieving objectivity [6]. The implications of failure vary depending on the local circumstances, with some instances being attributed to approximately objective factors.

In research endeavors aiming to minimize subjective errors, positive artifacts manifest as intrusions within the visual domain of a natural phenomenon. Conversely, negative artifacts signify the ongoing search for an elusive

object, highlighting the investigation process itself. However, mere search efforts are not enough to avoid errors; achieving success requires controlling circumstances to achieve the desired result. Negative artifacts represent the potential existence of “hidden” elements, much like the artifacts themselves, which conceal their presence until technical modifications unveil their existence during testing. Consequently, negative artifacts set the stage for actualizing previously unforeseen objects under specific circumstances. When errors occur, they are attributed to subjective factors hindering the accurate representation of the object itself. Furthermore, tools and equipment have imperfections, defects, and associated errors. As elaborated below, the materiality inherent in scientific practices profoundly influences the attainment of scientific objectivity and bestows epistemic significance upon the technologies utilized in its attainment.

MODERN MEDICAL PRACTICES: DISTRIBUTED AGENCY AND THE EPISTEMIC STATUS OF AI

The rejection of the traditional cognitive subject–object model within scientific research, marked by the “material turn,” attributes all aspects of the research process and its outcome (i.e., scientific knowledge) to social characteristics. “Forgetting artifacts (in the sense of tangible objects) has led to the creation of another kind of artifact (in the sense of an illusion): a society sustained solely by social constructs” [7]. As a result, the knowledge we obtain is determined by the social processes involved in its production. This acquired knowledge represents the final result of the scientist’s work. In classical science, knowledge acquires a logical form because it aligns with the object studied. However, in contemporary practice, knowledge attains scientific status not only because of its logical correspondence with reality but also due to its functional utility within society as an artifact. This functional success often marks the social processes involved in its production. The scientist’s role is not merely detached from the research object but involves a specific form of subjectivity characterized by submission to the object’s resistance to complete control. This dynamic creates a sense of scientific subjectlessness in which the scientist is an evaluator of an ever-present object but lacks ultimate authority in its judgment. B. Latour argues that within the realm of science, there is no concept of authoritative finality found in legal proceedings (“the authority of the adjudicated case (*res judicata*)” [8]). However, he introduces the notion of an independent hybrid entity as a third party in decision-making processes. These hybrids serve as representatives speaking on behalf of scientists who, in turn, speak on

⁶ For example, Galileo’s experimental method, in contrast to Bacon’s empirical approach, facilitates the integration of speculative frameworks and empirical models through technical engagement act. Simultaneously, the observation of sunspots through a telescope had to be justified as a product of observation rather than an artifact generated by the telescope itself.

behalf of “things” or objects under study. In this context, the scientist’s role shifts from attempting to dominate the object to enabling it to express itself (“make it speak”). The facts play a dual role; they represent what they speak about and determine the truth of their statements [8; p. 82].⁷

For Latour, science is not just discourse; it is primarily a network of practices for fact production. In the context of understanding science as technoscience, where technology and engineering are not merely applied to science but are integral to development, AI occupies a unique epistemic status as an agent (or actor) that cannot be excluded from the scientific practice where it operates. Consequently, the question of innovation in science during the digital age is closely related to how scientific identity is evolving. What actions must scientists take to nurture science, and how does objectivity as an epistemic principle fare in today’s landscape? By the end of the 20th century, the emergence of new technologies and a hybridized approach to scientific inquiry relegated what was once considered a method of representing nature to a secondary role. The integration of natural elements and human-made artifacts in the scientific realm, particularly in creating images at the atomic scale, shifts the focus from representation to presentation strategies.⁸ Within the context of nanotechnology’s evolution from the 20th to the 21st century, Daston and Galison introduced the concept of “image-as-tool” to describe a new approach to scientific visualization. This conceptual shift redefines modern scientific images; and transcends mere representations to become active tools for manipulating and exploring the depicted objects.⁹ In this sense, the most profound change, delineating the

shift from representative to representational strategies, unfolds precisely within the domain of the scientific self. In this amalgamation of disciplines, the delineation between the scientist and engineer, once sharply defined, gradually fades. As this convergence solidifies into a unified scientific-engineering identity, a new perspective on images emerges. No longer mere representations, these images assume an active role as tools that are seamlessly integrated into the scientific apparatus. They are similar to computer screens that reveal the intricate maneuvers of a robot performing surgery from a remote location, adjustments made to satellites orbiting in space, the processes involved in chemical reactions, or the delicate task of defusing a bomb [1]. Modern scientific practices, especially those driven by AI, increasingly seek to minimize human subjectivity¹⁰ in creating and observing objective images. This trend extends to the point of potentially eradicating the human self from the process to prevent any potential interference or misinterpretation of the observed phenomenon. This prompts a critical question: are these new technologies posing a threat to the traditional scientific self, historically guided by objectivity as an epistemic virtue? Moreover, are we witnessing the emergence of new epistemic regimes that break the link between ethics and epistemology? Alternatively, could algorithms replacing human agency in scientific endeavors be the new custodians of epistemic virtues?

For example, the image of a disciplined, meticulous observer who refrains from intervening in the process but only impartially records and accurately interprets observed phenomena embodies the essence of the “objective” scientist.

⁷ It is no coincidence that the realms of law and science are closely linked; they both share a common virtue, impartiality, achieved through meticulous distance and precision. Each area has its unique language and mode of thinking. For example, Latour suggests viewing the Council of State as a laboratory in the search for objectivity pursued by scientists. “The role of the *conseiller du gouvernement* is similar to that of a scientist to the extent that they speak and publish under their own name; similarly, scientist all possess elements similar to the *conseiller du gouvernement*, seeing themselves as enlightening the world. The *conseiller du gouvernement* is, thus a strange and complex hybrid, embodying the sovereignty of *lex animata*, law incarnated in a person, yet their declarations bind only themselves <...> the *conseiller du gouvernement* is a unique exemplar of producing objections, or, of objectivity” [8]. The fundamental link between legal and scientific endeavors lies in the art of manipulation of texts and records in a broader sense.

⁸ This transition is characterized by the following state: “On one side are the older atlases that aimed, through representation, at fidelity to nature. Capturing nature accurately on the page might align with the 18th-century concept of truth-to-nature, yet it could also adhere to 19th century’s mechanical objectivity or 20th-century trained judgment. On the other side are the newer forms of image galleries that serves as presentations, where the presentational strategy can include either new entities (such as rearranged nanotubes, DNA strands, or diodes) or the presentations’ explicit embrace of deliberate enhancements to clarify, persuade, delight—and sometimes, market” [1]. O.E. Stolyarova notes that by highlighting these two strategies (representational and presentational) Daston and Galison implicitly create an ontology of “collective formation” [9], with epistemological implications that, using I. Hacking terminology, involve intervention as the formation of the new rather than the reproduction of the existing. This pragmatically interpreted constructivism imposes an ontological framework on our theorizing and practice, shaping what is termed as “second nature.” In modern epistemology, the concept of the subject is evolving; the disembodied subject is giving way to the embodied subject. The outcomes of the embodied subject’s engagement with the world no longer merely yield subjective images of an objective reality but rather artifacts that, according to Latour, expand our capacities and connect us with other individuals and social groups, thereby changing our needs.

⁹ The ability of modern scientists to manipulate nanoobjects and their nanoimages is in itself amazing. However, equally surprising is the fact that “produced by an atomic force microscope that measures the force between a tiny probe and a surface over which the probe scans, [these figures are] not intended to depict a “natural” phenomenon. Instead, this and similar haptic images are part and parcel of the fabrication process itself” [1].

¹⁰ Note that the desire to minimize the self not only goes along with the desire to minimize subjectivity and thus errors related to the human factor, but “makes some routine procedures unnecessary for HCPs. The reduction of time and material costs is another important advantage of using AI in medicine” [10].

To achieve objective (i.e., to uphold the epistemic virtue of objectivity), it was not simply about advancing scientific knowledge, but primarily about relinquishing personal biases and desires, such as refraining from altering a photograph. In this context, a change in epistemic virtues is not just a change in scientific practices but a also change in the ethical guidelines that guide a scientist's behavior. However, as observed by Daston and Galison, "Yet these three virtues all served, each in its way, a common goal: what we have called a faithful representation of nature" [1]. In the 20th century, traditional methods of representing nature, which seemed self-evident, were pushed into the background with the emergence of new technologies. This shift also transformed and significantly broadened scientific identity, now acknowledging neural networks and AI-based technologies as essential non-negotiable elements in decision-making processes.¹¹

As visual representation in scientific endeavors is increasingly interconnected with computers and computational formats, their digital materiality requires a special approach.¹² In the not-so-distant past, as the 20th century transitioned into the 21st century, there was a belief that the role of the scientist-observer would eventually be supplanted by enhanced algorithms and imaging technologies devoid of human intervention. However, the creation of new digital atlases, in contrast with earlier brain mapping methods, now imposes new requirements on exercising control and limiting personal biases in pursuit of what is termed "digital objectivity"¹³ [12]. Digital scans are integral to a complex infrastructure that provides visual knowledge in a manner distinctly different from merely assessing mechanically generated objective representations by an observer. Alongside the objective perspective, a relational viewpoint becomes imperative, treating the image as a dataset intertwined with the object under investigation. Within the realm of big data in science, the pursuit of delivering intricate and exhaustive data representations often results in a selective reflection of significant information. This selectivity is largely shaped by the technologies and platforms used for data collection, as well as the fundamental ontological perspectives guiding

the data. In essence, the data suggest a selective view that is tuned certainly and limited to the use of certain tools [13].

Emerging methodologies in data analysis, such as advances in machine learning, computer vision, and innovative visualization techniques, are revolutionizing modern scientific investigations. In fields like nanoscience, where the emphasis lies on discovering and exploring new phenomena, a unique form of visualization is crucial to capture these phenomena effectively. However, the question of whether a new mode of representation is emerging may not have a straightforward answer; there could be uncertainty regarding the novelty of such methods. Nevertheless, there is no doubt that in fields like nanotechnology and other dynamic scientific fields, strict adherence to replicating the exact properties of the study object is no longer the dominant requirement for the object under study. Digital atlases, beyond attaining mechanical objectivity through scanning and visualization technologies, "are shaped by the deployment of computer-supported statistical and quantitative tools, serving as additional means for validation and ensuring objectivity" [12]. In these contexts, the assumption is that digital imaging can promote the epistemic ideal of objectivity by using automated processes, thereby reducing the need for human intervention in data processing.

Finally, the evolution of scientific value manifests itself in the transformation of the epistemic virtue of objectivity. Initially rooted in the historical ideal of science, objectivity now assumes a new form as a scientific value, intertwined with the pursuit of refining artifacts toward a more instrumental image. This transition from representation to presentation becomes a turning point in the history of visual practices and objectivity, highlighting the coupling of representation practices with their construction process; the accuracy of photographic images does not delegate objectivity to technology as a desire to minimize subjectivity.¹⁴

Introducing AI-based technologies and computer vision aims to standardize image streams for primary and automated defects and pathology detection, as well as upscale screening programs. For example, a biomedical

¹¹ For example, observing and visualizing using digital atlases of the brain is mainly done behind a computer monitor [11]. This implies changing the relationship between the observer, the object observed, the technologies used, and the institutional arrangements that enable the practice of surveillance. A digital atlas, in contrast to the atlases discussed by Daston and Galison, takes on the characteristics of a tool that is not so much a representative as a presenter, since it can both represent and be used to improve representations.

¹² For example, a brain scan result is not a static snapshot, and some of Daston and Galison's assumptions about mechanical objectivity do not directly apply to brain scans [11]. Advances in computer technology have integrated brain scans into a digital and networked context, leading to brain scans less representative but more presentative.

¹³ During the 1990s, known as the Decade of the Brain, numerous digital and electronic resources were developed to facilitate the organization and integration of various neuroscience sub-fields. This approach, termed neuroinformatics, aims to rationalize and integrate sub-fields. In the process of developing atlases, the definition of objective neuroscientific knowledge undergoes significant redefinition. This redefinition is influenced by the technological possibilities of these tools and the standardization constraints inherent in projects involving multiple measurements. The term "digital objectivity" is proposed to describe a specific configuration of ideals, methodologies, and cognitive objects in modern cyberscience [12].

¹⁴ For example, R. Buiani details cases where technology falls short in detecting significant differences, prompting researchers to manually enhance, highlight, and organize image elements [14].

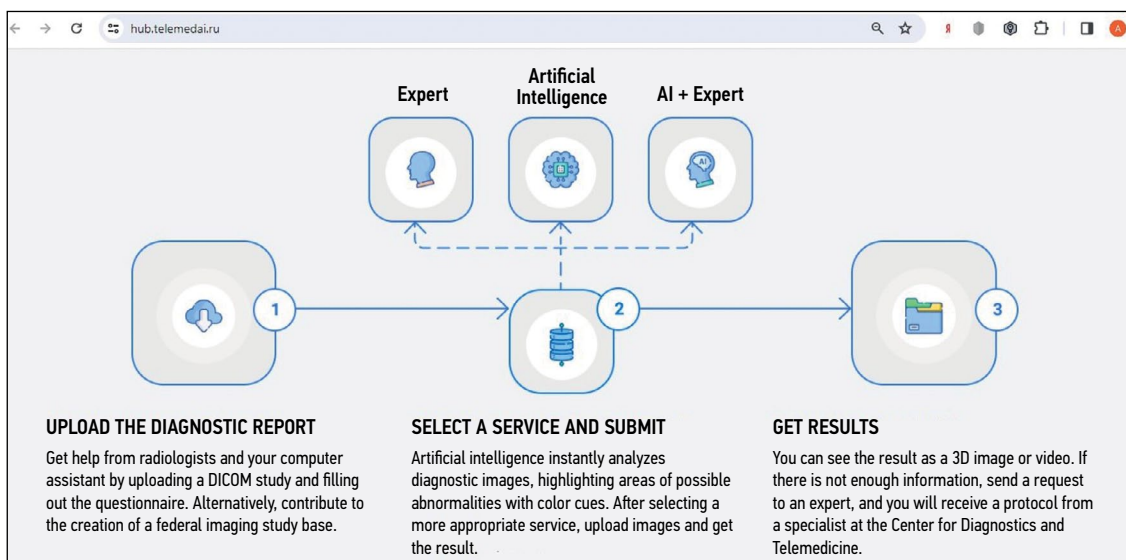


Fig. 1. Screenshot of the project page for a telemedicine platform for HCPs describing the examination process.

The screenshot shows a 'Our Services' section with three service cards. The first card, 'AI SERVICE + EXPERT', describes computer vision algorithms used by Moscow radiologists to analyze patient images in under 15 minutes, with a 'Read more' button. The second card, 'EXPERT', describes an agreement for healthcare organizations to gain access to services of the Moscow Reference Center for Diagnostic Radiology, with a 'Read more' button. The third card describes a federal project to upload COVID-19 confirmed CT data into a nationwide database, with a 'Read more' button.

Fig. 2. Screenshot of the project page for a telemedicine platform for HCPs describing the services provided.

image analysis service in Moscow explores using AI data analysis for decision support in healthcare.¹⁵ This service minimizes diagnostic error risks but does not eliminate false positives. HCPs are integral to the process but work alongside AI in a hybrid model. HCPs' objectivity in decision-making hinges on data analysis from the service. Initially, however, HCPs' expert opinion is not required, only becoming necessary if the AI results are unsatisfactory. Computer vision algorithms analyze images, reviewed by experts if necessary (Fig. 1 and 2).

ETHICAL CHALLENGES: PROS AND CONS OF AGENCY AND THE SUBJECTIVITY OF AI

AI-based robots are already providing significant assistance to both HCPs and patients in diagnosis, therapy, and surgery. Russia has embraced robotic medical systems, as exemplified by the Assisted Surgical Technologies robotic surgeon.¹⁶ In therapy, a preliminary diagnosis is traditionally made by a primary care physician. However, robots are already doing this job; special

¹⁵ The study includes three projects: Experiment on the Use of Innovative Computer Vision Technologies for Medical Image Analysis and Further Application in the Moscow Healthcare System; HUB AI Consultant (service for automatic X-ray analysis for HCPs), and Speech Recognition Technologies in Healthcare using an AI-based technology for automatic conversion of spoken speech to text to help HCPs to voice control a workstation and dictate diagnostic findings instead of typing them manually [15].

¹⁶ https://new.fips.ru/registers-doc-view/fips_servlet?DB=RUPAT&DocNumber=2715400&TypeFile=html [Accessed 09 February 2024].

sensors placed on the patient's body gather all the information and transmit it to the HCP in case of abnormalities. The system can diagnose in place of an HCP. The Russian RoboScan diagnostic system performs automated ultrasound scanning.

The proper use of neural networks in medical practice, such as detecting pulmonary COVID-19 lesions, helps in reducing tomography's radiation dose. A pre-trained neural network model acts as an expert,¹⁷ enhancing objectivity by standardizing data collection, initial diagnosis, and, sometimes, preliminary decisions. This shift reduces the burden on HCPs, allowing them to focus on data analysis, interpretation, and conclusion. Increasingly, HCPs delegate responsibilities to AI, including data processing, diagnosis, treatment planning, patient interaction, and decision-making. However, this trend prompts questions about AI's potential to completely replace HCPs and the ethical challenges that may arise [15–16].

What are the consequences of misdiagnosis or failure to detect a pathology, and who bears the responsibility for these decisions? Russia stands among the pioneers globally in recognizing the risks and threats, delineated in the AI Code of Ethics [17], as threats to human rights and freedoms, associated with the digitization and application of AI technologies within the medical field. These threats to discrimination, loss of privacy, loss of control over AI, potential harm to individuals stemming from AI errors, and misuse of AI. For example, the Russian Service for Surveillance in Healthcare recently suspended the use of a system designed to analyze computed tomography images, known as Botkin.AI, citing concerns over "the threat to the life and health of citizens."¹⁸

The traditional domain of decision-making, once the sole purview of human experts, is now shifting toward AI systems. Given that achieving a Technosphere similar to nature necessitates delegating decision-making authority to technical systems, this trend is expected to continue over the next 10–20 years.¹⁹ Concurrently, the digital transformation of modern medicine is occurring not only at a procedural level but also at the communicative interface, where HCPs and patients may find themselves separated by technological barriers. Indeed, digital technologies usher in an era of expanded network space, significantly augmenting the potential to

bridge and surpass existing, notable geographical boundaries between HCPs and patients. This paradigm shift also opens avenues for the potential replacement and displacement of expert HCPs from their traditional professional domains, yet, among these transformations, opportunities have emerged to form a networked collective expert subject through digital laboratories and multidisciplinary discussions [19].

The inclination toward substituting and partially displacing the expert functions of the HCP with quasi-expert functions using digital technologies indicates a new form of communication. This shift moves from the traditional dynamic between an expert (medical professional) and a layman (a nonspecialist patient) to a hybrid model of "doctor plus software to patient" and, in the long run, to a "software to patient" communicative model. This evolution challenges the expert status of HCPs, transferring the role of possessing absolute or near-absolute knowledge to digital programs. Although this model is technocentric, it also, to some extent, becomes patient-centered by leveling the physician's role [19; pp. 166–167]. The delegation of expert functions to technologies reflects a broader trend aimed at mitigating diagnostic errors by HCPs. The higher accuracy of AI in diagnosing pathology or predicting disease risks fuels the desire to substitute expert functions with algorithms. Consequently, decision support systems claim to be not merely human tools but full-fledged actors performing complex procedures. This progression diminishes irreparable biases. Technology assumes the responsibility of making judgments about the reality it perceives [19; pp. 167–168]. In this context, it is foreseeable that if the trend persists to limit human involvement in favor of AI, we will inevitably confront the need to view AI-based technologies not merely as tools but as entities with full agency and subjectivity, accompanied by their advantages and drawbacks.

ADDITIONAL INFORMATION

Funding source. This article was not supported by any external sources of funding.

Competing interests. The author declares that she has no competing interests.

¹⁷ "The proposed method reduces the total number of X-ray projections and the radiation dose required for COVID-19 detection without significantly affecting the prediction accuracy. The proposed protocol was evaluated on 163 patients from the COVID-CTset dataset and achieved an average dose reduction of 15.1%, while the average reduction in prediction accuracy was only 1.9%. Pareto optimality was improved compared to the fixed protocol" [16].

¹⁸ <https://www.kommersant.ru/doc/6350252> [Accessed 09 February 2024].

¹⁹ No comprehensive studies have yet explored the perspectives of HCPs and patients regarding their the implementation of AI into medical practice. However, a recent public opinion survey, the first of its kind to assess HCPs' interest in using AI in medicine and healthcare while also identifying challenges and prospects, reached an optimistic conclusion. According to the survey, "Russian HCPs are supportive of AI in medicine. Most respondents believe that AI will not replace them in the future but will instead serve as valuable tool for optimizing organizational processes, research, and diseases diagnosis." According to the report, several potential challenges in using AI were highlighted by respondents. These include concerns about the lack of flexibility and limited applicability in controversial situations (cited by 64% and 60% of respondents, respectively). Additionally, 56% believe that decision-making using AI could be challenging when there is no sufficient information available for analysis. One-third of HCPs expressed worry about the involvement of specialists with limited experience in AI. Notably, 89% of respondents believe that HCPs should be actively participate in the development of AI for medicine and healthcare. Interesting, only 20 respondents (6.6%) agreed that AI could replace them at work. However, a significant majority (76%) of respondents believe that in the future, doctors who use AI will replace those who do not" [18].

REFERENCES

1. Daston L, Galison P. *Objectivity*. Ivanov KV, editor. Moscow: Novoe literaturnoe obozrenie; 2018. (In Russ). EDN: PIXKTY
2. Mol A. *The body multiple ontology in medical practice*. Gavrilenko SM, Pisarev AA, editors. Perm: Gile Press; 2017. EDN: LJXRPN
3. Daston L, Galison P. The Image of Objectivity. *Representations*. 1992;(40):81–128. doi: 10.2307/2928741
4. Maslanov EV. Artifact: culture and nature. In: *Epistemology today. ideas, problems, discussions*. Kasavin IT, Alekseeva DA, Antonovskii AYu, et al, editors. N. Novgorod: Lobachevsky State University of Nizhni Novgorod; 2018. P:295–299. EDN: ELYPTC
5. Baker LR. Ontological significance of artifacts. In: *Ontologies of artifacts: the interaction of "natural" and "artificial" components of the lifeworld*. Stolyarova OE, editor. Moscow: Izdatel'skii Dom «Delo» RANKhiGS; 2012. P:18–33. (In Russ).
6. Lynch M. *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. London/Boston/Melbourne: Routledge & Kegan Paul; 1985.
7. Latour B, Vakhshtain V, Smirnov A. On interobjectivity. *Russian sociological review*. 2007;6(2):79–96. (In Russ). EDN: JWURYH
8. Latour B. Scientific objects and legal objectivity. *Kul'tivator*. 2011;(2):74–95. (In Russ).
9. Stolyarova OE. The historical context of science: material culture and ontologies. *Epistemology and philosophy of science*. 2011;30(4):32–50. (In Russ). EDN: OPDQIX
10. Alekseeva MG, Zubov AI, Novikov MYu. Artificial intelligence in medicine. *International Research Journal*. 2022;(7(121)):10–13. EDN: JMMMDf doi: 10.23670/IRJ.2022.121.7.038
11. Beaulieu A, de Rijcke S. Networked Neuroscience: Brain Scans and Visual Knowing at the Intersection of Atlases and Databases. Coopmans C, Woolgar S, editors. In: *Representation in Scientific Practice Revisited*. Coopmans C, Vertesi J, Lynch M, Woolgar S, editors. Cambridge: MIT Press; 2014. P:131–152. doi: 10.7551/mitpress/9780262525381.003.0007
12. Beaulieu A. Voxels in the Brain: Neuroscience, Informatics and Changing Notions of Objectivity. *Social Studies of Science*. 2001;31(5):635–680. doi: 10.1177/030631201031005001
13. Kitchin R. Big data, new epistemologies and paradigm shifts. *Sociology: methodology, methods, mathematical modeling (4M)*. 2017;(44):111–152. EDN: YMAFTQ
14. Buiani R. Innovation and Compliance in Making and Perceiving the Scientific Visualization of Viruses. *Canadian Journal of Communication*. 2014;39(4):539–556. doi: 10.22230/cjc.2014v39n4a2738
15. Center for diagnostics and telemedicine [Internet]. Moscow; c2013–2023 [cited 2023 Nov 29]. Available from: <https://mosmed.ai/>
16. Bulatov KB, Ingacheva AS, Gilmanov MI, et al. Reducing radiation dose for NN-based COVID-19 detection in helical chest CT using real-time monitored reconstruction. *Expert Systems with Applications*. 2023;229 Part A. doi: 10.1016/j.eswa.2023.120425
17. AI ethics code [Internet]. AI Alliance Russia, c2020–2024 [cited 2024 Feb 9]. Available from: <https://ethics.a-ai.ru/>
18. Orlova IA, Akopyan ZhA, Plisyuk AG, et al. Opinion research among Russian Physicians on the application of technologies using artificial intelligence in the field of medicine and health care. *BMC Health Services Research*. 2023;23(1):749. doi: 10.1186/s12913-023-09493-6
19. Popova OV. Digitalization and transformation of medicine: problems and prospects for development. In: *Modern problems of socio-technical-anthroposphere: a collective monograph*. Budanov VG, editor. Kursk: Universitetskaya kniga; 2022. P:153–171. (In Russ).

СПИСОК ЛИТЕРАТУРЫ

1. Дастон Л., Галисон П. Объективность / под ред. К.В. Иванова. Москва: Новое литературное обозрение, 2018. EDN: PIXKTY
2. Мол А. Множественное тело: Онтология в медицинской практике / под ред. С.М. Гавриленко, А.А. Писарева. Пермь : Гиле Пресс, 2017. EDN: LJXRPN
3. Daston L., Galison P. The Image of Objectivity // *Representations*. 1992. N 40. P. 81–128. doi: 10.2307/2928741
4. Масланов Е.В. Артефакт: культура и природа. В: Эпистемология сегодня. Идеи, проблемы, дискуссии / под ред. И.Т. Касавина, Д.А. Алексеевой, А.Ю. Антоновского, и др. Н. Новгород : Национальный исследовательский Нижегородский государственный госуниверситет им. Н.И. Лобачевского, 2018. С. 295–299. EDN: ELYPTC
5. Бейкер Л.Р. Онтологическая значимость артефактов. В: Онтологии артефактов: взаимодействие «естественных» и «искусственных» компонентов жизненного мира / под ред. О.Е. Столяровой. Москва : Издательский дом «Дело» РАНХиГС, 2012. С. 18–33.
6. Lynch M. *Art and Artifact in Laboratory Science: A Study of Shop Work and Shop Talk in a Research Laboratory*. London/Boston/Melbourne : Routledge & Kegan Paul, 1985.
7. Латур Б., Вахштайн В., Смирнов А. Об интеробъективности // *Социологическое обозрение*. 2007. Т. 6, № 2. С. 79–96. EDN: JWURYH
8. Латур Б. Научные объекты и правовая объективность // *Культиватор*. 2011. № 2. С. 74–95.
9. Столярова О.Е. Исторический контекст науки: материальная культура и онтологии // *Эпистемология и философия науки*. 2011. Т. 30, № 4. С. 32–50. EDN: OPDQIX
10. Алексеева М.Г., Zubov A.I., Новиков М.Ю. Искусственный интеллект в медицине // *Международный научно-исследовательский журнал*. 2022. № 7 (121). С. 10–13. EDN: JMMMDf doi: 10.23670/IRJ.2022.121.7.038
11. Beaulieu A., de Rijcke S. Networked Neuroscience: Brain Scans and Visual Knowing at the Intersection of Atlases and Databases. Coopmans C., Woolgar S., editors. In: *Representation in Scientific Practice Revisited*. Coopmans C., Vertesi J., Lynch M., Woolgar S., editors. Cambridge : MIT Press, 2014. P. 131–152. doi: 10.7551/mitpress/9780262525381.003.0007
12. Beaulieu A. Voxels in the Brain: Neuroscience, Informatics and Changing Notions of Objectivity // *Social Studies of Science*. 2001. Vol. 31, N 5. P. 635–680. doi: 10.1177/030631201031005001
13. Китчин Р. Большие данные, новые эпистемологии и смена парадигм // *Социология: методология, методы, математическое моделирование*. 2017. № 44. С. 111–152. EDN: YMAFTQ
14. Buiani R. Innovation and Compliance in Making and Perceiving the Scientific Visualization of Viruses. *Canadian*

Journal of Communication. 2014. Vol. 39, N 4. P. 539–556. doi: 10.22230/cjc.2014v39n4a2738

15. Центр диагностики и телемедицины [Internet]. Москва; c2013-2023 [дата обращения: 29.11.2023]. Доступ по ссылке: <https://mosmed.ai/>

16. Bulatov K.B., Ingacheva A.S., Gilmanov M.I., et al. Reducing radiation dose for NN-based COVID-19 detection in helical chest CT using real-time monitored reconstruction // Expert Systems with Applications. 2023. Vol. 229 Part A. doi: 10.1016/j.eswa.2023.120425

17. Кодекс этики в сфере ИИ [Internet]. AI Alliance Russia, c2020-2024 [дата обращения: 09.02.2024]. Доступ по ссылке: <https://ethics.a-ai.ru/>

18. Orlova I.A., Akopyan Zh.A., Plisyuk A.G., et al. Opinion research among Russian Physicians on the application of technologies using artificial intelligence in the field of medicine and health care // BMC Health Services Research. 2023. Vol. 23, N 1. P. 749. doi: 10.1186/s12913-023-09493-6

19. Попова О.В. Цифровизация и трансформация медицины: проблемы и перспективы развития. В: Современные проблемы социо-техно-антропосферы: коллективная монография / под ред. В.Г. Буданова. Курск : Университетская книга, 2022. С. 153–171.

AUTHOR'S INFO

Angelina V. Baeva, Cand. Sci. (Philosophy);
address: 27-4 Lomonosovsky Prospekt,
119192, Moscow, Russia;
ORCID: 0009-0005-5871-6217;
eLibrary SPIN: 2951-1427;
e-mail: a-baeva93@mail.ru

ОБ АВТОРЕ

Баева Ангелина Викторовна, канд. философ. наук;
адрес: Россия, 119192, Москва,
Ломоносовский проспект, д. 27, к. 4;
ORCID: 0009-0005-5871-6217;
eLibrary SPIN: 2951-1427;
e-mail: a-baeva93@mail.ru