Philosophical Problems in Science (2025), **79**, 73–90
(Zagadnienia Filozoficzne w Nauce)
doi:10.59203/zfn.79.730

ARTICLE

More than a tool: Artificial Intelligence in a philosophical history of technologies

Miroslav Vacura

Prague University of Economics and Business, Czech Republic
Email: vacuram@vse.cz

Abstract

This article examines artificial intelligence and shows how it is embedded in a history of technology. It develops a philosophical classification of tools as technologies based on the degree and nature of mediation that technology provides between humans and the world. Drawing on phenomenological and post-phenomenological conceptions—in particular the work of Ihde, Heidegger and Merleau-Ponty—it identifies five distinct stages in the historical development of technology, ranging from simple passive extensions of the body to advanced artificial intelligence systems. Each stage is defined by a qualitatively different mode of technological mediation between humans and the world, each transforming human perception and intentionality in increasingly complex ways. The most recent contemporary fifth stage in the evolution of technology, is characterized by the use of artificial neural networks, and the article demonstrates that a significant break from lower artificial intelligence systems separates this technology. Unlike simpler machines, these systems exhibit partial autonomy, are not epistemically transparent, and their outputs are unpredictable, forcing us to rethink their instrumental nature and the nature of instrumentality itself.

Keywords: postphenomenology, technological mediation, artificial intelligence, artificial neural networks

Introduction

One of the recurring themes of philosophical inquiry is the relationship between humans and technology, from ancient considerations of tools for crafts (τέχνη, *techné*) to contemporary debates about artificial intelligence and robotic systems (James 2019). This article considers artificial intelligence in the context of the historical development of tools and technologies, and proposes a systematic typology of technological development, based on phenomenological and post-phenomenological analyses and drawing on insights from cognitive science. Instead of the classification of technologies according to their social role and context, which is common among many authors (Mumford 1934; Marx 1990; Nolan and Lenski 2009), this study adopts a functional-structural approach based on emphasizing the nature of the mediation between human intentionality, technology and the world.

Taking as a starting point the works of Don Ihde, Martin Heidegger, Maurice Merleau-Ponty, James J. Gibson and others, our study explores the evolution of the human relationship to tools, or technology in general, from simple hand tools to contemporary advanced artificial intelligence. Each stage of described historical trajectory shows the fundamentally different nature of tools or technologies, showing a progression from tools functioning as simple extensions of limbs to tools increasingly gaining independence and autonomy. Mapping this evolution not only allows us to better clarify the philosophical conception of tools and technologies, but also to provide a framework for understanding the ontological context of AI technologies.

The final part of this study focuses on advanced artificial intelligence systems based on artificial neural networks, showing that they represent a fundamentally new form of technology. These systems exhibit a high degree of autonomy, are epistemically largely unpredictable, and their behavior is nontransparent.

These systems go beyond the classical notion of tools, which understands them as an extension of the human body and intention. As a result, we show that it will be necessary to rethink a number of philosophical categories associated with the relationship between humans and technology. We conclude the text by speculating on a hypothetical sixth stage of technological evolution: the emergence of artificial consciousness.

Philosophical Approaches to Tools and Technological Mediation

The proposed typology of technologies and its conceptual framework is based on an exploration of several influential philosophical perspectives on tools and technological mediation. Drawing on phenomenology, postphenomenology and Gibson's ecological approach, this section delineates the essential components of the relationship between humans, tools and world. This analysis provides a starting point for understanding how technology transforms human perception, action and intentionality, and places in context the subsequent distinction between different stages of technological development.

If we wish to clarify the nature of the relationship between humans and technology, we must start with their distinction from a non-technological, natural, approach to the world. Ihde (1979, p.18) distinguishes two approaches to the world. The first is the non-technological approach to the world according to the scheme *Human → World*. The second is approach in which the instrument has a mediating role in the relationship between humans and the world: *Human—Instrument—World*. For Ihde, the simplest example of technology is the dentist's probe, a pointed piece of stainless steel that the dentist uses to probe his patient's teeth.

Other authors, e.g. Sullins (2011, p.152) in this context talk about the trinity of *user, tool and victim*—the term victim shifts the discussion to the moral domain, where we usually talk about situations where the user uses technology to injure a victim. This is, of course, a very specific approach that has its place in moral theory, but in this text we will draw on Ihde's approach, which is not so directly normatively loaded.

Ihde analyses the use of technology in general—a dentist's probe used to check the health of a patient's teeth is an informational tool—for it is used to collect information about the patient's teeth. In doing so, the dentist does not feel his probe with his hand but feels the tooth with the probe. Ihde's dentist is similar to the blind man, in the example given by Merleau-Ponty, who uses his stick to feel the space where he is walking—the stick is thus an extension of the blind man's touch: "The blind man's stick has ceased to be an object for him, and is no longer perceived for itself; its point has become an area of sensitivity, extending the scope and active radius of touch, and providing a parallel to sight." (Merleau-Ponty 2005, p.165).

The realization of the intention to touch thus occurs at the end of the dentist's probe, at the end of the blind man's cane. Ihde (1979, p.19) says that the dentist feels the surface of the tooth at the end of the probe, not the probe at the point where it touches his palm—this is touch "at a distance". It turns out that not only sight and hearing are senses that can mediate sensory perceptions at a distance, but this is also the case for touch, which of course needs technological mediation. The technical enhancement of touch can take several forms based on different stages of technological development.

This type of tools is also thematized by Heidegger (2010, p.68) from which Ihde draws inspiration when he distinguishes different ways of human experience of the world. Human activity in the world largely consists of skillfully handling the things that are in the world. Useful things are in many cases simple tools—utensils for writing, utensils for sewing, utensils for working—but then there are often more complex technologies. These things have a pragmatic character, Heidegger refers to the Greek term πράγματα (*pragmata*), which refers to "things" in general, but is derived from the word πρᾶξις (*praxis*), denoting human activity or action. In this world view we thus understand things, as useful things, as tools [*Zeug*].

In this sense, it is possible that the Greeks, who understood things as πράγματα (*pragmata*), understood them primarily through their practical function, through their utility to themselves. We can consider in this context that they did not have a concept of a neutral thing that did not have a pragmatic purpose. A thing, the moment it becomes an object of human thought, always takes on a pragmatic aspect at the same time—at least in the background there is a consideration of what this thing can be useful for in human πρᾶξις (*praxis*) (Aristotle 1926, VI.2–5; Heidegger 2010, §15–18).

Heidegger emphasizes the relational character of these things—he says that a useful thing is essentially "something in order to [*um-zu*]..." (Heidegger 2010, §15), it thus involves a reference [*Verweisung*] to something else outside of itself. A useful thing is never available on its own, it is always embedded in a network of all useful things that only make it what it is: "There always belongs to the being of a useful thing a totality of useful things in which this useful thing can be what it is." (Heidegger 2010, §15).

Another related approach to this understanding of tools is the concept of affordances, developed by Gibson (2015) as part of his broader ecological approach, which was influenced by the work of Gestalt psychologists but had a number of original moments. The concept of affordances is based on the observation that when a person (or an animal) moves around in an environment, they do not perceive their surroundings homogeneously as a neutral space. The person perceives affordances on objects—places where an action is possible, places that the object "offers" to the agent, depending on the physical properties of both the agent and the object. Gibson argues that affordances are consciously perceived directly, they are not an afterthought derived from primarily perceived properties such as size, shape, and color. The agent directly perceives the possible points on which action is offered without any additional conscious processing of the perceived material.

A tool is an object that has affordances, offering us some "in order to" something else. We can generalize and extend the notion of a tool and consequently something else need not be another object, it can be another state of affairs. The handle on the door offers itself to us as a "tool" whose use leads to a different state of affairs, one where the door is open.

Postphenomenological analyses also show how tools relate to society and the culture in which they are used. For example, Verbeek (2005, 2011) argues that technologies not only mediate action and perception, but also influence the moral space of human culture. Tools modify the experience of individual actions and influence which actions are considered morally acceptable and how responsibility is distributed within society.

Similarly, Rosenberger (2017) deals with technological mediations in the specific social contexts of their use. He emphasizes that the role of tools in society depends on the cultural environment and social practices that can influence how people learn to use tools and how they are subsequently integrated into human activities. Each tool allows for multiple uses, and specific practices are always based on social and cultural contexts.

A Functional Typology of Technological Development

In the following text we map the historical development of tools and technologies in general and divide it into six stages, which can also serve as a basis for the classification of current tools and technologies.

There have been a number of attempts in the history of philosophy to structure history in a similar way, but the distinction we propose below is based on the degree and nature of mediation that technology provides between humans and the world rather than other factors, and allows us to understand the breakthrough significance of AI technology.

For example, Comte (1830) in his famous distinction of the three stages was primarily focused on the stages of scientific knowledge, based on the way natural processes are explained and what role supernatural divine forces play in this explanation. For Marx (1990), the historical development of the social order, which was correlated with the modes of production, was of primary interest. Mumford (1934) distinguished the phases of the development of society according to the dominantly used materials and energy sources into eotechnic, paleotechnic and neotechnic phases. In contrast, McLuhan (2002) and similarly Ong (2012) divided technological development on the basis of the media and communication technologies used. Closer to our approach is the concept of Nolan and Lenski (2009), who distinguish phases of societal development by production techniques into Hunting and Gathering Societies, Horticultural and Pastoral Societies, Agrarian Societies, Industrial Societies and Post-industrial Societies. In their approach, the type of technology used already plays a more important role although here again it is significantly linked to the nature of the society. Similarly, Toffler (1990) distinguishes three waves of technological development: agricultural revolution, industrial revolution and post-industrial or information revolution.

In our distinction, we start strictly from technological leaps, which are based on an understanding of technology as a generalized form of human tool, and in this understanding of technology we draw primarily on the aforementioned post-phenomenological analyses of Ihde (1979), which draws primarily on Husserl and Heidegger.

Although these stages are presented in the following text as a historical sequence of development, they are rather historically emerged forms of technological mediation that can continue to coexist in contemporary society without excluding each other.

The order of the stages is not determined by moral considerations or any notions of ethical inferiority or superiority. Their order is based on when they appear in human history. The evaluative dimension contained in the typology relates exclusively to the type of technological mediation. Later stages surpass previous stages primarily in the sense that they further expand human capabilities, technological means gain greater autonomy, and they further transform the epistemic conditions of human access to the world and to these tools themselves. This hierarchy is therefore more historical-ontological and epistemic than axiological or ethical.

The First Stage

The earliest, first types of technological tools were purely static in nature, they were *extensions of the limbs* with possible enhancements of the sense of touch, such as just the stick, the mallet or hammer, or today the dentist's probe.

From a historical perspective, the first documented cases of this form of technological mediation can be identified in the late Paleolithic and early Neolithic periods. During these periods, the first stone tools appeared in archaeological finds—these were clubs and simple hammers. Examples include the Oldowan and Acheulean tool industries, in which tools function as direct extensions of the limbs, primarily improving strength or reach, without the use of more complex mechanical principles (Bower 1977).

However, this type of tool is not exclusively used by humans, but is also capable of being used by great apes. For example, Köhler (1976) has shown in his experiments with chimpanzees that they are able to use this type of tool. Human use of these early tools is distinguished by its stability within early communities and integration into their shared practices. Thus, it is only with humans that we can speak of the historically oldest permanent form of technological mediation.

Already Kapp (2015, p.40) understood these tools as "organ projections" (*Organprojektion*) of human faculties, i.e. for example the hammer as an enhancement and extension of the arm, optimized for a specific purpose. Heidegger (2010, p.69) calls such a tool "ready-to-hand", the hammer is ready to be put to work, no significant cognitive work is required. We do not approach the tool as a theoretical object with properties, we understand the tool through its use, through the activity we perform with it, that is, through πρᾶξις (*praxis*). When we use a hammer, we do not think about it, we take it in our hand and use it. The "ready-to-hand" tools are perceived in the activity as an extension of our body, they are directly integrated into the activity.

But we can approach the hammer in a second way, which Heidegger calls "present-at-hand"—in which case we stop, for example when the hammer breaks, and look at it theoretically as an object, rather than seeing it merely as a tool. In this stance, we try to understand the hammer through intellectual analysis and as such it is not available to animals. The essential point is that what makes an object a tool or technology is not only its properties, but also people's attitude toward that object.

The Second Stage

The second stage in the development of tools is represented by simple, non-powered mechanisms, such as a lever, sometimes consisting of several structural elements, such as a pulley.

Most are *simple machines* whose basic functional principle is that they *change the direction or magnitude of a force* based on simple physical principles. More complex mechanisms may involve, for example, interlocking gears. Their systematic use requires cognitive power that only humans have among animals. Most of these machines are not even capable of systematic use by apes. Chimpanzees, for example, are able to use a rod as a lever in some experimental settings, but this appears to be the result of random experiments rather than a systematic understanding of the principle of lever operation (Malherbe et al. 2024).

From a historical perspective, the emergence of this stage can be traced back to the beginning of the Neolithic period, to the development of the first agrarian communities in classical antiquity. At that time, simple machines were not only used in practice, but were also described theoretically in the form of early mechanical theories, e.g., in Greek engineering (Koloski-Ostrow 2008). Although these technologies do not use an engine, they require an abstract understanding of the mechanical principles of operation and application of force, which indicates a fundamental cognitive and cultural shift beyond technical mediation in the form of mere physical extension of the limbs.

The use of even the simplest mechanisms such as a lever is cognitively demanding. This is shown by the classic balance scale experiment proposed by Siegler (1976, p.481). In this experiment, children's understanding of proportional reasoning and their problem-solving strategies regarding the magnitude of force are studied. During the experiment, children are presented with a balance scale with different weights on either side and are asked to predict whether the scale will flatten or tilt to one side when released. More complex configurations of weights require more sophisticated reasoning strategies and the formulation of more complex rules. Siegler identified four stages of reasoning and rule complexity that children are capable of as they increase in age: stage 1 (approx. 4–5 years), stage 2 (approx. 6–7 years), stage 3 (approx. 8–9 years) and stage 4 (approx. 10+ years). This demonstrates that effective use of even simple machines has a non-negligible cognitive demand.

This experiment shows that reasoning regarding simple machines and other stage two tools requires significantly higher cognitive abilities than the first type of tools. A lever is a simple machine that operates in a similar way to Siegler's balance scale; it is a bar used in a specific way. This example also demonstrates that it is not just the tool itself that is important, but also its use—using the rod as a lever, using redirection of force, makes it a second stage tool.

At the same time, these are still often extensions of our limbs, especially in the case of simple machines like levers or pulleys. More complex mechanisms that require gears still often have to be driven by our own power, but their integration with our body is lower. However, they still have affordances and are seen as something in order to.

The Third Stage

The third stage of tool development comes when some *independent source of motion*, electric or otherwise, is available. While a pulley or lever can only increase the force by using the basic laws of physics, with the use of an additional source of power it is possible to increase the mediating effect far beyond simple machines.

Although the first machines using water power were already in use in ancient times (Wilson 2002), historically we can generally locate this stage in terms of its broader social scope and impact at the beginning of the modern era and the Industrial Revolution (Ashton 1997). From the 18th century onwards, other external energy sources were systematically exploited, notably steam, followed by electricity and combustion engines. This transformation fundamentally changed the way and extent to which human activity affects the world. New types of machines no longer merely redirected human physical strength based on the laws of mechanics, but significantly amplified it, both in terms of its power and duration, using an external energy source. This change enabled forms of technological mediation that significantly exceeded the limits of natural human muscle power, yet remained entirely controlled by human will in every movement.

Examples include steam engines, locomotives, diesel-powered construction machinery, such as cranes, but also motorcycles or automobiles, for example. The word "automobile" is formed from the Greek word αὐτός (*autos*) meaning "self" and the Latin word *mobilis* meaning "mobile". They are therefore machines that have their own source of movement within them, self-propelled machines. However, they have no logic or will of their own; the self-source of motion is only intended to reinforce the intention and action of the human being. This is most evident in machines such as the construction crane, but it also applies to the traditional automobile (not modern automobiles equipped with sophisticated electronics or even artificial intelligence), as the latter only simplifies and accelerates human movement/running. In a traditional car, the gears are mechanical, the machine only supplies the extra force that sets the whole thing in motion. No inherent logic of the machine enters into controlling direction or speed.

Even in the case of machines like the crane, what was stated above in the case of the dentist and his probe applies. The centre of attention of the crane operator is the point where the object being carried is suspended. The experienced crane operator does not control the crane lever, but the crane arm and its suspension. The point where his hands touch the crane control levers is completely out of his attention, just as a walking man is unaware of the contraction of the thigh muscle. The crane is an extension of the crane operator's limbs, part of his extended self.

This is similar in the case of machines that enable us to move faster, such as automobiles. Merleau-Ponty (2005, p.165) gives the following example, "If I am in the habit of driving a car, I enter a narrow opening and see that I can «get through» without comparing the width of the opening with that of the wings, just as I go through a doorway without checking the width of the doorway against that of my body." (see also Grünbaum 1930).

The Fourth Stage

The fourth stage of instrument development consists of integrating the actual (usually electronic) logic into the instrument. The tool is no longer a mere mechanical mediator transmitting directly the user's intention, but can actively participate in the transformation mediating between man and the world.

These types of instruments have a character that the ancient Greeks referred to with the word αὐτόματον (*automaton*)—they act on the basis of their own will, a logic that is, however, very simple and straightforward. For example, Homer in *The Iliad* (Homer 2007, 5.749, 18.376) speaks of automatic gates or automatic movement of tripods intended for the gods.

Historically, however, this stage appeared much later, with the development of cybernetics in the mid-twentieth century and the development of analog and, in particular, digital computing technology. Automated machines using electronic control systems represent another form of technological mediation, where each movement is no longer controlled directly by human will, but where a human-independent control system is built directly into the machine itself.

Such a control system can take many forms, and these approaches are generally referred to as symbolic artificial intelligence. They include if-then rule systems and expert systems, logic-based knowledge representation methods (e.g., Prolog), formal ontologies, and a whole range of others (Russell and Norvig 2016). These systems can function independently of humans in the long term, learn, and process new information. Their activity is still fully determined by epistemically transparent deterministic operations in quantities that are comprehensible to humans and is therefore fully predictable and interpretable.

A modern example is an industrial robot, which works autonomously but does not require the constant human presence, instead having a fixed program that precisely defines each of its repetitive movements. The essential point is that the resulting action of the tool is determined by a clearly predictable output of the system, based on a clear, precisely defined sequence of rules, computational instructions, or straightforward connections of logic circuits.

In this case, technology becomes partially independent of humans—humans enter their goals into the machine in the form of instructions and settings, and the technology then works independently, performing preset operations and activities in a given order. The function of the machine is completely epistemically transparent; every movement and action has a clearly defined cause that can be identified at the level of a specific operation, which has a clear interpretation and behind which there is some intention of the human author. These systems are completely heteronomous, controlled by external input instructions, and have no autonomy. Such a machine can be very complex, with large sets of instructions and configuration data. For example, there are plans to create Robot Scientists that would be capable of automating all aspects of the scientific discovery process (Sparkes et al. 2010). However, the machine is still understood as a tool—something in order to—but affordances may be lacking, for example, in an industrial robot.

When we talk about independence from humans, we must distinguish between two forms of independence: operational/functional autonomy—the ability to act without direct human intervention; socio-epistemic autonomy—independence from human practices, cultural and social norms, and the resulting interpretive structures. In this case, we are referring to independence in the first sense. The concept of technological independence described in this stage of the analysis cannot be understood as separating the tool from human social contexts. It is operational autonomy in the narrow sense, meaning that the system can function over longer periods of time based on a set of rules and change its operations according to them without direct human supervision and active intervention. This form of autonomy is embedded in socio-cultural frameworks created by humans, particularly through decisions that are reflected in the very design of the sequence of rules according to which the machine operates.

Research in the field of the philosophy and sociology of science and technology (STS) shows that even very complex computing systems remain embedded in human socio-cultural practices and institutional structures. For example, Leonelli (2016) showed that both computational systems and the data used in them are linked to a number of human practices (annotation, categorization, interpretation, data management) and are therefore not solely dependent on the functioning of the technology itself. However, recognizing this social embeddedness does not conflict with the fact of increased operational autonomy and independence from humans. On the contrary, it clarifies that operational autonomy at the level of machine functioning and its individual activities can be accompanied by increased interconnection with human practices at the socio-epistemic level.

The Fifth Stage

The fifth and currently highest level of tool development is represented by technologies equipped with advanced artificial intelligence based on artificial neural networks. These incorporate a high degree of autonomy and unpredictability. This is also a fundamental difference distinguishing this degree of technological mediation—it is epistemically opaque, meaning that the activity of these machines is not only independent, but also impossible to fully interpret and understand—these systems are often referred to as black boxes.

Historically, we can locate this stage at the turn of the 20th and 21st centuries, and it is closely linked to the development of deep neural network technology and machine learning architectures based on it. The very concept of artificial neural networks originated in the mid-20th century and developed more or less in parallel with symbolic approaches. Artificial neural networks (and connectionist models) are therefore neither historically nor conceptually newer than symbolic systems, but they were limited by theoretical and practical constraints. Theoretical limitations were gradually overcome from the 1980s onwards, when the mathematical foundations of modern neural networks were laid. Subsequently, advances in the availability of raw computing power and reductions in the cost of computer hardware made it possible to build large artificial neural networks of previously unimaginable complexity, with trillions of connections (parameters) between nodes, enabling the current development of artificial intelligence based on these approaches.

In traditional terminology, the term “artificial intelligence” includes both older symbolic (rule-based and similar) systems and new systems based on large artificial neural networks. According to our classification, symbolic systems belong to the fourth tool class, while advanced systems based on artificial neural networks belong to the fifth tool class. The use of the term artificial intelligence for such a wide range of systems largely obscures how significant shift in technology the introduction of large artificial neural networks represents.

Although, at a practical level, current artificial neural network-based artificial intelligence systems are programmed using standard deterministic programming languages and algorithms, yet, as a result of the vast quantity of artificial neurons and their interconnections (in the trillions) and learning capabilities, the outputs of these systems are often inexplicable in the standard way used in the case of symbolic artificial intelligence.

Artificial intelligence is usually divided into weak and strong (Russell and Norvig 2016, p.1020). The term Weak AI refers to systems focused on solving a single narrowly defined task, such as driving a vehicle, predicting the weather, or controlling the electricity supply. Conversely, the term Strong AI refers to hypothetical systems that solve complex tasks requiring the use of intelligence as well as or better than a human, but are also general-purpose, i.e., can solve the same breadth of task types as a human, without the need for prior learning of individual narrowly defined task types. Strong artificial intelligence is also referred to as Artificial General Intelligence (AGI). Hypothetical, also not yet existing, artificial intelligence with significantly higher capabilities than humans is referred to as Artificial Superintelligence (Bostrom 2014).

Previously, the Turing test was considered the criterion for achieving advanced artificial intelligence (Turing 1950). The current most advanced artificial intelligence systems—large language models—have successfully passed this test (Biever 2023; Jones and Bergen 2025). However, if we classify large language models according to distinction of strong and weak intelligence, the consensus is that large language models are a very advanced form of weak AI that is focused on processing and producing high quality textual output. Large language models are based on probability theory. The first step in their use is to process extremely large sets of textual data from which probabilistic information about the relationships between words and expressions are extracted, and these are then encoded into the parameters of an artificial neural network. The resulting large language model can then be used to construct meaningful language outputs based on the input texts.

Thus, language models predict the likelihood of the occurrence of a subsequent linguistic expression on the basis of textual (context) in the range of up to several thousand words. An artificial neural network successful in this prediction typically has a range of several billion parameters.

The well-known application of large-scale language models is in popular chat applications, however, this technology has been applied in a large number of other applications. Chat programs (e.g., OpenAI ChatGPT) allow to generate coherent, meaningful and contextually relevant answers to the questions asked and to have a meaningful conversation. Large language models are based on probability theory (Brown et al. 2020), but they are able to synthesize texts in an innovative way that produces meaningful results. However, large language models do not understand language in the same sense that humans do. Some authors refer to them as stochastic parrots that merely repeat text with probabilistic word combinations without understanding its meaning (Bender et al. 2021).

Although the capability of large language models is based on probability the resulting text is not a mere rearrangement of the input data or a copy of it, but shows a high degree of originality and competence (Floridi 2023). The use of probability and artificial neural network technology does not diminish the capabilities of large language models, which is why other authors reject this comparison to a parrot (Arkoudas 2023). All these factors are the reason why large language models are usually classified as very advanced weak artificial intelligence.

At the same time, it should be emphasized that large language models lack the subjective conscious phenomenal experience, semantic grounding in experienced reality, and intentionality of the human kind that are typical of human thought and cognition and thus of the human way of being. Unlike humans, language models do not have a subjective "existence in the world"; they are not "at home" in the world (Pacovská 2024). The general consensus is that large language models do not have consciousness (Goertzel 2023; Vervoort, Mizyakov, and Ugleva 2023).

The question is to what extent can we consider large language models as tools. For example, Hongladarom and van der Vaeren (2024) say that ChatGPT goes beyond the classical understanding of the technological mediation of reality to the human, arguing that it acts as a kind of hermeneutic agent. In this sense they claim that ChatGPT can to some extent "perform the human". Similarly, Thai Vo-Nhu (2025) rejects an instrumentalist approach to this technology, arguing that it cannot be considered just another neutral tool.

In this sense, large language models are transcending the capabilities of traditional technologies and reaching a new level of autonomy and independence from humans. While an industrial robot following precise rules can also work independently, its function is limited by predefined deterministic procedures that are precisely predicted and planned by humans. A traditional industrial robot is a tool that is several stages more advanced than a hammer, but its function is based on human planning. Man has pre-planned and thought through every aspect of its function in detail in his imagination, and every movement somehow accomplishes the purpose for which it was built and programmed. The traditional industrial robot and other similar technologies that belong to the fourth stage of development are merely tools working alone, their function fully subordinated to human goals, planning and imagination.

However, systems based on artificial neural networks go far beyond the functioning of rule-based systems or straightforward algorithms. Their behavior is neither fully predictable nor plannable. We can design and train an artificial neural network with a goal in mind, but the resulting behavior may always exhibit unexpected anomalies of deviation, sometimes harmful, but at other times highly original and beneficial. How it internally arrives at its decisions is opaque, often referred to as black-box systems or unexplainable artificial intelligence.

Despite their sometimes outstanding performance, current systems based on large neural networks have a number of shortcomings. For example, large language models are trained on large volumes of text data and therefore adopt and sometimes even amplify social biases contained in the training data (Bender et al. 2021; Weidinger et al. 2021). Large language models also tend to generate credible-looking but factually incorrect outputs—so-called hallucinations—in certain contexts, which stem from their reliance on probabilistic calculations rather than actual understanding of content (Marcus and Davis 2020; Ji et al. 2023).

Another problem is that how they internally arrive at their decisions is opaque, often referred to as black box systems or unexplainable artificial intelligence. The characterization of these systems as black boxes does not refer to a single phenomenon, but encompasses several dimensions of opacity. First, we are talking about technical opacity, which is due to the quantitative scope of the models, as they involve trillions of connections between nodes, making it impossible to directly control the model. Second, we are talking about epistemic opacity, where it is not possible to fully reconstruct the semantic causal paths leading from the input to the output, i.e., to explain and interpret why the neural network produced the output it did. Thirdly, we talk about phenomenological opacity—a neural network is not a transparent mediator of human-defined intentions, but works as an autonomous machine that generates outputs based on internal logic, which, although created in a human-defined learning process, is inaccessible to human interpretation.

These types of opacity have been comprehensively analyzed in the literature. Pasquale (2016) and O'Neil (2016) focus on the social risks of opacity. Zerilli et al. (2019) deal with epistemic opacity and the difference between explainability and understanding, and Rudin (2019) calls for the use of information systems whose decisions can be fully interpreted. At the same time, opacity motivates extensive research initiatives aimed at making the results of these systems easier to interpret and explain—explainability is one of the main areas of research in artificial neural networks (Gilpin et al. 2018).

In general, epistemic opacity is considered a fundamentally new characteristic of these systems, which is one of the reasons for their inclusion in a separate stage of technological mediation development.

Can we still consider such epistemically opaque systems to be tools? Rather, we can compare the relationship between humans and modern artificial intelligence based on artificial neural networks to the relationship between humans and dogs (Sullins 2011, p.152). Humans have been breeding dogs for their purposes for thousands of years, and if by technology we mean the manipulation of nature to achieve human goals, then according to Sullins, domesticated dogs also fall under technology. Yet dogs have a certain degree of natural intelligence and we systematically train them to improve their abilities in tasks we define. Such a task can be both attacking another human and helping the visually impaired. Dogs' behavior is not always completely predictable, and it is not possible to train a dog to completely control and predict all aspects of its behavior. The dog is always a partially autonomous, individual being. Modern artificial intelligence systems based on artificial neural networks, or in the future robots that will incorporate such systems, are much more like dogs than older industrial robots. At the same time, similar to dogs in the past, robots based on artificial intelligence may act as assistants and guides for people in the future, which is why some authors demand that these technologies be generally available (Špecián and Špeciánová 2025). We therefore list them as a fifth separate stage of technological development in our classification to highlight their uniqueness and transformative nature with respect to society.

Are these artificial intelligence systems tools? In a sense, yes—but in the same way that dogs are tools, i.e., they are trained for our purposes and our goals, which they help us achieve. Unlike the lower-level tools, they are not just a tool, but have a degree of autonomy and independence beyond that, and in this they transcend the trait of tools.

Let us now return to Merleau-Ponty's conception of the tool quoted above. For the blind man, the cane is not an object, it is not perceived; the blind man feels directly with the tip of the cane, not at the point of contact of the skin of the hand with the cane. The tip of the cane is the sensitive area with which the blind person palpates the surface in front of him. It is an extension that extends the range and active range of the blind person's touch. Can a system based on artificial neural networks become a tool in this way? Can it become an extension of the natural abilities of man? In certain cases, yes, if it fully supports the human user—we are referring to various smart assistance systems that anticipate the user's intentions and correct their direct actions to match their intentions. However, in most of their applications, such as conversational language models, their autonomous nature will often be apparent and will at least partially remove the human user from their straightforward use.

The Sixth Stage

In this context, can we think about a hypothetical sixth, even higher level of technology? Above, we made a comparison between a dog and an advanced artificial intelligence system. However, there is one important difference that we have not mentioned—the dog has a certain level of consciousness, whereas the scientific consensus for artificial intelligence systems based on artificial neural networks is that they do not possess consciousness (Goertzel 2023; Vervoort, Mizyakov, and Ugleva 2023).

Thus, as a sixth, hypothetical level of technology, we can consider a technology possessing artificial consciousness. It is important to stress again that we must distinguish the question of artificial intelligence from the question of artificial consciousness. Systems of strong artificial intelligence or even systems of artificial superintelligence may exceed humans in solving most tasks, but they may not have human-like consciousness.

We understand this stage as a borderline case, based on the conceptual structure of the proposed typology. The first five stages concern technologies that are mediating in nature—they are linked to human intentionality, but with each subsequent stage, this link loosens and the tools become ontologically independent, epistemically opaque, and transition from heteronomy to autonomy. However, these tools are still not subjects of experience; they remain unconscious technological mediators of human intentionality. If we consider the hypothetical possibility that technology could acquire artificial consciousness, we understand this as a way of exploring the extreme limits of the hierarchical scheme described above. A description from the perspective of purely instrumental mediation ceases to be adequate here, and categories associated with subjectivity are required to reflect the ethical issues associated with any entity possessing consciousness. mutual agency.

At this stage, we use the term "consciousness" in a phenomenal sense, which means that we understand it as a state in which a machine has subjective experience (Chalmers 1995) or a state of "what-it-is-like" (Nagel 1974). This is not just about the sophistication of a machine's functioning or the degree of its autonomy, but rather a hypothetical system that has conscious experience. This stage is currently speculative and is a borderline case, an ideal form, not a technological milestone. We do not prefer any of the competing theories of consciousness here—global workspace theories (Baars 1995; Dehaene 2014), higher-order theories (Rosenthal 2010), and integrated-information theory (Tononi 2008). However, the possibility of this stage assumes that consciousness is physically realizable in non-biological systems, which is not a claim accepted by all authors (Block 2009; Chalmers 2010). If it were not valid, this stage would never be realistically achievable through non-biological technologies.

The sixth stage, as a borderline case, also has a methodological role in this classification: it takes into account the difference between autonomy or unpredictability and consciousness. The fifth stage is characterized by a qualitative shift towards systems that are epistemically opaque, whose actions cannot be interpreted straightforwardly because they are based on learned patterns encoded in trillions of connections in an artificial neural network. These systems may outwardly show signs of a certain autonomy, i.e., they follow rules that they themselves have created in the process of learning from a huge amount of empirical data. This behavior may tempt some interpreters to assign agency or subjectivity in a strong sense—that is, to assign consciousness. By explicitly separating fifth-stage systems, capable of autonomous technological mediation of human intentionality, from hypothetical sixth-stage systems, capable of conscious subjectivity, we methodologically avoid confusion at this level.

Outline of the Proposed Stages

The stages described above are conceived as ideal types rather than historically accurate reconstructions of the sequence of scientific progress. The criteria distinguishing the stages are functional and epistemic, which means that some technologies may fall into borderline areas between stages. The stages are cumulative, meaning that technologies characteristic of higher stages often include attributes of lower stages. The following table clearly shows the characteristics of the stages described above.

Table 1. Stages in the development of technology

No.	Characteristics of the stage
1	Extension of limbs.
2	Simple mechanisms that change the direction or magnitude of a force based on simple physical principles.
3	Mechanisms using an independent source of motion.
4	Technology that integrates a control system, including rules, logic, or algorithms.
5	Technologies equipped with advanced artificial intelligence based on large artificial neural networks.
6	Technology with artificial consciousness.

The order in which the individual stages of technological mediation are presented reflects the increase in the scope and depth of mediation. The tools typical of the first stages of historical development directly expanded the capabilities of the human body, but they were still fully controlled by human will. Later technologies increasingly lead to the independence of tools from human will and independent activity, initially based on logic or rules, but later also on increasingly independent decision-making. This change also alters the structure of action and responsibility in ways that are ethically problematic in many cases.

This is also related to the trust with which we relate to the means of technological mediation, which is an implicit factor present in all stages of the proposed typology, a problem that has been explored by a number of authors (Lee and See 2004; Hoff and Bashir 2015). In the first forms of tool use, we can understand trust as embodied. In the case of a dentist's probe, we rely on it to faithfully transmit tactile information, so it is usable and works properly, and therefore can be trusted in the sense of a bodily extension. As a result of this implicit embodied trust, it recedes from thematic attention and is transparent in the sense explained by Ihde. Similarly, in stages two and three, trust is unquestionable, as tools and machines operate on causally epistemically transparent principles. This comprehensibility directly leads to trust, which, although lower than in the first case, is still stable. We trust a lever mechanism, a traditional mechanical machine, or a crane because their movements are predictable and understandable; in the case of incomprehensible activity, it is a malfunction that can be relatively easily diagnosed, understood, and eliminated based on causal principles.

In the case of fourth-generation technologies, which have internal logic, trust becomes more conditional. Once configured and launched, these systems can operate independently, without human intervention, but their symbolic, logical or rule-based nature ensures epistemic transparency. As in the previous case, in the event of unexpected or incomprehensible activity, it is possible to stop the machine, fully diagnose and understand the causes of the malfunction, and restore trust.

Technologies based on artificial neural networks falling within the fifth stage pose a fundamentally new problem of trust. These systems can no longer be trusted on the basis of predictability, epistemic transparency, and understanding of the causal chains of their internal functioning. Although these systems operate on causal principles, their quantitative scale (trillions of connections between nodes) makes it impossible to fully understand the causality of their actions, rendering them largely opaque to users. This and their autonomy undermine the basis on which simpler types of tools and machines were trusted as forms of extension of human intentionality. The problem of trust that arises from this situation is not only psychological, but also ontological. Such systems no longer function as technological means of mediation between people and the world, but achieve a degree of independence and autonomy as a result of which their actions can only be predicted probabilistically, and trust in them is thus considerably limited.

Conclusion

Drawing on phenomenological and postphenomenological frameworks, this paper assumes that tools and technologies are not merely neutral objects, but actively shape human perception, possibilities of action and corporeality. On this basis, we propose a typology of technologies based on the degree and nature by which technologies mediate between human intentionality and the world. This typology starts from simple hand tools and continues to contemporary artificial intelligence systems.

All stages of technological development—from early tools that exist only as extensions of limbs by means of rods, or to strengthen or refine them as in the case of levers and other mechanisms, to more complex machines that have their own power source and machines with built-in logic, to autonomous artificial intelligence systems—represent a fundamental advance in the relationship between tools and the human body.

The conclusion of our argument is that current artificial intelligence systems based on large artificial neural networks are not just an improvement on older, symbolic systems, but are a qualitatively new form of technology that is partially autonomous, fully unpredictable, and largely opaque, and thus to some extent uncontrollable by humans. In this sense, these AI systems go beyond the classical conception of tools as extensions of the human body or intentionality and force us to rethink existing conceptual categories.

At the same time, these systems do not yet possess consciousness and thus do not represent conscious agents comparable to humans. The hypothetical emergence of a technology equipped with artificial consciousness—if it ever occurs—would be another major ontological breakthrough, and we would understand such a technology as a sixth, fundamentally different type of technology, with its emergence leading to major philosophical implications.

Acknowledgement

The research reported in this paper has been supported by The Czech Science Foundation (GACR) grant no. 24-11697S.

References

Aristotle. 1926. *Nicomachean Ethics.* Translated by H. Rackham. Cambridge: Harvard University Press.

Arkoudas, Konstantine. 2023. ChatGPT is no Stochastic Parrot. But it also Claims that 1 is Greater than 1. *Philosophy & Technology* 36 (3): 54. https://doi.org/10.1007/s13347-023-00619-6.

Ashton, Thomas S. 1997. *The Industrial Revolution 1760-1830.* Oxford: Oxford University PressNew York, NY. https://doi.org/10.1093/oso/9780192892898.001.0001.

Baars, Bernard. 1995. *A Cognitive Theory of Consciousness.* Cambridge: Cambridge University Press.

Bender, Emily M., Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency,* 610–623. Virtual Event Canada: ACM. https://doi.org/10.1145/3442188.3445922.

Biever, Celeste. 2023. ChatGPT broke the Turing test — the race is on for new ways to assess AI. *Nature* 619 (7971): 686–689. https://doi.org/10.1038/d41586-023-02361-7.

Block, Ned. 2009. Comparing the Major Theories of Consciousness. In *The Cognitive Neurosciences,* 4th ed., edited by Michael S. Gazzaniga. Cambridge: MIT Press. https://doi.org/10.7551/mitpress/8029.003.0099.

Bostrom, Nick. 2014. *Superintelligence: paths, dangers, strategies.* Oxford: Oxford University Press.

Bower, John R. F. 1977. Attributes of Oldowan and Lower Acheulean Tools: "Tradition" and Design in the Early Lower Paleolithic. *The South African Archaeological Bulletin* 32:113–126. https://doi.org/10.2307/3888658.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, et al. 2020. *Language Models are Few-Shot Learners.* Version Number: 4. https://doi.org/10.48550/ARXIV.2005.14165.

Chalmers, David. 1995. Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies* 2 (3): 200–219.

Chalmers, David. 2010. The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies* 17 (9-10): 7–65.

Comte, Auguste. 1830. *Cours de Philosophie Positive.* Paris: Bachelier.

Dehaene, Stanislas. 2014. *Consciousness and the brain: deciphering how the brain codes our thoughts.* New York: Penguin Books.

Floridi, Luciano. 2023. AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models. *Philosophy & Technology* 36 (1): 15. https://doi.org/10.1007/s13347-023-00621-y.

Gibson, James J. 2015. *The Ecological Approach to Visual Perception: Classic Edition.* Hove: Psychology Press. https://doi.org/10.4324/9781315740218.

Gilpin, Leilani H., David Bau, Ben Z. Yuan, Ayesha Bajwa, Michael Specter, and Lalana Kagal. 2018. Explaining Explanations: An Overview of Interpretability of Machine Learning. In *2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA),* 80–89. Turin, Italy: IEEE. https://doi.org/10.1109/DSAA.2018.00018.

Goertzel, Ben. 2023. *Generative AI vs. AGI: The Cognitive Strengths and Weaknesses of Modern LLMs.* Version Number: 1. https://doi.org/10.48550/ARXIV.2309.10371.

Grünbaum, Abraham Anton. 1930. Aphasie und Motorik. *Zeitschrift für die gesamte Neurologie und Psychiatrie* 130 (1): 385–412.

Heidegger, Martin. 2010. *Being and time.* Translated by J. Stambaugh. Albany: State University of New York Press, SUNY Press.

Hoff, Kevin Anthony, and Masooda Bashir. 2015. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57 (3): 407–434. https://doi.org/10.1177/0018720814547570.

Homer. 2007. *The Iliad.* 2nd ed. Translated by I.C. Johnston. Arlington: Richer Resources Publications.

Hongladarom, Soraj, and Auriane Van Der Vaeren. 2024. ChatGPT, Postphenomenology, and the Human-Technology-Nature Relations. *Journal of Human-Technology Relations* 2. https://doi.org/10.59490/jhtr.2024.2.7386.

Ihde, Don. 1979. *Technics and praxis. Boston studies in the philosophy of science.* Dordrecht; Holland; Boston: D. Reidel Publishing Company.

James, Ian. 2019. Tekhne. In *Oxford Research Encyclopedia of Literature.* Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780190201098.013.121.

Ji, Ziwei, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. Survey of Hallucination in Natural Language Generation. *ACM Computing Surveys* 55 (12): 1–38. https://doi.org/10.1145/3571730.

Jones, Cameron R., and Benjamin K. Bergen. 2025. *Large Language Models Pass the Turing Test.* Version Number: 1. https://doi.org/10.48550/ARXIV.2503.23674.

Kapp, Ernst. 2015. *Grundlinien einer Philosophie der Technik: zur Entstehungsgeschichte der Kultur aus neuen Gesichtspunkten.* Hamburg: F. Meiner.

Köhler, Wolfgang. 1976. *The mentality of apes.* New York: Liveright.

Koloski-Ostrow, Ann Olga. 2008. New Perspectives on Ancient Technology and Engineering in Greece and Rome. *Technology and Culture* 49 (4): 1025–1030.

Lee, John D., and K. A. See. 2004. Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46 (1): 50–80. https://doi.org/10.1518/hfes.46.1.50_30392.

Leonelli, Sabina. 2016. *Data-Centric Biology: A Philosophical Study.* Chicago: University of Chicago Press.

Malherbe, Mathieu, Liran Samuni, Sonja J. Ebel, Kathrin S. Kopp, Catherine Crockford, and Roman M. Wittig. 2024. Protracted development of stick tool use skills extends into adulthood in wild western chimpanzees. Edited by Frans B. M. De Waal. *PLOS Biology* 22 (5): e3002609. https://doi.org/10.1371/journal.pbio.3002609.

Marcus, Gary, and Ernest Davis. 2020. GPT-3, Bloviator: OpenAI's language generator has no idea what it's talking about. *Technology Review* 294. https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion accessed January 9, 2026.

Marx, Karl. 1990. *Capital: a critique of political economy, Volume 1.* Edited by B. Fowkes. London; New York: Penguin Books in association with New Left Review.

McLuhan, Marshall. 2002. *Understanding media: the extensions of man.* Cambridge: MIT Press.

Merleau-Ponty, Maurice. 2005. *Phenomenology of perception.* Edited by C. Smith. London; New York: Routledge.

Mumford, Lewis. 1934. *Technics and Civilization.* London: Routledge & Kegan Paul PLC.

Nagel, Thomas. 1974. What Is It Like to Be a Bat? *The Philosophical Review* 83 (4): 435–450. https://doi.org/10.2307/2183914.

Vo-Nhu, Thai. 2025. *The Thief of Virtue: "AI slop" is more than just bad content.* https://blog.apaonline.org/2025/12/24/the-thief-of-virtue-ai-slop-is-more-than-just-bad-content accessed January 10, 2026.

Nolan, Patrick, and Gerhard Lenski. 2009. *Human societies: an introduction to macrosociology.* Boulder: Paradigm Publishers.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.* New York: Crown Publishing Group.

Ong, Walter J. 2012. *Orality and literacy: the technologizing of the word. 30th anniversary ed.* London: Routledge.

Pacovská, Kamila. 2024. Being at Home in the World. In *Platonism,* edited by Herbert Hrachovec and Jakub Mácha, 247–268. Berlin: De Gruyter. https://doi.org/10.1515/9783111386294-016.

Pasquale, Frank. 2016. *The black box society: the secret algorithms that control money and information.* Cambridge: Harvard University Press.

Rosenberger, Robert. 2017. *Callous Objects: Designs Against the Homeless.* Minneapolis: University of Minnesota Press.

Rosenthal, David. 2010. *Consciousness and Mind.* Oxford: Clarendon Press.

Rudin, Cynthia. 2019. *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead.* Publisher: arXiv Version Number: 3. https://arxiv.org/abs/1811.10154 accessed March 14, 2026.

Russell, Stuart J., and Peter Norvig. 2016. *Artificial intelligence: a modern approach.* Pearson Education.

Siegler, Robert S. 1976. Three aspects of cognitive development. *Cognitive Psychology* 8 (4): 481–520. https://doi.org/10.1016/0010-0285(76)90016-5.

Sparkes, Andrew, Wayne Aubrey, Emma Byrne, Amanda Clare, Muhammed N Khan, Maria Liakata, Magdalena Markham, et al. 2010. Towards Robot Scientists for autonomous scientific discovery. *Automated Experimentation* 2 (1): 1. https://doi.org/10.1186/1759-4499-2-1.

Špecián, Petr, and Jitka Špeciánová. 2025. Universal Basic AI Access: Countering the Digital Divide. *Acta Informatica Pragensia* 14 (2): 272–281. https://doi.org/10.18267/j.aip.270.

Sullins, John P. 2011. When Is a Robot a Moral Agent? In *Machine ethics,* edited by M. Anderson and S.L. Anderson, 151–161. Cambridge: Cambridge University Press.

Toffler, Alvin. 1990. *The third wave.* New York: Bantam Books.

Tononi, Giulio. 2008. Consciousness as Integrated Information: a Provisional Manifesto. *The Biological Bulletin* 215 (3): 216–242. https://doi.org/10.2307/25470707.

Turing, Alan M. 1950. Computing Machinery and Intelligence. *Mind* 59 (236): 433–460.

Verbeek, Peter-Paul. 2005. *What things do: Philosophical reflections on technology, agency, and design.* Pennsylvania State University Press.

Verbeek, Peter-Paul. 2011. *Moralizing technology: understanding and designing the morality of things.* Chicago: The University of Chicago press.

Vervoort, Louis, Vitaliy Mizyakov, and Anastasia Ugleva. 2023. *A criterion for Artificial General Intelligence: hypothetic-deductive reasoning, tested on ChatGPT.* Version Number: 1. https://doi.org/10.48550/ARXIV.2308.02950.

Weidinger, Laura, John Mellor, Maribeth Rauh, Conor Griffin, Jonathan Uesato, Po-Sen Huang, Myra Cheng, et al. 2021. *Ethical and social risks of harm from Language Models.* Version Number: 1. https://doi.org/10.48550/ARXIV.2112.04359.

Wilson, Andrew. 2002. Machines, Power and the Ancient Economy. *Journal of Roman Studies* 92:1–32. https://doi.org/10.2307/3184857.

Zerilli, John, Alistair Knott, James Maclaurin, and Colin Gavaghan. 2019. Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard? *Philosophy & Technology* 32 (4): 661–683. https://doi.org/10.1007/s13347-018-0330-6.