

Thinking About AI with Kant

A Philosophical Perspective on Artificial Intelligence, Social Media, and Human Autonomy

Introduction

Contemporary society possesses an ever-growing capacity to process information, perform actions, share experiences, and maintain social interactions within the same amount of time. Technological developments make it possible not only to communicate more quickly, but also to carry out an increasing number of activities simultaneously. As a result, a subjective experience may arise in which time and life themselves seem to accelerate.

This development can be understood through the concept of social acceleration, as developed, among others, by Hartmut Rosa. Technological acceleration, social acceleration, and an acceleration of the pace of life reinforce one another. Yet the increase in possibilities does not automatically lead to an intrinsically deeper relationship with the world. On the contrary, acceleration can go hand in hand with alienation and with a problematic relationship between human beings and the world.

One of the most far-reaching technological developments at present is artificial intelligence (AI). AI systems are capable of processing vast amounts of data, recognizing patterns, making predictions, and generating recommendations. Combined with social media, this creates an environment in which human preferences are not merely recorded, but also predicted, influenced, and shaped.

The philosophical question that follows is what it means when human activities are increasingly organized by systems of prediction, selection, and optimization. People train AI in order to gain more control over life, and this has consequences for thinking and behavior because AI becomes more directive and answers the question: "What will happen?" If this question is given greater weight, what does that mean for human autonomy? This leads to the further question: "What ought to happen, and what must human beings do to make it happen?"

This article examines this question from the perspective of the philosophy of Immanuel Kant. Kant is relevant here because he makes a fundamental distinction between the conditions under which knowledge and experience are possible on the one hand, and practical reason, which concerns freedom, responsibility, and morality, on the other. His philosophy makes it possible to distinguish between knowing, predicting, and making moral judgments.

The argument forms part of a broader philosophical inquiry in which Plato and Heidegger can also be employed as perspectives on artificial intelligence. In this first part, Kant is central. The intention is not to apply Kant's philosophy directly as though he had developed a contemporary theory of AI. Rather, the aim is to offer a contemporary interpretation of his thought, consciously building a bridge between his transcendental philosophy and contemporary technological developments.

The central question is:

What happens to human autonomy when systems developed by human beings become increasingly capable of predicting, optimizing, and influencing human thought and behavior?

1. Social Acceleration and the Relationship to the World

Hartmut Rosa describes modern society through the concept of social acceleration. Among other things, he distinguishes technological acceleration, the acceleration of social change, and an acceleration of the pace of life. The result is a society in which more and more activities are carried out within the same amount of time.

Acceleration, however, has a paradoxical character. Technical possibilities increase, but this does not necessarily increase the experienced quality of our relationship with the world. Rosa connects this problem with the concept of alienation. The world may become increasingly accessible, available, and manageable, while the meaningful relationship with it may actually weaken.

Opposed to this stands the concept of resonance. Resonance refers to a relationship with the world in which a person experiences being touched or affected by it while simultaneously responding to it. Such a relationship can arise in contact with nature, music, art, other people, or social and religious experiences.

This contrast between availability and resonance forms an important starting point for the analysis that follows. For what happens when technology not only increases the speed of human interaction, but also increasingly determines which information, images, people, and experiences appear before us? (Incidentally, the word availability, or standing-reserve, will also arise in the section on Heidegger.)

The question therefore shifts from technological acceleration to the conditions under which reality appears to human beings. At this point, Kant offers a fundamental perspective.

2. Kant and the Conditions of Experience

In the Critique of Pure Reason, Immanuel Kant develops a philosophy in which the conditions of human knowledge are central. His project is not primarily a description of individual objects, but an investigation into the conditions under which objective experience is possible at all.

Kant distinguishes between sensibility and understanding. Sensibility receives that which appears to us. According to Kant, space and time are not properties that we abstract from experience afterward, but a priori forms of our intuition. They constitute the conditions under which something can appear to us as an object of experience at all.

The understanding then has the task of ordering the manifold of intuition. For this it possesses pure concepts, the categories. One of these is causality. For Kant, causality is not simply an empirical pattern that we derive from reality after repeated observation. It belongs to the structures that are constitutive of possible experience.

Here an important boundary emerges. Human beings do not know reality from a standpoint outside every possible experience. We know reality as it can appear to us within the conditions of human experience. Kant therefore distinguishes between the phenomenon (*Erscheinung*) and the thing in itself (*Ding an sich*).

This boundary does not mean that empirical reality is unreal. Kant is precisely an empirical realist: objects within the experienced world are real as objects of possible experience. His point is that our knowledge is always bound to the conditions under which something can appear to us.

This idea is relevant to the contemporary debate about AI. AI systems also produce representations of the world. They process data, select patterns, and generate predictions. The question therefore concerns not only which representations a system produces, but also how such representations subsequently influence our relationship with reality.

3. From Knowing to Predicting

It is important here to distinguish between knowing, understanding, explaining, and predicting.

A system may be able to predict with great accuracy what is likely to happen on the basis of data, without this establishing that the system understands why something happens. Machine learning makes this distinction particularly clear. An algorithm can recognize statistical patterns and correlations that make it possible to predict future behavior without possessing a causal theory explaining why that behavior occurs.

Kant offers a critical perspective in this context. For Kant, the category of causality concerns the conditions under which we can understand events in the experienced world as objectively and causally connected, rather than merely correlating and predicting. Modern science has subsequently operationalized and optimized these structures on a large scale in models, measuring instruments, and predictive theories. Kant places great value precisely on the development of judgment rather than on optimization and following recommendations.

A historical development therefore becomes visible: structures that, for Kant, describe the conditions of possible experience are increasingly employed in modern science as instruments for measuring, modeling, and predicting phenomena.

This is not in itself a problem. Science and technology have been extraordinarily successful precisely because they can describe and predict phenomena. The philosophical problem arises when the predictive capacity of such systems is extended into a form of normative authority.

The question “What is likely to happen?” is, after all, not the same as the question “What ought I to do?”

4. From Science to Technology

Modern science has profoundly changed the human relationship to reality. Time, space, causality, and quantifiability are not only concepts with which phenomena are described, but have also become operational instruments.

In economics, psychology, medicine, military science, engineering, and the social sciences, models are developed to predict behavior and events. The development of digital technology has further intensified this movement. More and more aspects of human existence can be recorded as data.

A new relationship between human beings and reality therefore emerges. What previously could only be experienced or described can now increasingly be measured, stored, compared, and predicted.

AI is a particular expression of this development. AI systems can discover patterns in quantities of data that are impossible for an individual human being to survey. As a result, they can generate predictions about preferences, behavior, and probabilities.

The technical ability to predict something can then easily be connected to a next step: optimization.

If a system can predict what someone is likely to click on, it can then determine which message is best to offer. If it can predict which video is likely to be watched, it can adjust the sequence of videos. If it can predict which product is likely to be purchased, it can tailor the presentation of products accordingly.

A chain therefore emerges from prediction to influence.

This chain is not necessarily problematic. Recommendations can be useful, and technology can significantly expand human capabilities. The philosophical problem arises when this technical logic imperceptibly becomes a norm for human action.

5. Social Media and the Production of Representations

Social media provide a particularly clear example of this development. People constantly post photographs, videos, opinions, and experiences. In doing so, experiences are transformed into digital representations.

What is distinctive is that these representations do not merely report on an experience afterward. They can begin to influence the experience itself.

For example, a place may be visited not only because it is interesting, but also because it is suitable for sharing experiences through photographs and videos. An event acquires meaning not only through the fact that it is lived, but also through the possibility of representing it digitally. The representation is then evaluated by others through likes, comments, and other forms of interaction.

A feedback loop can thereby arise: experience – representation – evaluation – adjustment of behavior – new experience – new representation.

The representation thus becomes part of the reality that it originally seemed merely to record.

This also has consequences for the experience of space and time. Space can increasingly appear as a collection of locations that are available, accessible, and shareable. Time can appear as a continuous stream of moments that must be captured, shared, and relived.

The question is then no longer simply, “What am I experiencing?” but also, “How can I represent this experience, and how will this representation be received?”

A possible tension between experience and representation therefore emerges. Representation can enrich experience, but it can also begin to structure it.

6. AI and the Predictable Human

AI intensifies this development because it does not merely record what people have done; it also attempts to predict what they will do next.

For example, a recommendation system may establish that people who view certain content have an increased probability of subsequently viewing similar content. On that basis, the system can recommend new content.

The algorithm does not need to understand why someone finds something interesting. It only needs to predict sufficiently accurately which behavior is likely to follow.

Here a philosophical tension arises.

When behavior is increasingly approached as a collection of patterns that can be predicted and optimized, there is a risk that human beings will begin to see themselves as predictable objects as well.

The human being is then represented as a collection of preferences, behaviors, and probabilities: “You viewed this, so you will probably find that interesting as well.” The prediction can subsequently influence behavior. When the system measures that behavior again, it may see it as confirmation of its

earlier prediction.

A self-reinforcing cycle thereby emerges.

7. The Ouroboros: When the Model Acts Back on Reality

This phenomenon can be represented by the image of the Ouroboros, the serpent eating its own tail.

Human beings develop models in order to understand and predict reality better. They then use those models to influence behavior. The changed behavior produces new data. These data are processed by the model again.

A feedback loop thus emerges: measurement, analysis, prediction, influence, behavior, measurement, and so on.

The model consequently no longer merely describes reality, but also helps determine how that reality develops.

The system can thereby make its own predictions come true.

For example, when an algorithm predicts that someone is interested in a particular type of content and then primarily offers that content, the probability increases that the user will actually consume it. The system subsequently sees that consumption as confirmation of its original prediction.

A closed cycle can thus arise in which human preferences and algorithmic predictions continually reinforce one another.

The Ouroboros is therefore not merely a metaphor for technological repetition. It refers to the philosophical problem that human beings create a model of themselves and their behavior and are then partly shaped by that model. Human beings are, to a lesser extent, shaped by God from clay, but shape themselves.

8. Kant and the Question of Autonomy

To understand the significance of this development, we must move from Kant's theoretical philosophy to his practical philosophy.

In the Critique of Practical Reason, the central question is not what is likely to happen, but how human beings ought to act. For Kant, morality is connected with autonomy. Autonomy means that the rational human being gives himself the moral law and is not simply determined by external causes, impulses, or interests. This is a crucial distinction.

A person may know what is likely to happen and still have to decide what ought to be done. The predictive rationality of AI therefore cannot simply be equated with practical reason. An algorithm can predict which choice a person is likely to make. It may even predict which choice will be most effective or successful within a given objective. But this does not establish which choice ought morally to be made.

Here lies a fundamental difference between instrumental rationality and practical reason.

Instrumental rationality asks, "What works?"

Practical reason asks, "What ought to happen?" and "What ought I to do?"

A recommendation system may be excellent at determining which content is likely to hold a user's attention for the longest time. But this does not establish whether it is morally right to offer that content.

In this context, it is also understandable that attention is regularly given to the ethical dimensions of social media and AI.

The same applies to a system that can predict which behavioral change will be economically or politically effective. Effectiveness is not yet moral justification.

9. Sapere aude: Preserving One's Own Judgment

Kant's famous call *sapere aude* ("dare to know") can in this context be read as a call to make independent use of one's own reason.

In a contemporary interpretation, this idea can be extended to the relationship between human beings and algorithms. People can allow themselves to be supported by technological systems without outsourcing their own capacity for judgment to them.

The problem, then, is not that people use AI.

The problem arises when the prediction of a system imperceptibly takes the place of one's own judgment.

When someone thinks, "The algorithm knows what I like, so this must be valuable to me," a prediction can easily be elevated into a norm.

The step from what is probable to what is desirable is then no longer made explicit.

This is precisely where Kant is relevant. The autonomous person is not someone who must do everything independently, but someone who remains responsible for their judgment.

AI can assist in answering the question, "What is likely to happen?"

But responsibility for the question, "What ought I to do?" cannot simply be outsourced.

10. From Autonomy to Heteronomy

When the norm for action is increasingly derived from predictions, recommendations, and measurable results, there is a risk of heteronomy.

For Kant, heteronomy means that the will finds its determining ground outside itself. The human being then acts not from a moral law given by itself, but is determined by external factors.

In a digital environment, this can take on a new form. Likes, reach, recommendations, rankings, and predictions can function as external measures of success. The individual may then adapt their behavior to what the system rewards.

The logic shifts from "What do I believe I ought to do?" to "What works according to the system?" And then: "What is rewarded by others and by the system?" The normative question is thereby replaced by a functional one.

This does not mean that every social-media user acts heteronomously. The point is more a structural risk: when an environment continually provides feedback about the probability and valuation of behavior, that feedback can increasingly function as a norm for the behavior itself.

The paradox is that a technology intended to increase freedom of choice can simultaneously create an environment in which choices become increasingly predictable and influenceable.

11. The Paradox of Representation

A broader paradox therefore emerges.

AI and social media can enrich life. They make communication possible, provide access to information, and can bring people into contact with experiences, ideas, and other people who might otherwise remain beyond their reach.

At the same time, these systems can increasingly transform reality into representations that are measurable, comparable, and optimizable.

What stands out receives attention and is measured, optimized, recommended, and influenced, as indicated above, just as human behavior is.

Reality does not necessarily become poorer as a result, but it does risk being approached increasingly through what is measurable and predictable.

Here lies the connection with Rosa's analysis of social acceleration. The technological expansion of possibilities can be accompanied by a loss of resonance when the relationship with the world is organized primarily through control, availability, and prediction.

The question is therefore not only how much human beings can know, but also which form of knowing becomes decisive for their relationship with the world.

12. The Limits of Knowledge as a Protection of Freedom

Kant's limit on human knowledge can be read here in a new sense.

The limit is not merely a restriction. It can also serve as a protection against the illusion that everything relevant can be fully known, predicted, and controlled.

When human beings recognize that their theoretical knowledge is limited, space remains for another kind of relationship with reality. In Kant's practical philosophy, practical reason becomes precisely of fundamental importance. Human beings are not merely beings who perceive, explain, and predict, but also beings who act and bear responsibility.

A distinction therefore emerges between two forms of rationality.

On the one hand, there is a rationality that seeks regularity, predictability, and control.

On the other hand, there is practical rationality, which asks about norm, responsibility, and duty.

AI can considerably strengthen the first form of rationality. The philosophical question, however, is what happens when this strengthening is understood as though it also replaces the second form of rationality.

Precisely because a system can predict so much, it becomes more important to determine what prediction cannot determine.

13. The Question of Meaning

This also brings the question of identity and meaning into view.

When identity is increasingly fed back through digital representations and measurable reactions, individuals may become inclined to understand themselves through external confirmation. The number of likes, reactions, followers, or recommendations may then function as a secondary measure of meaning.

From a Kantian perspective, however, such confirmation can never constitute the basis of human dignity.

Human beings have value not because they generate a great deal of attention, receive many views, or are predicted by a system to be interesting. Human dignity is connected with the capacity to act autonomously as a rational being.

An important distinction therefore emerges between identity as an empirical profile and personal autonomy as a normative status.

An algorithm can create an increasingly refined profile of a person. It can recognize patterns in preferences, behavior, and social relationships. But no profile thereby fully captures the normative person.

The human being is not identical with the probabilities that a system can calculate about them, and should therefore not be reduced to them.

14. A Contemporary Reading of Kant

It would be incorrect to present Kant directly as a critic of AI. He wrote in an entirely different historical and technological context and, naturally, could not have foreseen today's digital society.

The value of Kant therefore lies elsewhere. His philosophy provides concepts with which a contemporary problem can be analyzed.

His distinction between sensibility, understanding, and reason makes it possible to distinguish different levels of the human relationship to reality. His analysis of causality helps us understand the difference between experience and scientific modeling. His practical philosophy then makes clear that predicting and explaining are not the same as making moral judgments.

From this interpretation, contemporary technological development can be understood as follows: Human beings develop systems that can measure, model, and predict reality with increasing precision. These systems are then used to influence human behavior. The risk is that human beings begin increasingly to understand themselves according to the categories of the system they themselves have designed.

The paradox is that human beings attempt to increase their own freedom by developing ever more knowledge and control, while at the same time running the risk of subjecting their own actions to the models and predictions that emerge from that knowledge. In this connection, we have referred above to the Ouroboros.

15. Conclusion: From Prediction to Judgment

The philosophical significance of artificial intelligence therefore lies not solely in the question of how intelligent machines become. The more important question is what happens to human thinking when machines become increasingly capable of tasks that were previously regarded as forms of human intelligence.

AI can recognize patterns, make predictions, organize information, and generate recommendations. These capabilities can support and expand human action.

But from a Kantian perspective, a fundamental distinction remains between knowing what is likely to happen and judging what ought to happen.

When the first question begins to dominate the second, a risk to human autonomy emerges. Human beings may increasingly allow their own judgment to be guided by what is measurable, predictable, and optimizable. The feedback from systems can then influence the behavior that they initially only attempted to predict. Thus an Ouroboros emerges: human beings create a model of their behavior, the model influences the behavior, and the changed behavior subsequently confirms the model.

The fundamental question is therefore not whether AI can replace human beings. The more fundamental question is whether human beings are willing to transfer parts of their own capacity for judgment to systems because those systems are better at predicting what is likely to work.

Kant offers a principled answer here: human beings are not only knowing and predicting beings, but also autonomous acting beings. The question of human freedom therefore cannot be reduced to which choice is most probable, efficient, or successful.

The question that ultimately remains is not "What will I probably do?" but "What ought I to do?"

Attention therefore shifts from prediction to responsibility, from optimization to judgment, and from heteronomy to autonomy.

In this shift perhaps lies Kant's most contemporary significance for thinking about artificial intelligence and social media.

Erik Heemskerk, August 29, 2026