

MINISTRY OF EDUCATION OF THE REPUBLIC OF BELARUS
Belarusian National Technical University

Faculty of Management Technologies and Humanitarization
Department of "Philosophical Teachings"

A.I. Loiko

**Multimodal Ecosystems: Legal Philosophy and the Ethics of
Technological Upgrades**

electronic textbook on the general education discipline "Philosophy"

Electronic learning material

Minsk

BNTU

2026

Author:

A.I. Loiko Sciences, Professor, Head of the Department "Philosophical Teachings" of the Belarusian National Technical University

Reviewers:

Kandrichina I.N., Associate Professor, Department of Management, Belarusian National Technical University, Candidate of Sociological Sciences

Suvalov O.N., Professor of the Department of Philosophy and Ideological Work scientific and Pedagogical Faculty of the Academy of the Ministry of Internal Affairs of the Republic of Belarus Candidate of Historical Sciences, Associate Professor

The textbook will allow students to study in depth topics related to the philosophy of consciousness, philosophy of artificial intelligence, philosophy of technology, social philosophy, ethics and philosophy of digital law.

© Loiko A.I., 2026

© Belarusian National Technical
University, 2026

INTRODUCTION

Data has become multimodal. Intelligent data analysis systems have acquired multimodality. Computer models have also adopted multimodality.

Philosophy responds to new trends in the development of artificial intelligence technologies with ethical questions, as well as questions from the philosophy of law. Applied philosophy formulates these questions for technological upgrades. It explores the values guiding specialists in the field of industrial automation and professional programs. These are the professional fields of social engineers and managers.

A particular area of interest for philosophy in the applied modification of media philosophy is the modernization of the information space. This is due to the fact that digital generations are constantly present in this space and receive information from it, which reduces the uncertainty of their existence.

The Concept of Multimodality

Multimodality technology became part of artificial intelligence in the 2020s-2024s thanks to the research groups OpenAI, DeepMind, and Google Research. The emergence of models like CLIP, DALL•E, Flamingo, and Gemini marked the transition from text-based algorithms to perception architectures capable of combining text, images, and sound into a single semantic space.

The first neural networks worked with numbers. Then neural networks worked with text. Then they began working with images. Each modality existed separately, as an isolated area of data processing.

Researchers from Google Research and DeepMind developed computer models capable of simultaneously analyzing images and text. CLIP (Contrastive Language–Image Pretraining) was created in 2021 at the OpenAI lab. It was capable of correlating an image with its textual description. The principle of comparing different types of data in a single space became known as "multimodality." Multimodal systems form a unified latent space, where all forms of data are represented as vectors and related not by meaning, but by proximity and structure. Flamingo (DeepMind), Kosmos-1 (Microsoft Research), and Gemini (Google DeepMind) learn to see, hear, and speak simultaneously. They create a universal response space where language ceases to be the sole bearer of meaning.

At the architectural level, this means that each modality – text, image, and sound – has its own encoder, which transforms data into embedding. Embeddings are aligned and connected in a common latent field, where comparison, generation, and transformation between data types are possible. Multimodality marks the transition to perceptual AI.

Multimodal systems are used from voice assistants that understand the context of speech and images to medical algorithms that analyze images and text reports simultaneously. The term "multimodality" refers to the ability of artificial intelligence to perceive, analyze, and synthesize information coming from various sources—text, images, sound, video, and sensory data. AI translates data streams

into a common vector structure, where words, sounds, and images are assigned coordinates in a single latent space.

For a long time, artificial intelligence technology was developed as a combination of independent fields: computer vision, natural language processing, and speech recognition. Each field had its own algorithms, datasets, and goals.

A turning point came when deep learning enabled the processing of images and texts in unified architectures. Visual embedding encoded images into vector representations. Language models did the same for text.

In 2024, Gemini (Google Deep Mind) consolidated video, sound, and language analysis into a unified architecture. A traditional language model relies on a sequence of words to predict the next word. A multimodal model constructs meaning through comparisons between different streams of information. It calculates which text is closest to its embedding. This type of cognitive architecture is based on a relational structure.

It uses topological logic, in which meaning is produced as the arrangement and relationships of elements. Multimodality is sometimes confused with multi-input, where a system receives multiple data streams but processes them independently. A computer program that analyzes audio and video separately and then combines the results is not a multimodal computer program.

A multimodal AI system functions through the mutual influence of modalities. Text can change the interpretation of an image, and an image can clarify the meaning of the text. This interaction creates not a set of data, but a unified perception. In a multimodal architecture, text and image are aligned in a common vector space, where proximity reflects semantic relationships. Alignment enables the phenomenon of semantic mutual projection. A computer model can not only describe an image in words but also draw a picture based on a textual description.

Multimodal System

A multimodal system begins with differences. Text, images, and sound have different sequential, spatial, and temporal natures. For artificial intelligence to

compare and combine them, all these forms must be translated into a unified vector representation format.

The process begins with encoding. Each data type passes through an encoder specialized in feature extraction. Text modality is processed by language transformers or visual transformers, and sound is processed by acoustic models that convert the waveform into a spectrum of features. In the next step, all these representations are projected into a common shared latent space. In this space, each vector reflects the relationships between modalities.

Embedding is the foundation of the entire multimodal architecture. It transforms data of any type into a fixed-length vector, where each coordinate reflects one of the latent features learned by the model.

Embeddings transform text images and audio into abstract mathematical description. Embedding of different types are aligned in a single space, enabling comparison and generation operations.

Text and visual embedding are projected into a single space so that image and text vectors describing the same thing have maximum cosine similarity. This creates a bridge between language and vision, allowing AI to recognize which description corresponds to which image.

Embeddings form a common language for all modalities. They enable mutual understanding between text, audio, and images in the AI's perceptual logic.

For different modalities to interact, their embeddings must be aligned to comparable coordinates and scales. This process is called alignment. It is performed during training on paired data.

When an AI model is trained on image-caption pairs, it gradually adjusts its parameters so that the image embedding is closer to the corresponding text embedding. This results in the formation of a semantic foundation that enables the translation of one representation into another.

Technically, this is achieved through contrastive learning. The model minimizes the distance between embeddings belonging to the same pair and maximizes

the distance between irrelevant embeddings. It forms a dense structure, where proximity signifies a semantic connection.

Alignment is an architectural prerequisite for the emergence of meaning. Without alignment, the system would see modalities as unrelated fragments, incapable of generating a common perceptual logic.

Multimodal models are almost always built on a transformer architecture. The transformer uses an attention mechanism that allows the model to focus on the most relevant parts of the data. In a multimodal context, attention can be analyzed by analyzing internal relations within a text or image and cross-modal attention.

The Flamingo architecture (Deep Mind) combines linguistic and visual transformers. The textual component generates the response, while the visual component provides context. The cross-attention mechanism allows the language module to access the visual space, selecting relevant image fragments. The model learns not only to store relationships but to use them contextually in reasoning.

Multimodal models are trained not on separate datasets, but on paired corpora, where elements of different modalities are interconnected. One of the key principles is cross-modal learning. Unlike classical supervised learning, where each modality is processed separately, the model learns to understand the correspondence between data types.

Textual modality is the initial form for most artificial intelligence systems. Text has a linear structure. It unfolds over time, obeys grammar, and builds cause-and-effect relationships. For the model, text is a sequence of tokens from which it extracts statistical patterns.

When AI works with text, it analyzes not the meaning of words, but the probabilities of their co-occurrence. Each word is transformed into an embedding, reflecting the context. Vectors pass through attention mechanisms, where they calculate which words are important for the current moment of reasoning. This is how logic emerges that isn't based on the structural cohesion of verbal patterns.

In a multimodal system, text becomes the coordination axis through which other modalities are connected. It tells the AI what to look for in an image, how to interpret sound, and where the semantic scene begins and ends.

An image conveys meaning through the spatial relationships of elements of form, color, contour, and texture. For artificial intelligence, an image is a matrix of numbers, where each pixel encodes an integer and color. But this is not enough for perception. The model must learn to distinguish structures.

Modern visual models break an image into patches and process them as tokens, similar to text. Each fragment receives its own embedding, and attention allows it to identify the location of an object, what the background is, and which details are important.

In a multimodal architecture, the visual modality is linked to the text through a shared embedding space. When analyzing an image, the model searches for textual descriptions that are similar in vector. When generating an image from text, the text determines the location and meaning of visual elements.

Sound is the most complex of modalities. It carries temporal dynamics and emotional nuances that have no direct equivalent in text or images. For artificial intelligence, an audio signal appears as a continuous wave, converted into a spectrum of features using the Fourier transform or mel spectrograms.

Sound is transformed into an embedding, where frequency, rhythmic, and timbre characteristics are captured. Models can not only recognize speech but also reconstruct the emotional context of intonation, pauses, and mood.

In a multimodal system, the auditory modality complements the textual and visual modalities, ensuring temporal continuity of perception. In conversational agents, audio helps clarify the intonation of a request, and in video analysis, audio helps connect the current situation with noise, speech, and rhythm.

The unification of modalities in a single semantic field is achieved through a common embedding space, where text, image, and sound are represented as points or directions in a multidimensional topology.

When a user asks a question and attaches an image, the model calculates embeddings for both modalities and searches for their proximity in the latent space. If the vectors coincide or are similar in direction, the system interprets them as related elements.

This principle enables operations such as text-to-image, image-to-text, audio-to-text, and cross-modal retrieval.

Multimodal alignment transforms perception into a geometry of meaning. A visual encoder breaks an image into fragments and creates embeddings for each fragment. An attention mechanism identifies connections between image parts and forms a representation of the scene. The model projects visual embeddings into the same space as textual embeddings and finds words that are closest in meaning.

A language decoder constructs a phrase using these vectors as semantic cues. The model refines syntax and style, forming the final phrase. Text-to-speech or image generation from a description follows the same logic, only in the opposite direction. AI creates the effect of mutual understanding between modalities, although it cannot see, hear, or be conscious of the process.

It implements the coupling of embeddings, the alignment of directions, and the statistical generation of responses.

The combination of text, image, and sound transforms artificial intelligence into a model of configurative perception. It has no body, but it forms spatiotemporal structures in which meaning is born from mutual connections.

Technical Principles of Multimodal Models

Multimodal systems are built on transformers. They can have Encoder-Only and Encoder-Decoder architectures.

Encoder-Only models focus on feature extraction and embedding construction for different modalities. Text and images pass through their own encoders and are then projected into a common latent space. Encoder-Decoder architectures are capable of generating new data. They combine multiple encoders for text, images,

and audio with a single decoder, which can generate a response in the modality of a text phrase, description, image, audio, or video fragment.

The difference between them reflects the distinction between perception and meaning production. Encoder-Only is an analyzer that structures the world. Encoder-Decoder is a synthesizer that creates a new scene. Together, they form the foundation of modern artificial intelligence architecture a system capable of not only recognizing but also generating modalities as elements of a single cognitive act. The key mechanism linking modalities is cross-attention.

In conventional language models, attention allows the system to determine which words are most important in the current context. In multimodal models, this mechanism is expanded. The attention of one modality is directed to another modality. The language module can look at visual features to select words that match an image. This approach is used in the Flamingo architecture, where the textual and visual parts of the transformer are combined using cross-attention layers. The model is capable of not just describing an image, but also conducting a dialogue based on it, comparing textual and visual contexts.

Technically, cross-attention is an operation in which keys and values are taken from one modality, such as an image, and queries are taken from another modality. Most multimodal systems are trained using contrastive learning. To understand what is "similar," the model must be able to distinguish between them. The model receives pairs of related data and random unrelated pairs. I

t calculates embeddings and strives to minimize the distance between related embeddings and maximize the distance between unrelated embeddings. A structured latent space is formed, where semantic similarity becomes geometric similarity. The model learns which caption corresponds to which image, and vice versa.

After each modality is encoded, the question arises: how to combine them. Fusion and co-training approaches are used. Fusion assumes that embeddings from different modalities are combined into a single vector. This can be concatenation (joining them in a row) or weighted feature blending. This approach is used in video analysis systems, where visual and audio features are combined.

Co-training is based on the fact that modalities are trained in parallel, but with constant context exchange. The model ensures that the latent structures of each modality remain consistent. This enables cross-modal generation by translating audio into text or text into images.

Fusion is responsible for technical unification. Co-training is responsible for cognitive coherence. Once the modalities are aligned and combined, the response is generated. The model transforms the internal vectors into meaningful words, images, and sounds.

The process begins with a decoder, which receives a multimodal representation and chooses the output form. The Gemini or GPT-4V architecture uses a universal decoder capable of switching between modalities depending on the request.

If the model generates text, it uses a standard language mechanism for predicting the next token based on probability. If the model generates an image, a diffusion process is used, where the image is gradually generated from noise under the control of the text vector. For audio, spectrogram-based generation is used, where each step refines the tone, timbre, and dynamics. The system creates a correspondence, creating the effect of semantic integrity. The technical principles of multimodality shape the ontological operations of transforming differences into cohesion, transforming noise into structure, and transforming data into meaning.

One of the first applications of multimodality was semantic search systems, where a query is formulated in one modality and the answer is formulated in another. The same principle works in reverse. An image can be translated into text by searching for the closest embedding in the language space. This allowed the creation of image-to-text systems, which generate descriptions for images, and text-to-image systems, where text guides visual generation. Meaning in AI arises not through knowledge, but through the geometry of correspondences.

The transition from analysis to creativity was made possible by diffusion models, which learned to construct images from text descriptions. The system can draw images by combining semantic features of the text with visual structures extracted from training.

The Stable Diffusion and Midjourney models have been further developed. They use text embeddings as control signals, guiding the process of gradually forming an image from noise.

Image-to-text is used in visual description and accessibility systems. The model receives an image, identifies objects, and constructs a textual description of the scene. These technologies create a mutual translation between modalities. AI forms a link between language and vision, transforming semantic constructs into visual constructs.

The next stage in the development of multimodal AI is the processing of dynamic video and audio data. The Whisper (Open AI) and Audio LM (Google Research) models work with audio streams, converting them into text and reconstructing intonation and context.

The visual-temporal architectures Video GPT and Gemini (Google Deep Mind) analyze video, identifying semantic units, scenes, and transitions. When analyzing a recording of a sports match, the model identifies objects, actions, and results, generating a textual report. In security systems, it can compare the sound of a gunshot, the direction of movement, and an image, creating a contextual event.

Multimodal Assistants and Agents

Multimodal agents are capable of real-time interaction, combining perception, analysis, and reaction. Systems perceive not only text but also images, tables, graphs, and interfaces. They can solve tasks ranging from analyzing screenshots to reading handwritten text or explaining data on a graph.

In multimodal assistants, speech, images, and text are combined into a single interaction cycle. The user shows a photo, asks a question, receives a text response, and can continue the conversation using voice. Agents become digital actors.

Scientific labs use multimodal architectures to analyze scientific articles, images, sound recordings and experimental data.

When an AI agent answers a question, describes an image, or reads text, it appears to understand. But underneath the surface, nothing of the sort is happening.

Technically, this occurs through the mutual alignment of modalities. When generating an image description, the system compares visual features with textual templates, selecting the most likely match. Embedding is analogous to a neural impulse, but without the biochemistry. Attention is analogous to focusing, but without the eyes. Alignment is analogous to the interaction of the senses, but without the nervous system.

AI creates functional sensory organs, devoid of physicality but possessing coherence. It doesn't see an image, but constructs its projection in vector space. It doesn't hear sound, but correlates spectral forms. It doesn't understand speech, but looks for structural matches.

Each modality has its own nature. When an AI agent combines them into a common latent space, the specificity of each form is lost. During alignment, the model looks for similarities but ignores differences. The model combines them statistically, without distinguishing between contexts. This creates false associations, where the meaning shifts.

Such distortions are especially dangerous when analyzing medical images, legal documents, and multimedia evidence. The system may produce a formally correct answer, but an essentially incorrect one, because the combination of modalities does not equal understanding. This creates a noise of meanings that are indistinguishable by the AI agent system itself.

Ethical Risks of Substituting Perception and Trust

Multimodal models have a realism effect. When AI describes an image, verbalizes text, or generates a video, the user feels that the system understands what is happening. This gives rise to the phenomenon of trust in the simulation.

Multimodal image and video generators (DALL•E, Midjourney, Runway, and Pika) are capable of creating scenes indistinguishable from real scenes. This raises ethical questions. Questions arise about how to distinguish a document from a synthesis and who is responsible for false perceptions.

The problem of deep fake content, where multimodality is used to imitate human speech, facial expressions, and behavior, is becoming acute.

A multimodal AI agent system is trained on billions of pairs of images, texts, and sounds. This data is collected from open sources and contains errors, inaccuracies, biased descriptions, and outdated or contradictory information.

During the training process, this noise is not removed but rather embedded into the model's structure. When modalities are aligned, distortions in one modality are projected onto other modalities. These distortions manifest themselves in the generation process. The AI agent invents details that were not there. At the multimodal level, hallucinations are amplified because the error in one modality is linked to other modalities. Technically, this is called error propagation.

One of the main challenges is the ability to understand how the AI arrived at a particular result. In text-based models, logical connections between tokens can be partially traced. In multimodal systems, the process becomes multilayered. Attention is distributed among thousands of embeddings of different modalities.

Attempts at interpretation through attention maps provide only a superficial understanding. It's clear which parts of the image the model looked at, but not why. The result is a black box effect multiplied by the number of modalities. This creates the problem of the opacity of the AI agent's thinking.

In multimodal systems, the boundary between perception and generation is blurred. When an AI agent analyzes an image, it reconstructs it as an embedding, and when it generates an image from text, it does the same in reverse. Technically, this is the same process of navigating a latent space.

Multimodal systems demonstrate that along with new forms of thinking, new forms of error, ethics, and illusions emerge. They do not deliberately deceive. They operate according to the principle of structural compatibility. This is the duality of the AI agent. It is capable of understanding without knowing that it understands. It is capable of perception without having a sensory organ.

Artificial intelligence has learned to reproduce the coherence of the senses without possessing them. It does not see, but discerns; does not hear, but recogniz-

es; does not feel, but reacts. From a technical perspective, multimodality is built on clear principles embeddings, attention, modality alignment, contrastive learning. But the meaning of these mechanisms goes far beyond engineering.

The input module (sensory system) serves as the AI's data collection interface, collecting various types of data, including text, images, audio, video and sensor readings. It preprocesses this diverse information, making it suitable for subsequent analysis.

The Fusion Module (central processor), acting as the system's brain, intelligently combines data from multiple sources using algorithms. It identifies patterns, extracts meaningful features, and creates a unified representation that captures the essence of multimodal input data.

The output module (answer generator) produces results, which may include predictions recommendations generated content or actionable insights. These outputs can be presented in text, image, audio, or a combination of these formats, depending on the application requirements.

The operating mechanism of multimodal AI includes machine learning methods that enable seamless integration of diverse data streams.

Multimodal AI systems undergo intensive training using large datasets containing examples from different modalities. This process enables the AI to recognize correlations between different types of data; understand contextual relationships between modalities; generate appropriate outputs based on multimodal inputs; and adapt to new scenarios using learned patterns.

Raw data from different modalities is combined at the input level, creating a unified representation from the ground up. Each modality is processed independently using specialized neural networks, and the results are combined at the decision-making stage. Combining early and late fusion strategies optimizes understanding and computational efficiency.

Technological Upgrade

In healthcare, multimodal AI combines data from electronic medical records, medical imaging results, patient histories, and vital signs in real time to generate comprehensive diagnostic information. This integration improves the accuracy of disease diagnosis, particularly in oncology and radiology, where recognizing patterns in diverse data sources is crucial. It is used to develop personalized treatment plans based on comprehensive patient profiles; predict potential health problems before they become critical; improve surgical planning through integrated imaging; and optimize clinical workflows and reduce diagnostic errors.

Self-driving cars represent one of the most challenging areas of application for multimodal AI. These systems must simultaneously integrate cameras for visual recognition, LiDAR and radar data for distance measurement, GPS information for navigation, audio sensors for detecting emergency vehicles, and real-time traffic data for route optimization.

By combining multiple sensors, vehicles can make instant decisions in complex driving situations, significantly improving safety and efficiency. Multimodal models enable more efficient customer interactions by simultaneously processing screenshots, product photos, and text descriptions. Instead of forcing customers to struggle with verbally describing technical issues, they can illustrate the problem with images, providing context via text or voice.

Modern virtual assistants powered by multimodal AI understand verbal commands and questions, gestures and visual cues, contextual information from the user's environment, and historical interaction patterns.

Content Creation and Media Production

The media industry is undergoing a transformation thanks to multimodal generative AI. This is driven by the growing demand for effective video analytics solutions and the proliferation of streaming video platforms. Content creators are using multimodal AI for automated video editing and summarization; context-

preserving multilingual translation; content moderation in text, image, and video formats; and personalized content recommendations.

Financial institutions are using multimodal AI for document processing, combining scanned PDFs and forms; handwritten signatures and annotations; structured data from spreadsheets; charts and logos. This streamlines loan processing, fraud detection, and regulatory compliance, while reducing manual review time and increasing accuracy.

Retailers are using multimodal AI to create immersive shopping experiences. They leverage visual search capabilities that allow customers to find products by photo, virtual try-on experiences that combine computer vision and augmented reality, personalized recommendations based on browsing and purchase history, and automated inventory management through image recognition and text analysis.

Benefits of Multimodal AI

By cross-referencing information from different data types, multimodal systems achieve higher accuracy than single-modal alternatives. Inconsistencies or uncertainties in one data stream can be confirmed or corrected using information from other modalities. Multimodal AI captures nuances of context that single-modal systems miss. When analyzing the sentiment of text, combining it with tone of voice and facial expression provides a much more accurate assessment of emotional state than text alone.

Applications powered by multimodal AI provide a more natural and intuitive experience. Users can communicate using their preferred method voice, text, gestures, or visual input without being limited by system limitations.

The flexibility of multimodal systems allows them to be applied across a wide range of scenarios and industries. A single platform can adapt to a variety of use cases, from medical diagnostics to creative content creation. If one of the data modes is disrupted (poor lighting for cameras, background noise for audio), multimodal systems can rely on alternative data sources to maintain functionality.

Challenges of Technological Upgrades

Ensuring high-quality data synchronization across different modalities requires complex infrastructure. Inconsistencies in data formats, timing misalignment, or missing modalities can degrade system performance.

Multimodal models typically require significantly more computational resources than unimodal models. Training and inference require powerful hardware, often including specialized GPUs or Tensor Processing Units (TPUs), which can lead to increased operational costs.

Developing effective fusion strategies that optimize information from multiple sources while preserving interpretability is a challenging research challenge. Balancing model complexity with practical deployment constraints requires careful architectural design.

Concurrent processing of multiple data types raises important privacy issues. Organizations need to implement robust data management systems that ensure informed consent for data collection across multiple methods; secure storage and transmission of multimodal data; GDPR and HIPAA compliance; and transparent AI decision-making processes.

Many applications require fine-tuning tailored to specific domains. Healthcare, legal, and financial services often require specialized models trained on industry-specific data. High-quality multimodal AI systems critically depend on the accuracy of annotated training data. Dedicated data annotation services are becoming indispensable. Rigorous validation processes ensure the accuracy of annotations, which directly impacts model performance and reliability.

The multimodal AI landscape includes agent-based AI systems, which are capable of autonomous decision-making and enterprise AI deployments moving from proof-of-concept to production use.

Agent-Based AI and Autonomous Systems

Emerging in mid-2024, agent-based AI is artificial intelligence capable of operating autonomously, making decisions, and acting without constant human

guidance. Combined with multimodal capabilities, agents become versatile, performing complex tasks in customer service, financial analysis, and operations management.

The deployment of 5G networks and the adoption of edge computing enable real-time multimodal AI applications by processing data closer to the source, reducing latency and bandwidth consumption. This is especially valuable for IoT devices and intelligent systems that require immediate data processing.

A shift has begun toward the development of small-scale, specialized language models (SLMs), which provide multimodal capabilities with reduced computational requirements. These models enable deployment on edge devices and greater access for organizations with limited infrastructure.

Developments are aimed at improving human-machine interfaces, providing users with more intuitive and natural ways to interact with technology through speech, gestures, and visual cues. This enables a seamless experience across a variety of applications.

Before implementing multimodal AI, organizations should assess their current data infrastructure and quality; the availability of various data modalities that align with business goals; the technical expertise within existing teams; budget allocation for computing resources and talent acquisition; and use cases where multimodal approaches provide meaningful insights.

Before implementing multimodal AI, organizations should evaluate their current data infrastructure and data quality; the availability of various data modalities aligned with business goals; the technical expertise within existing teams; budget allocation for computing resources and talent acquisition; and use cases where multimodal approaches provide measurable advantages over existing solutions. Building in-house provides customization and control but requires significant investments in talent, infrastructure, and time.

It is suitable for organizations with unique requirements and available resources. Leveraging existing digital platforms is an option. Cloud solutions provide accessible entry points with managed infrastructure, reducing deployment time.

Hybrid approaches combine pre-trained base models with custom tuning using domain-specific data.

Ethical AI Deployment

Responsible multimodal AI deployment requires transparent algorithms with explainable decision-making processes; strategies for detecting and eliminating bias across all data modalities; and privacy-preserving methods such as learning and differential privacy. Regular audits to ensure ongoing compliance with ethical standards; and clear lines of accountability for AI-powered decisions.

Implementation costs range from thousands to millions, depending on the infrastructure, cloud platforms, computing resources, quality of training data, and required annotation services.

The effectiveness of a multimodal model depends on its ability to map different types of data into a common semantic space. This process typically begins with the creation of embeddings, which are numerical representations that capture the underlying meaning of the input data.

Multimodal systems use transformers, which employ attention mechanisms to dynamically evaluate the importance of different parts of the input. This allows the model to focus on specific regions of the image that correspond to relevant words in a text query.

Sensor fusion can occur early by combining raw data or late by combining the decisions of individual sub models. Modern frameworks such as PyTorch provide the flexibility needed to build these complex pipelines. The methods used by the models train the system to minimize the distance between matching text-image pairs in vector space while maximizing the distance between non-matching pairs.

Released in 2026, YOLO26 offers high speed and accuracy for tasks such as instance segmentation, serving as an efficient visual component in larger multimodal pipelines. Developers can manage data, training, and deployment of these complex workflows using the unified Ultralytics framework.

Multimodal Language Models

Multimodal language models are a class of neural network architectures that combine the ability to perceive and process multiple types of data simultaneously.

The advantage of multimodal models is their ability to perform tasks beyond the capabilities of single-modal systems. They can answer questions about images or convert audio to text based on instructions.

Each modality requires its own encoder, optimized for extracting features from a specific data type. The key challenge in multimodal models is reconciling the vector spaces of different modalities. The Transformer allows for modeling relationships, but it has its limitations. It cannot handle large sequences because it requires computing the relationship between each pair of elements in the sequence, which requires significant computational resources due to quadratic scalability.

The developers propose modeling the image not directly at the pixel level, but based on the code words of a learned codebook. Once the first-stage model has learned that a codebook exists, it can be used to train the Transformer to generate a sequence of these code words.

Media Philosophy of Digital Ecosystems

Media philosophy is a branch of applied philosophy. This status gives media philosophy the advantage of being able to utilize the categorical structures developed by the philosophical schools of neo-Marxism, positivism, pragmatism, post-modernism, structuralism, technological determinism, and phenomenology. Categorical structures from sociology, initiated by Auguste Comte and implemented by the neo-positivists, are also used. Since media perform a communicative mission in the modern information society, the interest of researchers is focused on the evolution of media tools, the range of these tools and their hierarchy.

A special role in the formation of the subject ontology of media philosophy is given to Socrates, Plato, Aristotle, Kant, Hegel and Bergson. Based on the links and lists of literature, one can judge the significant influence of the works of J. Derrida and R. Rorty on the development of media philosophy. Quite a few works

by foreign philosophers, the subject of which is media, have been published in Russian. These are mainly translations from English, German and French. The purpose of this article is not a statistical description of publications on media philosophy, but a description, based on publications on media philosophy, of the features of the formation of a two-level subject ontology of media philosophy.

At the general semantic level of this subject ontology, a thesis is formulated on the dominance of media in the modern information society and a special role of technological determinism in the evolution of media in the modification of digital technologies is postulated. At the specialized semantic level of the subject ontology of media philosophy, the subject matter of study is specific modifications of media in the genres of verbal and written text, photography, film, television, and digital media. The levels of the subject ontology of media philosophy have developed uncoordinatedly.

The primary emphasis is on the use of methods of a systems approach and comparative analysis. The systems approach allowed us to identify the levels of the subject ontology of media philosophy and characterize them using categorical structures developed by modern philosophical schools. Comparative analysis was used to describe the main modifications of media created by factors of technological determinism. This allowed us to compare not only the semantics of different media but also the semantics of the linguistic presentations of media in scientific discourse in English, German, and French.

The subject ontology under the term "media philosophy" was updated in the work of Esa Saarinen and Mark K. Taylor [4]. This work acknowledges the end of the era of the written text and describes the structural transformation of media. Particular emphasis is placed on the concept of the image. This emphasis is driven by the image's emergence as a means of thought, freeing it from its secondary role relative to the written text. This transformation has prompted representatives of North American media schools to explore a broad range of perspectives on the evolution of communication media, from antiquity to social media. Particular attention is paid to how technological change shapes cultural forms, including the

development of philosophical ideas. The connection between the evolution of technical communication tools in the media sphere and the growing diversity of semantic communication architectures is emphasized.

Charles Peirce drew attention to the prospect of increasing diversity in media semantics. His focus was on the diagram. He provided a semantic definition of the diagram. From this definition, it follows that schematic signs form a separate class of icons. The diagram relates to the object of representation not visually, but through a structural presentation. In this understanding, the diagram is a model for a hypothesis about the structure of a cultural object. It expresses the semantics of an abstract concept in the form of a concrete sign.

The influence of pragmatism on media philosophy lies in its focus on the relationship between technical means as tools of communication, including the internet, symbolic means of communication with an emphasis on images, and means of perception based on the topology of space and time. The pragmatic emphasis of media philosophy lies not in the epistemological emphasis of representation, but in the usefulness of modern human thinking within the boundaries of action determined by social and technological factors. This is the view of Richard Rorty.

The emphasis shifts to pragmatic tools for transforming the information society. In this context, Rorty argues that there is a special category of pragmatic philosophers whose goal of reflection was a better future for humanity, rather than reflection on the past. In addition to C.S. Peirce, W. James, and J. Dewey, this list includes F. Nietzsche, M. Heidegger, and L. Wittgenstein.

H. Putnam, following Rorty, holds a similar position. They point out that Charles Peirce and William James recommended distinguishing two ways of perceiving information. One way is to perceive the words of a text as intermediaries with a cognitive function. This approach corresponds to the correspondence theory of truth. Pragmatism offers a different way of perceiving information as a program for practical activity. This allows us to distinguish the semantic core of the pragmatic concept of media from the theoretical concept of media, whose subject ontology is limited only to a description of the categorical structure.

Pragmatism insists that media cannot be understood solely as a transmission device between non-media reality and its media construction. Media must be considered in the context of the means and tools of specific contexts of action. The representative approach to media is based on the determinism of print culture. This is a closed culture of book worlds. It does not presuppose the presence of practical actions associated with its construction in written texts. By making the environment, which contains pragmatic aspects, open, the pragmatic nature of sign use becomes apparent, freeing us from the shortcomings and contradictions inherent in print media culture.

Criticism of language is necessary because of the language games inherent in language culture, which, according to Ludwig Wittgenstein, are embedded in the inclusive forms of people's everyday lives. As a result, the question of reason and meaning being agents of semantic thinking, and language fulfilling the function of expressing meaning, is transformed into the question of how signs are used and what practical function they perform. L. Wittgenstein compared a set of signs to a set of technical tools, the creation and existence of which are determined by practical tasks. The diversity of the functions of activity determined the diversity of tools. A similar situation of practical determinacy arose with the diversity of words, signs, and meanings. L. Wittgenstein writes that, unlike mechanical tools, which differ in physical properties, words have a uniform appearance. This is confusing.

Representatives of pragmatism, when analyzing media philosophy, found a methodology similar to theirs in the works of F. Nietzsche. According to the author of the revaluation of values, the printed text and the science and journalism associated with it have created a situation of paralysis for themselves. Texts produce no impression other than criticism. Moreover, criticism generates criticism of itself, and this continues ad infinitum. This abstract fascination with criticism has no practical effect on human activity.

Digital communication technologies have transformed the situation in favor of media pragmatism. Interpersonal synchronous communication on multi-user

digital platforms has transformed the function of writing. It is now used not only for information but also as an institution for coordinating joint social action. Data, information, and knowledge function as socially determined structures. They pragmatically collaborate with people's individual goals and desires. This expands the format of writing. As a result, actions based on non-speech acts are perceived in the context of the social action of coordinating communicative participants.

The pragmatization of the use of internet symbols is reflected in the hyper-textual structure of social networks. Hypertexts, using intertextual links, transform them into content. This content has an open-ended nature, deploying semiotics and being open to graphics. As a result, the role of hermeneutic sensitivity is growing. The perception of text has become intertextual. This is due to the increased role of interpreting text in its various syntaxes within the paradigm of social action and communication. Signs are programmed connections. A sign is read not in the context of its meaning, but in the context of the semiotics of connection. The sign is perceived as a potential connection that must be recognized in a media sphere of possible connections.

Hypertext contains a set of possible linear structures. These require individual choice. As a result, instead of a centralized media system, one finds oneself in a decentralized media system of multiple other participants and contexts.

William Flusser played a special role in the formation of a general subject ontology of media philosophy. He understood that he was at the threshold of a new historical era of media, when the monopoly and uniformity of the diversity of written text were giving way to technical images and the computational technologies of technical images.

This was already the era of digital technologies, and it required a new media philosophy. Its task would include studying how digital practices transform the figurative, textual, and numerical perceptions of social media users.

Particular emphasis is placed on the photographic message. To this end, Flusser wrote the monograph "Toward a Philosophy of Photography". In it, he examined the relationship between written text and the image as text. The monograph

also describes the possible evolution of philosophy as a media tool. However, there is no consensus on the question of what this evolution of philosophy and media will be. In this regard, Frank Hartmann suggests not succumbing to the influence of a superficial interpretation of technological determinism on the media sphere.

It is necessary to avoid value judgments in matters related to the content of the subject ontology. This ontology cannot include discussions; it must contain a categorical structure for describing media as a system of objects and the mutual relationships inherent in these objects, based on the principle of feedback.

Harold Innis argues that a superficial interpretation of the subject ontology of media philosophy is unacceptable. The superficial level of media perception is shaped by the entertainment industry, which operates on themes of play and war. As a result, according to Paul Virilio, an invisible militarization of civil society is occurring.

Alain Badiou does not want philosophy to be transformed into a space of play and language. A reassessment of the institutional status of philosophy under the influence of superficial deformations of social communication could prove the end of media philosophy itself. The risks for philosophy stem from the fact that in studying twentieth-century media phenomena, there is no reference to the philosophical tradition. Marshall McLuhan calls this an ontological shift. In this context, R. Debray developed mediaology.

It describes communication and cultural forms. He uses a method of media analysis in the context of educational, political, symbolic, and technological components involved in the transformation of information through spatial and temporal environments in mediaspheres, graphospheres, videospheres, and hyperspheres. J. Bolter and R. Gruzin called this sphere of spheres the culture of remediation. This culture operates with images constructed from other images and texts. These are technologies of visual design and programming. They shape the perspective of the postlinguistic media era.

J. Steiner, in his book "Real Presence," notes the declining role of language and text in the structure of new media. Hypertext is also losing ground. The iconic

turn became a theme, discovered by O. Neurath. E. Cassirer's assertion that people communicate through symbolic systems became relevant. V. Flusser believed that iconic culture is a step toward the numerosphere of digital culture.

B. Stiegler's "Organology" focuses on the generative interactions between biological, technical, and social organs. Michel Serres proceeds from the premise that the human body and its organs have different nature. The body has a biological basis. Organs have a cultural basis. Organs include media. In this context, the question arises of how organs transform the cognitive functions of the human brain. It is important to understand that the construction of a technical information carrier does not eliminate the natural information carriers found in the ecosystem. Greenland's glaciers provide an example. They contain information about the history of the planet's biosphere. This information is important for media ecology.

Going beyond the written text is also characteristic of representatives of the Frankfurt School's neo-Marxism. Matthias Vogel, in his book "Means of Reason," substantiated the thesis that music and painting fulfill important functions in the criteria of rationality and ethics. Media are presented as a certain amount of mass media activity with mental entities. Mark K. Taylor and Esa Saarinen expanded their understanding of media in their book "Imologies: Philosophy of Media." This expansion touched on the phenomenon of the observer. It turns out that the dominance of the observer's perspective is eroding.

In P. Virilio's "Dromology," media are viewed as transportation systems. When the observer's role collapses and the intentions of technological determinism return, hermeneutic media philosophy becomes relevant. Gianni Vattimo utilized the methodological resources of this philosophy in analyzing the formation of a general subject ontology of media philosophy.

Against the backdrop of the creative industry's rapid adoption of virtual reality technologies in the field of cultural gaming, the philosophy of Immanuel Kant has become in demand. According to this philosophy, human thought does not have direct access to reality. It has access only to the world of phenomena. This world is contained within the boundaries of one's subjective worldview. This world

allows the individual not only to passively exist within the space of possible worlds but also to actively influence their semantics through their presence. In this status, they are supported by a priori forms of reason, represented by categories. A priori forms do not follow from experience and are not dependent on it, but rather condition it. Any specific definition of mediation presupposes the fact of mediation. In this context, interest in cognitive cultural artifacts is growing. They exemplify the objectification of subjective experience through the constant presence of modern humans in space and time with digital devices.

This phenomenon of digital generations has become the subject of study in digital anthropology and digital ethnography. It appears that, in one perspective, the status of the observer has begun to lose ground, while in another perspective it has become a priority in the urban communication environment, as evidenced by the growing interest of researchers in digital users.

The interest stems from the fact that this user has become institutionalized on social networks. This interest is no longer intrinsic, but rather due to the role they have played in shaping the entire system of social network institutions and subcultures. This is a new historical modification of media, which has acquired the ontological status of existence based on internet technologies. These technologies shape the architecture of social communication and are a fundamental value for digital generations.

Limitations in relation to internet technologies have become one of the reasons for potential protest expectations. The strengthening of digital media's position in the worldview of digital generations was facilitated by the epidemiological factors of quarantine isolation. They accelerated the digital generations' assimilation of digital media institutions. Thanks to this, these generations were able to avoid loneliness, find work, and enjoy a wide range of opportunities for organizing their leisure time.

The phenomenon of photography has had the most profound impact on the transformation of the subject ontology of media philosophy. This phenomenon consists in the fact that photography was not considered ontologically. In 1980,

Roland Barthes, in his book *Camera Lucida*, posed the question of the essence of photography. He began to study how photography differed from other media. He discovered that photography is a message without a code.

A photograph is an exact copy of a real object. It follows that the signifier is also the signified. This property precludes various interpretations, since the situation is fixed. But R. Barthes makes an unexpected assumption. In his view, the absence of a code becomes the code of photography. The meaning of a photograph is objective, since it is identical to the real situation of space and time. This is due to the fact that the object of the photograph must necessarily exist. This is what distinguishes photography from literature and cinema.

R. Barthes's subject matter became how the viewer perceives a photograph and the ontology of physical reality it contains. He identified the aspect of *punctum*, which focuses on the photographer's intention and the meaning of the photograph. In this case, he refers to a situation in which the photographer has no script for creating the photograph. This is documentary photography. *Studium* denotes a situation in which a script, determined by the commission for the photograph and the creative process, is used to create the photograph. This is artistic photography.

One of the first fundamental functions of photography was memory. Then the function of emotional evaluative feedback emerged. Research into the phenomenon of photography in the context of identity became important in the context of the growing emancipation of modern society.

Digital technologies emphasized the role of photography as a means of communication and connection between people. In mobile phone formats, photography became a language of communication at the level of the current moment and situation. Words began to be combined with video. As a result, photography was transformed into a technology of video messaging. The digital state became an ontological reality. In this state, photography acquired the instrumental function of communication and continuous presence in dialogue. Digital photography is the result of processing a sample of visual information, displayed by a digital camera or scanner from a paper photograph.

Digital media technologies use computational processes. These processes distinguish them from analog media technologies. Nelson Goodman writes about this difference in his book "Languages of Art". D.N. Rodovik notes that analog photography is an isomorphic transcription of its object. Digital photography is the result of data output through a symbolically mediated connection with the object. Digital photography simulates analog photography.

This raises concerns about the process of digital image creation. This is due to the risk of significant structural changes to analog images during their digital transformation. This situation raises the question of the ontology of photography.

The works of Peter Bazin, Susan Sontag, and Roland Barthes can provide assistance. This is not an ontology of objective physical reality, but an ontology of communicative action shaped by intentionality. A trace of phenomenology is discernible here.

Analog photography has an ontological foundation and physical embodiment. Digital photography lacks physical embodiment and an ontological foundation. William J. Mitchell takes this distinction as a basis in his book "The Altered Eye". A new status has been chosen for digital photography. It is designated as simulacrum. Digital photography does not guarantee the existence of an essence. It creates risks of manipulation of meanings. Jean Baudrillard writes about this.

The subject ontology of media philosophy has found itself in the context of the formation of categorical structures and the almost simultaneous transformation of media. As a result, media philosophy found itself within very limited time frames for reflecting on the processes of the media sphere's transition from analog to digital technologies. Nevertheless, media philosophy has managed to maintain an applied focus and draw on the ontological advances of fundamental philosophical schools. Pragmatism, which is focused on applied methodology, has played a particularly important role as a fundamental resource. Phenomenology and analytical philosophy, developed with the participation of Ludwig Wittgenstein, also played an important role.

Digital Epistemology

Digital epistemology addresses the question of how companies and banks can effectively use data, information, and knowledge to generate profit and solve related problems. In this perspective on the use of information resources, objectivity and reliability do not play a primary role. Moreover, information resources do not refer to all data, information, and knowledge, but only to that portion that has been digitalized by artificial intelligence technologies.

These technologies have become applied tools of the digital market economy and politics. Since digital generations access data, information, and knowledge through mobile phones integrated into online social networks, this is their primary news resource, which includes not only political news but also advertising, navigation, and news from peer interactions. Through digital technologies, digital generations gain access to scientific and professional knowledge.

A distinctive feature of digital technologies is that they not only provide users with social information resources but also receive free feedback from them, as users' use of social networks generates an increase in information resources, which is extremely beneficial for companies that own digital platforms and digital ecosystems. The primary social information resources for people are data, information, and knowledge. From a scientific perspective, data is information obtained through empirical methods of measurement, scientific observation, and experimentation, as well as methods of formal logic, particularly mathematical logic, such as arithmetic operations. In a narrower sense, data refers to signals, information, and messages in the form of a digital language compatible with artificial intelligence technologies for their storage, processing, and use.

In a broader ontological framework, information refers to the objective content of a signal, message, or piece of information. Deciphering this objective content is a task of modern science. Automated information collection and processing systems are used for this purpose.

Social information refers to information that people transmit to each other. In the information services market, media holdings are responsible for the rapid transmission of information about events.

Social information also refers to the meaning that people ascribe to signals, information, and messages at the everyday and user level. In this way, they interpret the data based on their own worldview and understanding of events.

Advertising is based on maximizing the user's worldview. It tends to "amplify" the impact on the consciousness of social media users. The primary emphasis is on the novelty of the information. Frequent repetition of advertising transforms it into information noise and irritates its users. The shift in modern human focus from the knowledge system to the social information system has affected digital generations in the form of clip-based thinking. This thinking is fragmented by the message architecture. It loses the integrity of the semantic perception of the information field. In this format, it becomes vulnerable to manipulation in the context of information warfare.

According to K. Shannon's definition, the useful function of information is that it reduces a person's uncertainty in specific matters related to their life.

Reference social information allows its user to reduce uncertainty when making decisions. Signal social information allows its user to stay informed about current events and quickly manage their activities.

Commercial information allows its user to track important economic indicators. Confidential information is closed to other users due to cybersecurity factors. Biographical information may be open to other users.

Social information is presented in the form of diagrams, sketches, and images. It is also presented in the form of graphs, charts, symbols, videos, photographs, and text. It is presented in the form of presentations, reviews, and samples, and is an important part of exhibition activities. Design plays a key role in the presentation of information.

Social information can have various levels of security clearance and be classified. Information and data with this status are the subject of heightened interest.

Social information can have various security clearances and be classified. Information and data with this status are the subject of heightened interest. This interest is pursued by intelligence, counterintelligence, industrial espionage, and hacking institutions.

Knowledge integrates data and information into a systemic description of processes, properties, and objects from the perspective of a general, subject, and target ontology, as well as an ontology of values. It configures data and information within the context of scientific explanation, engineering solutions, and practical action. Consequently, knowledge contains algorithms for solving typical problems, methods, models, and regulatory requirements. Currently, the use of knowledge is limited to education, professional activity, and scientific research.

In the digital economy, knowledge as a resource is used by engineering and consulting companies. These companies create innovative products based on knowledge in the form of projects and implement them in accordance with the client's needs, followed by maintenance of the infrastructure and communications they have created. These companies engage contractors and subcontractors.

Consulting companies engage in analytics, auditing, and business planning. There is a conceptual level of knowledge, which requires mastery of mathematical apparatus and specialized formalized languages. Its architecture is shaped by laws, theories, principles, postulates, ideals, and norms of scientific activity. There is an empirical level of knowledge, the architecture of which is shaped by facts, verification procedures, and methods of data collection and processing.

When data, information, and knowledge are translated into computing systems, this is done using crisp and fuzzy logic. Large language models use fuzzy logic. In this logic, the probability of a member of a set determines the degree of legitimacy of assigning this member to the set.

Crisp and fuzzy logic are also used in analytics functions. The main resource for them is formalized structures suitable for computation. Everything that is digitized and entered into the system becomes part of a unified computational world. AI does not know what it is analyzing. It learns data structures. It records statistical

relationships between words. It does not see an image. It recognizes pixel patterns. It does not hear music. It encodes frequency distributions. It forms data configurations in probability space. It discovers patterns of causal relationships. Where humans interpret, AI connects.

The more digital data, the more accurately a large language model functions. The more diverse the range of digital data, the richer its latent space. The higher the quality of the digital data, the more reliably learning occurs. Artificial intelligence does not create new meanings. It reveals hidden configurations of data structures. Data is defined as information in a formalized form accessible for analysis.

This can be numbers, text, images, sound, video streams, telemetry, or system logs. Data exists wherever something can be read by AI. It creates the ability for AI to think in the form of computation. 80% of data scientists' time is spent on data quality assurance processes. These processes include data cleaning, preparation, and validation.

Norbert Wiener and Peter Drucker pointed out that data, information, and knowledge form different levels of the cognitive structure of epistemology. This thesis is used in knowledge management theory. Data refers to raw observations, recordings, and measurements. The meaning and context within them have not yet been deciphered.

Information refers to data placed within a system of relationships, where an interpretive structure emerges between the elements. In this context, cybersecurity issues arise. These involve protecting the process of obtaining and storing data and information. Particular attention is paid to digital traces. Monitoring involves identifying vulnerabilities in the digital environment and protecting the epistemic process from unauthorized interference.

Cyber security

Cyber security includes a system of methods, technologies, and practices that ensure the protection of digital systems, social networks, computer programs, and data from attack, theft, damage, or unauthorized access. It protects personal

information, critical infrastructure, and financial resources based on the principles of confidentiality, integrity, and availability of data. Antivirus software, complex passwords, and two-factor authentication are used for this purpose.

Cyber security counters three types of threats. One of these is cybercrime. These are actions organized by one or more malicious actors with the goal of attacking a computer system to disrupt its operation and obtain financial gain.

Cyber attack methodology involves actions aimed at manipulating individual and public consciousness. Cyberattacks employ methods aimed at manipulating individual and public consciousness. They are part of an information war waged against social media users.

Cyberterrorism involves actions aimed at generating fear in social media users with the aim of recruiting them for terrorist activities.

Malware is a tool used by cybercriminals. It is distributed as files or email attachments. Viruses infect files with malicious code. To spread within a computer system, they replicate themselves.

Trojans disguise themselves as licensed software. Spyware secretly monitors social media user activity and collects information about them, such as credit card information. Ransomware encrypts files and data of social media users. Cybercriminals then demand a ransom for data recovery.

Criminals also use adware to distribute malware. They also employ botnets, SQL injections, and phishing. Their arsenal includes man-in-the-middle and denial-of-service attacks.

Data is generated everywhere. Text, images, motion, sound, and GPS data become part of the data stream that forms artificial intelligence training corpora. Language models are trained on text data from books, articles, websites, forums, scientific archives, correspondence, and technical documentation.

Visual data includes images and videos collected from open sources or generated by computer vision systems. Audio and speech data includes speech recordings, podcasts, subtitles, and audiobooks used to train speech recognition systems and voice assistants. Sensory and telemetry data includes signals from sensors,

cameras, navigation systems, and industrial and medical devices. User behavioral data includes clicks, views, routes, and reaction times. This data does not record content, but rather profiles of social media users' actions. It is used for recommendation systems and preference analysis.

Everything is transformed into data through the process of digitization. Text, images, sound, and movement have become computationally uniform. Every element of experience can be expressed numerically. Words are digitally transformed into tokens (a unit of text converted into an index). Images are digitally transformed into a matrix of pixels. Sound is digitally transformed into a frequency spectrogram. Movement is digitally transformed into coordinates and timestamps.

Manual tagging has become impossible. Automated methods have replaced it. Crawling automatically collects text from the internet. Sensor networks collect physical data from IoT (Internet of Things) devices and sensors. APIs and telemetry collect data through software interfaces that record the behavior of systems and social media users. Synthetic data generation aims to create artificial examples based on existing datasets.

Artificial intelligence collects data about itself. Models create synthetic texts, images, and scenarios that are used to train other intelligent systems. This is a form of recursive loop. AI generates data on which new intelligent systems are trained. This results in the formation of a meta-level of data.

AI operates only on what's in its data. Most training datasets are created in the United States and China. This means that global models contain media content. This raises the question of how well the data reflects the diversity of human experience. If a dataset doesn't include all natural languages, AI won't be able to understand it. If the data lacks the experiences of certain cultures, AI won't be able to reconstruct them. If the data is distorted by political or ideological filters, the model's behavior will also be distorted.

Control over data has become a form of power. Whoever decides what data is collected determines what knowledge will be. For photographic characteristics, the date taken, geographic location, and shooting device are important. For text

characteristics, the author, language, publication time, and license are important. For dataset characteristics, the source, volume, structure, and annotation method are important.

For data to train artificial intelligence, it must not only be collected but also explained. Annotation refers to the process of adding labels, categories, and tags that associate data with concepts. Without annotation, a neural network doesn't know what's being depicted or said. It only operates on numbers. Annotation enables supervised learning, where models match input data with correct answers. Annotation enables supervised learning, where models match input data with correct responses and learn from the model.

However, annotation introduces subjectivity. Different people perceive the same phenomenon differently. Therefore, even in machine learning, humans remain a source of bias. However, without human annotation, the creation of most modern AI models, from facial recognition systems to generative transformers, would be impossible.

Data format influences how data is read, interpreted, and combined. It defines the boundaries of artificial intelligence memory. It determines what can be extracted from the data. Therefore, data architecture plays a key role. However, at the initial stage of artificial intelligence's work with data, it is only a potential information resource. Vectorization makes this resource operational. It converts data into a numerical form accessible for processing by neural networks.

Vectors are called embeddings. They create a multidimensional map where similar elements are located close to each other, while distant elements are located further apart. This enables AI to compare, search, predict, and classify – all based on relationships between vectors, not on understanding.

For AI, meaning is associated with the structure of proximities in the embedding space. This space generates what humans call knowledge. After converting data into vectors, AI training begins. This is the process by which a language model searches for patterns by associating input data with internal parameters. In

classical machine learning, this occurs through optimizing the weights of numerical coefficients that regulate the strength of connections between neurons.

When a language model makes an error, it adjusts the weights using the backpropagation algorithm developed by Paul Werbos. What humans call knowledge exists in machine architecture as distributed memory.

These billions of weights record the statistical relationships between elements of the training set. Artificial intelligence is entirely dependent on the data it is trained on. This makes data quality a crucial factor. If the data is incomplete, noisy, or biased, the model will reproduce these errors in its output.

Therefore, models trained on limited corpora exhibit cultural or linguistic biases. When a language model answers a question or generates text, it activates weights modified during training. It reproduces patterns ingrained in the data. Its response is a reactivation of structural traces.

In modern intelligent systems, embeddings are stored separately from the main model. This allows models to access external knowledge sources, find similar cases, and generate responses more contextually relevant. The result is a hybrid memory. Part of this memory resides in the weights, and another part resides in the external data.

In digital ontology, information is a stable configuration of data. It does not exist prior to processing but is created at the moment of computation. Corporations that use artificial intelligence do more than just use data. They appropriate the process of observation. User data becomes raw material for training models, often without the consent of social media users. Shoshana Zuboff has dubbed this phenomenon "surveillance capitalism." It's a system where privacy becomes a source of profit.

Every search query, every photo, every pause before a click becomes an element of a statistical model that trains other intelligent systems. This has resulted in privacy threats in the age of artificial intelligence. Deanonimization refers to the ability of AI to reconstruct the characteristics of a specific individual from any-

mized data. Data leaks refer to the loss of control over AI repositories where personal information is collected.

Profiling refers to the construction of psychological and behavioral models of users for advertising, politics, or manipulation. Inference refers to the ability of AI to discover sensitive information, even if it's not explicitly provided on a social network. Data produces meaning. Thinking becomes a derivative of statistics. Every object leaves a digital trace.

Every action leaves a coordinate. Every state leaves a signal. AI technologies have shifted to navigating vast volumes of digital data. As a result, the digital environment has begun to influence human thinking. This is particularly evident in the role of digital platforms and digital ecosystems in organizing the modern information space.

Information is created by the digital context, as it predetermines the specificity of the epistemological object. Digital data is considered objective evidence. The linguistic units and their combinations contained in a text, as well as the pattern identified by a digital algorithm, often become self-sufficient evidence of meaning. But behind these neutral and objective corpora lie interpreters – designers of the analytical parameters that shape them.

Digital technologies reveal patterns but offer no explanations. Therefore, human intervention is typically embedded within a digital object, if we consider it as a history of human interaction with a technological device. A human being, as an agent, adapts to a technological system and inputs the data it can process. The technological system produces responses. What's important is not so much the digital object itself as the infrastructural system of relationships surrounding it.

Text processing can produce a number of different tokens for the same incorrectly recognized word, requiring additional manual editing. Ultimately, it becomes clear that a token, as a type of data, is not an objectively given phenomenon in the text, but a construct that emerges in the interaction between a human and a tool. The same word may be perceived differently in different contexts, but it will still be the same token.

Wordplay and common usage will be recognized as a single unit. A statistically significant relationship can be established between two tokens in a text, but this may not always be meaningful. Identifying words in a text that are most likely to appear together is like guessing the meaning of a grain of wheat: there is a coincidence between them, but not necessarily a cause-and-effect relationship.

Developing a user interface that enables interventions into the source data corpus may be the solution. The procedures for selecting, sifting, and processing data in this interface are transparent. Each intervention is offered to the user separately, and its potential limitations are considered in advance.

Social media, unlike newspapers or television, are open. You can simply register and start producing content. This eliminates the possibility of verifying information. Social media algorithms are aimed at maximizing user engagement.

This makes the social media business model effective because it makes money from advertising. People are highly interested in sensational news and conspiracy theories. Many users click on them, and the algorithm determines that such content generates engagement.

The algorithm doesn't care whether the news is fake or real. It only cares that it generates clicks. Humans are naturally attracted to such sensational news, and numerous studies confirm that fake news spreads faster than regular news.

Companies partially control algorithms. They can control them at the level of defining basic functions or influence the algorithm's behavior in specific situations. A digital platform can moderate search results on a website or selectively influence the algorithm's behavior.

But it's impossible to explain why a particular video appears in recommendations. There are an incredible number of parameters. But the developer can justify the reasons for this type of content appearing in recommendations.

In the US, digital platforms are protected by Section 230 of the Communications Decency Act. This exempts them from liability for their content. Social media platforms openly oppose and even combat the spread of disinformation. But at the same time, they profit from this content.

The GDPR (General Data Protection Regulation, in effect in the European Union) has completely changed the way users' information is used. Algorithms are designed to spread popular opinion. But they only use existing information, not create new information.

Objective algorithms don't exist. A recommendation system selects, say, ten pieces of content from billions. Clearly, there's no single, correct solution in this case. Some bias exists in what the algorithm selects for a user's feed. And this bias can be unintentional. On social media, user engagement is the simplest and most accessible metric on which to train an algorithm. Obtaining data on real user satisfaction is much more difficult.

Algorithms can't be objectively fair. Fairness can only work when a company advocates for an impartial approach to information, sets rules, and enforces them. But social networks don't disclose how they operate, so no one knows what principles they adhere to.

Therefore, there's no guarantee that fairness standards are met. Social networks don't have established standards because they make money from content. There are alternative services that are trying to move away from the passive model of information consumption. Television operates on this model. The viewer watches what's shown on the screen. At first, the internet worked differently.

Legal Philosophy and Intelligent Automation

Intelligent automation technologies are developing at an accelerated pace. Algorithms analyze data, predict behavior, make "suggestions" based on judgments, and, in some cases, effectively replace human initiative. The problem is that such AI activity is not yet fully covered by current legal norms. Furthermore, the law has not yet developed a universal response to the challenges posed by the autonomy, opacity, and learning ability of artificial intelligence systems.

Healthcare, security, the judicial system, transportation, and digital administration are the most vulnerable. It is here that algorithmic errors, built-in biases, or failures can have irreversible consequences. At the same time, concerns are

growing about invasion of privacy, the substitution of human responsibility, and the disruption of labor markets. Artificial intelligence is increasingly entering into relationships where law has traditionally protected the dignity, freedom, and autonomy of the individual, challenging jurisprudence to rethink its limits and functions in the digital age.

Modern approaches to regulating artificial intelligence are emerging both within national legal systems and in international practice. Despite attempts to develop universal ethical standards such as the UNESCO Recommendations or the Fourth World Health Organization Principles, a global consensus is lacking. In some countries, preventive regulation with a focus on human rights prevails, while others emphasize innovative freedom and self-regulation.

The concept of "artificial intelligence" (hereinafter AI) is both a well-studied and controversial scientific category, defined in various ways by scholars across various disciplines. The development of intelligent systems capable of autonomous decision-making requires not only scientific and technical understanding, but also legal, philosophical, and moral analysis. The historical emergence of AI demonstrates a logical evolution from abstract ideas about thinking machines to the creation of complex algorithms capable of learning and self-development. Questions related to the operating principles, regulation, and application of AI are becoming key in shaping the future of civilization.

The origins of AI as a scientific discipline date back to the 15th century, when the brilliant Renaissance thinker Leonardo da Vinci, although he did not formulate the concept of artificial intelligence in the modern sense, was already pondering mechanisms that imitate human behavior. His drawings of a mechanical knight, created around 1495, represented an automaton – a device capable of moving according to a predetermined program.

This device was not artificial intelligence in the modern sense, but the idea that mechanisms can replicate the actions of living beings anticipates some of the automation principles that underlie modern systems.

The scientific understanding of artificial intelligence began when computing technology provided the tools to implement such ideas. In 1950, Alan Turing published his paper "Computing Machinery and Intelligence," proposing a test for assessing a machine's ability to think. This marked the beginning of the development of artificial intelligence as a discipline.

The formal birth of AI is considered to be 1956. That year, the Dartmouth Conference took place, where John McCarthy, Marvin Minsky, Claude Shannon, and Nathan Rochester not only coined the term "Artificial Intelligence," or AI for short, but also outlined the main research directions.

From this point on, AI became an interdisciplinary field, combining logic, engineering, psychology, linguistics, and law. The Soviet school of artificial intelligence developed in parallel. Andrey Kolmogorov developed probability theory, which later formed the basis of machine learning algorithms, although it was not yet directly applicable to AI.

Theories of automatic control, logical models, and pattern recognition were also developing. Modern AI relies on the principles of self-learning and adaptability. Intelligent systems generate texts that are almost indistinguishable from those created by humans, raising both technical and ethical questions. Their operating principles are based on big data processing, the integration of methods (speech recognition, computer vision), and the flexibility of algorithms.

Progress comes with challenges. Autonomous systems, such as driverless cars or military drones, exhibit the "black box" problem, that is, the opacity of their decisions. Leonardo da Vinci's mechanical knight was completely predictable, which cannot be said of modern algorithms. Law must learn to respond to algorithmic actions. Understanding the historical evolution of AI is key to understanding the challenges that lie ahead in the fields of ethics, law, and industry practice.

Narrow AI is what people encounter every day. Algorithms that perform clearly defined functions, such as facial recognition on smartphones, navigation through city traffic, automatic text translation, and the generation of creative compositions—all these are examples of weak AI.

It is "smart" within a specific task, but cannot go beyond that. This kind of intelligence does not reason or reflect. It operates according to patterns, albeit complex ones.

General AI, according to the developers' vision, will have the multimodal ability to learn, adapt, and solve unfamiliar problems without step-by-step programming. It is based on complex mathematics that helps it find patterns, build connections, and make predictions. This kind of AI doesn't simply execute commands; it learns from examples. It can be compared to a student who looks at a multitude of problems and understands how to solve them, even if they have never seen them before.

One of the simplest manifestations of AI today is Reactive Machines. These are machines that respond to current stimuli but do not store past experience. The next level is formed by intelligent systems with Limited Memory AI. These algorithms are capable of analyzing data over a certain period of time and drawing conclusions based on it. This is precisely how modern autonomous cars operate, using cameras, radars, lidars, and other sensors to "understand" the road environment and road markings in real time, recognizing traffic lights, pedestrian behavior, road signs, and other objects. They learn, but only within a given amount of information, without long-term memory or context.

Deep Learning expert systems imitate the thinking of a specialist in a specific field. Such solutions are particularly in demand in medicine. In large clinics, you can find AI systems that help doctors make diagnoses based on symptoms, review medical histories, or study laboratory data. Such systems do not replace humans, but they are becoming an important decision-making tool, especially in conditions of overcrowding and staff shortages. If AI can "hear" text, it can increasingly "see" it. We're talking about Computer Vision, technologies that enable AI to interpret images and videos. This includes everything from facial recognition at airports to tumor diagnosis in X-rays, from product quality control in factories to drone operations in low-visibility conditions.

AI has conquered the field of robotics, where intelligence is embedded in physical devices. Industrial manipulators, autonomous aircraft, surgical assistants, robotic vacuum cleaners, and research vehicles operate, analyze, and make decisions. Natural language processing (NLP) AI systems make up a significant part of everyday experience. This digital ecosystem includes chatbots, translators, humanoid robots, and voice assistants. Interaction between humans and artificial intelligence increasingly takes the form of lively dialogue, albeit with the illusion of understanding.

An intelligent generative AI system creates texts that read like human writing and also creates images that can deceive the eye of a skilled artist. Music, code, and video are all the work of AI.

Given the widespread use of AI in various fields, the need for sector-specific regulation is emerging. Governments are beginning to formulate recommendations, laws, and codes that will apply not to AI generally, but specifically to specific areas such as healthcare, transportation, education, and finance. In medicine, artificial intelligence has revolutionized the field, as it can process massive amounts of data, improve diagnostic accuracy, personalize treatment, accelerate drug development, and optimize the work of medical institutions.

Artificial intelligence analyzes medical images, which help doctors not only speed up but also improve the accuracy of diagnoses, significantly reducing the risk of errors. Ethical and legal issues are particularly pressing in this area.

In the financial sector, artificial intelligence is used to analyze market data, manage risks, forecast financial flows, and automate lending. Banks use AI to identify suspicious transactions, helping combat fraud. This raises another important ethical question: "How can we ensure transparency in the operation of algorithms to avoid customer discrimination based on biased data?"

The manufacturing sector is actively using AI to optimize processes, automate assembly lines, manage supply chains, and predict equipment failures. Robots equipped with artificial intelligence perform tasks that previously required human labor, reducing costs but also creating new challenges in terms of retraining and

repurposing the workforce. One of the key issues facing companies is the ethical use of AI as a replacement for human labor.

The question arises: how can we ensure a transition without social unrest and support workers whose jobs may be replaced by robots?

In the security sector, AI has become an important tool for threat detection and crime prevention. Video surveillance systems, facial recognition, and analytics from cameras and sensors make it possible to identify suspicious behavior and prevent crimes. Such technologies often face accusations of privacy violations and lack of transparency in their algorithms. How can we determine where the right to security ends and an invasion of citizens' privacy begins? First and foremost, it's important to note the importance of data privacy.

Modern AI technologies rely on processing vast amounts of information, and as a result, there is a need to protect personal information. AI can collect and analyze data related to human behavior, preferences, and sometimes even personal life, which puts privacy at risk. An unprecedented rise in phone fraud has been recorded. Fraudsters posing as employees of banks, law enforcement agencies, and government agencies have defrauded citizens of significant sums.

The introduction of modern technologies into the justice sector represents one of the most sensitive and at the same time promising forms of digital transformation. AI systems are used in case summaries, automation, simplification of case searches, classification of offenses, as well as in office workflow and document management.

As AI becomes increasingly embedded in human life, the potential consequences of errors, abuses, and unlawful use grow. This requires states, developers, and society as a whole to not simply observe, but to actively and promptly respond to challenges that call into question legal sustainability, humanitarian principles, and fundamental human freedoms in the digital age.

The consequences of AI-driven labor displacement are worth considering. Automation in various industries will render many professions obsolete, creating new social divides. This is especially important in the context of economic inequal-

ity. If AI significantly replaces humans in jobs, how will the social security of those whose professions disappear be ensured? This issue remains relevant in all countries experiencing mass automation.

There are several international examples where ethical issues are becoming more pressing for individuals. A social credit system based on artificial intelligence is being actively implemented, raising serious concerns regarding individual freedom and the right to independence. This system analyzes citizens' behavior, their financial activity, and social connections, which could lead to serious consequences for those who do not meet predetermined standards.

One of the key areas of contemporary legal debate is the issue of assigning responsibility when using autonomous technologies, as well as the admissibility of granting AI legal subjectivity. The collection of articles "Artificial Intelligence and Law," published in the HSE24 journal, and the "Ethical Index of Artificial Intelligence Systems" methodology actively discuss models of functional and distributed responsibility arising from the introduction of AI into medical, financial, law enforcement, and other critical areas of human life. In the absence of a universal legal framework, the role of self-regulation is increasing in various countries. One of the priority areas of activity is combating the counterfeiting of digital information.

PAI is exploring ways to label synthetic content, particularly deepfake videos, to ensure transparency and prevent manipulation. This labeling serves as a digital "label," alerting the user that they are viewing the output of a neural network, not objective reality.

In 2021, UNESCO adopted the Recommendation on the Ethics of Artificial Intelligence. This is the first international document dedicated to universal principles for the ethical use of AI. The support of 193 states demonstrates a commitment to creating a common civilizational language on issues related to machine learning. The document prioritizes human values, dignity, freedom, and the right to self-determination above the functionality and efficiency of algorithms. UNESCO emphasizes that technology should not govern society. Its purpose is to serve the interests of people, not vice versa.

The UNESCO Recommendations center on universal principles that should guide the development and application of AI in any country, regardless of the level of digitalization and economic capabilities. These principles include respect, protection, and promotion of human rights and human dignity.

AI technologies must be limited to their intended use. Their use must be justified and safe. AI must be bias-free, ensure equal access, and consider the needs of vulnerable groups. AI technologies must be used in a sustainable manner, minimizing negative environmental impacts.

Algorithms must be understandable, traceable, and auditable, especially when making decisions that affect human rights and freedoms. There must always be an accountable entity for any AI action. AI systems must comply with privacy standards, legal data processing regulations, and the right to privacy. AI must not make final decisions in critical areas of life and health without human input.

AI development must take into account the diversity of languages, traditions, and conditions in different regions of the world. Stakeholder participation in policymaking and AI regulation is essential, including representatives of science, business, civil society, and governments. The World Health Organization (WHO) has played a significant role in developing an ethical framework for healthcare. The published global report emphasizes that AI has the potential to transform healthcare by improving diagnostic accuracy, clinical care, and research. However, it also raises risks related to computational bias, unethical data collection, cybersecurity threats, and environmental harm.

To mitigate these risks, the WHO has formulated six core principles, including protecting human autonomy; promoting human well-being and safety, as well as the public interest; ensuring transparency, explainability, and understandability; strengthening responsibility and accountability; ensuring inclusiveness and equity; and promoting responsive and sustainable AI. These principles will guide WHO's future work in supporting efforts to ensure that the full potential of AI in health and public health is harnessed for the benefit of all.

In February 2025, the French capital hosted the largest European summit dedicated to the legal and ethical aspects of AI. The discussions resulted in an international declaration supported by more than sixty countries, including France, China, Japan, and Australia. The document proclaimed the need for an open, ethical, and accessible approach to the development and application of AI. The United States and the United Kingdom refused to sign the declaration, citing the potential for it to hinder innovative development.

Against the backdrop of Europe's focus on preventive regulation and ethical frameworks, the United States' position demonstrates a different approach to legal oversight of artificial intelligence. In 2023, the Joe Biden administration introduced its first executive order aimed at ensuring the reliable and safe development of AI, emphasizing independent model verification, personal data protection, prevention of discrimination, and regulation of generative services.

In early 2025, Donald Trump rescinded this order, arguing that excessive measures were stifling innovation and undermining the country's technological leadership. Instead of strict regulation, the administration has focused on developing an AI platform "free from ideological bias" and creating a favorable climate for business. This position reflects the American legal tradition of avoiding centralized control of AI, relying on market incentives, competition, and self-regulation.

However, it also calls into question the readiness of the national legal system to address challenges posed by the lack of transparency in its methods, the risk of widespread disinformation, and copyright infringement.

In the international context, the United States is acting less as a signatory to global agreements than as an exporter of technological standards, primarily through the marketplace rather than through law.

The diversity of approaches reflects the ongoing search for the optimal balance between innovation and law, between efficiency and humanism, and between politics and cultural values. Control over artificial intelligence can no longer be reduced to formal regulation. It requires a flexible, well-thought-out, and, most importantly, collaborative strategy of action at both the political and industry levels.

A significant step in the legal regulation of artificial intelligence in recent years was the adoption by the European Union of the world's first comprehensive law directly dedicated to this technology. The document, officially numbered 2024/1689, was approved by the European Parliament and the Council in early 2024 and outlines the principles for the coexistence of AI and law in the 21st century. It is extraterritorial meaning that any developer of AI systems, regardless of their geographic location, is required to comply with the regulation if their product is to be used in the EU.

Thus, the law has become a de facto global benchmark, capable of influencing how artificial intelligence will be developed and applied worldwide. The purpose of the new law is to thoroughly codify the risks associated with its use.

The regulator abandoned a one-size-fits-all approach and instead introduced a multi-tiered classification, from minimal to unacceptable risk. For the first time, the law has adapted a risk grading tool from the field of biosecurity to digital technologies. Systems that fall into the high-risk category, such as AI in education, transportation, legal proceedings, or employment, are subject to rigorous scrutiny for compliance with basic rights and reliability standards. In other words, the closer the technology is to humans, the higher the bar.

The regulation includes a list of prohibited AI uses. This includes systems capable of manipulating human behavior against their will, such as those used for social scoring. The regulation also includes the use of AI for remote biometric facial recognition in real time in public places, except in cases strictly limited by law and subject to judicial authorization. An outright ban on such practices means that the law assumes the function of not just a supervisor, but a value filter.

Particular attention is paid to so-called general-purpose models, which underlie most modern generative AI systems. These are large-scale language or multimodal models capable of performing a wide range of tasks, from text generation to image analysis, regardless of the specific scenario.

The regulation introduces the concept of such models for the first time at the legislative level and describes a special procedure for their regulation. Developers

are required to disclose the purposes, scope, and sources of training, describe potential risks, take measures to mitigate them, and inform users of the synthetic nature of the content, if any. The law seeks to establish a regime of public accountability and transparency to prevent situations in which generative AI becomes a means of large-scale dissemination of fake news, authorship substitution, or manipulation of public opinion.

The document also calls for the creation of a unified database of high-risk AI, accessible to both government agencies and the public. This is one of the law's most original mechanisms. It aims to increase transparency, allowing users to learn which systems are used in a particular agency, school, bank, or hospital.

As a result, it establishes not only a technical but also a procedural standard of transparency, which is important for strengthening public trust in automated decisions. The law does not deny the right to innovation. It provides for mechanisms to support developers, including the creation of regulatory "sandboxes" within an experimental legal regime.

The essence of this regime is to allow for easing legal regulations so that companies can put their innovative developments into practice. Freed from restrictions, entrepreneurs gain the tools to test new technologies and implement improvements to existing products. This approach combines flexibility with responsibility, which is especially important in the context of rapidly changing technologies and risks.

Adversaries can exploit general AI to harm individuals, organizations, or society. There are risks due to malfunctions. The widespread adoption of general AI is associated with a number of broader systemic risks.

AI systems are capable of programming, generating photorealistic images on demand, and engaging in lengthy conversations. Since May 2024, new models have significantly improved their performance in scientific reasoning and programming tests. The harmful effects of general AI are already well understood. These include fraud, non-consensual distribution of intimate images, child abuse material, bias, reliability issues, and confidential data leaks. While mitigation

methods have been developed, none fully address the problem. Emerging evidence confirms the existence of subtler forms of discrimination associated with AI.

Risk management methods for general AI are still in their infancy, but are already showing progress. Technical approaches to risk assessment and mitigation have been developed, applicable to both developers and regulators, although they are still limited. Tools for interpreting AI decisions are particularly underdeveloped, but progress is being made in this area. At the same time, efforts to standardize risk management approaches internationally are intensifying.

The rapid and unpredictable pace of development in general-purpose AI poses a so-called "evidence dilemma" for policymakers. With limited data, they must balance proactive measures, which may prove premature, with the risk of delay, which could exacerbate the consequences. In response, companies and governments are creating early warning and risk management systems. Some implement preventative measures as new data becomes available, while others require proof of safety before launching models.

Key areas for further research remain the pace and methods of measuring the progress of general AI, establishing reasonable risk thresholds, improving policymakers' access to relevant information, assessing the risks of technology implementation, understanding the inner workings of models, and developing robust AI architectures.

AI development depends on human action, rather than being self-sustaining. The future of the technology remains uncertain, with both positive and negative scenarios possible. Overcoming this uncertainty requires informed decisions from society and states. International jurisprudence is gradually shaping a new direction in the application of law related to the use of artificial intelligence technologies.

This is an area where technology, intellectual property, and fundamental human rights intersect. Modern litigation focuses not only on the technical aspects of AI, but also on who is responsible for the algorithm's actions and the extent of its influence on social and legal institutions.

The number of administrative proceedings related to automated control will increase. A rise in criminal cases is also expected, particularly in the area of cyber-crime, where AI is increasingly being used as a tool for committing offenses, from phishing to the creation of false evidence. Court cases involving artificial intelligence not only reflect the growing number of digital conflicts but also mark the emergence of a new area of cyber law, where the key question arises whether a legal system based on human logic and the presumption of rationality can adequately respond to behavior generated by synthetic models.

The burden of consequences should fall on the individual or organization. This thesis is of paramount importance for future legal frameworks concerning civil and criminal liability in the field of AI.

Digital Law

Digital law should be understood as a set of legal norms and institutions regulating the diverse relationships associated with the implementation and use of digital technologies. However, these norms relate to various branches of law and are not united by a single subject or method of regulation. The emergence and need for further development of digital law is driven by the emergence and rapid development of relationships between new subjects and objects of digital legal relations, each with its own unique characteristics. Specifically, these include the legal status of subjects (users, operators) and the legal regime of objects of legal relations; the grounds for the emergence, modification, and termination of legal relations; and the specific content of legal relations mediating digital relations.

These also include unique sources, among which self-regulatory acts and technical standards play a significant role; a unique conceptual framework; questions of the legality of extending state sovereignty to cyberspace; and questions of identifying participants and ensuring the security of information exchange. A controversial question arises as to whether digital law can be viewed in two ways. On the one hand, can it be considered a separate system consisting of formal rules of digitalization similar to technical regulatory legal acts? On the other hand, can it be

viewed as a mechanism (technology) for implementing private and public rights and freedoms within the framework of existing branches of law?

Under current conditions, it is impossible to definitively formulate the subject matter of digital law due to the constant, rapid development of digital technologies. Currently, the subject matter of digital law includes a variety of social relations that arise, change, and cease in cyberspace. Legal norms and institutions of digital law relate to various branches of law and are not united by a single regulatory method, which is not mandatory for complex branches of law. At this stage, it can be classified as a complex branch of legislation.

Subjects of digital relations can be divided into holders of digital data, digital rights, and digital technologies. These are persons who create digital data and digital technologies and, on the basis of law or agreement, have the right to use them, as well as the right to permit or restrict access to them. These are users of digital data, digital rights, and digital technologies. These are entities that, by law or contract, acquire the right to use digital data, the results of its processing, and digital technologies to varying degrees and to varying extents.

These are digital intermediaries that provide information and other services, as well as work based on special contracts regarding digital data, rights, and technologies, ensuring the rights and legitimate interests of owners and users. The objects of digital legal relations are information in digital format, digital technologies, digital assets, and digital legal objects.

These include cryptocurrency, big data, accounts, domain names, digital intellectual property, smart contracts, personal data, bonus points, virtual gaming assets, and purely electronic forms of obligatory or corporate rights. They also include utilitarian digital rights, digital financial assets, and block chain, a technology that not only created digital rights but also made transactions with them possible, as well as the maintenance of various registries based on it.

To achieve the objectives of improving the legal regulation of digital relations, not only a specific regulatory and technical-legal framework for digital law is necessary, but also basic knowledge of economics and digital technologies. A

key tenet of digital law is that the rules of communication between parties are created, first and foremost, by the parties themselves. It is important that these rules do not violate the law.

Digital law includes the regulation of relations in the areas of IT projects, online publication, block chain, personal data processing, big data, and artificial intelligence. These areas intersect and are complemented by other regulatory components. Legislation related to digital law is constantly evolving.

Most people encounter digital law in their daily lives, but the entry into legal relations in this environment is usually unconscious, despite affecting many actions performed online and related to digital services. Cybercrimes, privacy-related offenses, and other events associated with the transition of many services and capabilities to the digital environment are increasingly occurring. Digital law is important for those launching IT projects, owning websites, or providing any services in the digital environment.

When faced with all of this, many people consider consulting specialists with expertise in digital rights. Many lawyers need to be familiar with digital law, as more and more legal relations are moving into the digital environment.

Belarusian legislation does not have separate sections specifically related to the regulation of the digital environment, and these norms are scattered across various laws. Digital law includes the regulation of relations in the areas of IT projects, online publication, block chain, personal data processing, big data, artificial intelligence, and other areas.

These areas overlap and are complemented by others, and ultimately, the scope of digital law expands as digitalization enters new areas of society.

The following sections can be roughly distinguished within the field of digital law: – Copyright of digital entities; – Software law; – Digital money law; – Digital transactions law; – Digital disputes law; – Software robot law; – Digital public administration law; – Digital government law; – Right of access to data and protection of access. Legislation related to digital law is constantly evolving. The

Republic of Belarus is actively developing e-government and creating a "digital state," which will allow for more effective solutions in various areas.

Digital law covers issues of electronic payments, electronic and mobile money, e-banking, consumer protection in payment markets, artificial intelligence, big data analytics, competition in the era of big data, block chain, crypto currency, smart contracts, digital identification and authentication, privacy, the intersection of competition law and intellectual property, and search engines.

Various types of fraudsters can attack citizens' digital rights. One way to infringe on a person's digital civil rights is to foist completely unnecessary digital services, such as spam, on them. Citizens' digital rights are protected by law. For example, Chapter 31 of the Criminal Code of the Republic of Belarus contains the following articles: – Article 349. Unauthorized access to computer information. – Article 350. Modification of computer information. – Article 351. Computer sabotage. – Article 352. Unauthorized Appropriation of Computer Information. – Article 353. Manufacture or Sale of Special Means for Gaining Unauthorized Access to a Computer System or Network. – Article 354. Development, Use, or Distribution of Malware. – Article 355. Violation of the Rules for Operating a Computer System or Network.

Digital law is difficult to recognize as an independent branch of law, given the obvious heterogeneity of the regulated sphere and the diversity of regulatory methods used. In summary, digital law should be understood as a set of legal norms and institutions governing a variety of relationships related to the implementation and use of digital technologies. However, these norms are not united by a single regulatory method and pertain to different branches of law.

Cyber Law

Cyber law regulates social relations arising in the digital space, including the Internet, through the use of digital technologies and information systems. The legal norms governing cyber law are diverse and present in such fields as intellectual property law (copyright protection in cyberspace); information law (protection of

personal data in the digital environment); administrative law (ensuring the functioning of e-government); international law (international protection of intellectual property rights); and criminal law (criminal liability for the creation, use, or distribution of malware). These social relations are united by a common domain, the digital space.

Computer law and cyber law are similar in content, but not identical concepts. Computer law aims to regulate social relations that arise during the operation of computing devices, computer systems, their network infrastructure, and software. Internet law is a set of legal norms that share a common subject matter and are integral parts of digital law, regulating social relations mediated by the functioning and use of the global information and telecommunications network, the Internet. Social relations aimed at ensuring a balance of interests between individuals, businesses, society, and the state in the context of global digitalization and arising during the functioning, use, and provision of virtual space, digital platforms, information and telecommunications networks, and their infrastructure are regulated by the norms of network law.

Social relations aimed at balancing the interests of individuals, businesses, society, and the state in the context of global digitalization and arising from the functioning, use, and provision of virtual space, digital platforms, information and telecommunications networks, and their infrastructure are regulated by the norms of network law. The legal regulation of information law focuses on social relations in the creation, receipt, transmission, processing, storage, distribution, and use of information.

The distinction between the concepts of "digital law," "cyberlaw," "computer law," "internet law," "network law," and "information law" is rather arbitrary due to the heterogeneity of the social relations that constitute the subject of legal regulation, as well as the interconnectedness and interweaving of elements of digital technologies. The differentiation of terms is generally determined by the goals and objectives of the researcher.

The subject matter of digital law is reduced to social relations mediated by the digital environment (the Internet), which determines the rather narrow range of social relations subject to regulation by digital law. Digital law focuses on a wide range of social relations.

The digital space is not limited to the internet; digital technologies permeate all spheres of human life. Consequently, digital legal relations extend beyond internet-based relationships and encompass a broader technological landscape. Social relations regulated by digital law can arise without the internet, including local databases; autonomous control systems (e.g., driverless taxis); the digitalization of forensic processes (e.g., forensic examinations without internet access); and government digital platforms (e.g., electronic voting). Thus, not all digital technologies require an internet connection, but the social relations that arise from them require legal regulation.

Digital information is information expressed in digital form (data, files, messages, images). Digital technology is the tools, processes, forms, and systems that enable the creation, processing, storage, transmission, and use of digital information. Digital information is the object, and digital technology is the means of its processing. The classification of digital rights as first-generation and second-generation rights remains controversial.

The importance of identifying digital rights as a new generation stems from the impossibility of fully protecting human rights through existing human rights in the digital environment, solely as a new form of expression, without emphasizing the independence of digital rights. Human existence within the framework of internet technologies occurs through indirect channels (programs, websites, social media platforms, investment platforms, virtual gaming services), which in most cases belong to private companies or whose owners are unknown.

This creates additional threats to comfortable and secure human activity in the information environment, the transformation of state obligations to protect human rights in the digital sphere, and the involvement of private companies among obligated entities in a broad cross-border context.

The object of these digital rights is information. Information is presented in a special digital form. Digital rights are realized through the use of digital technologies. Digital rights belong only to participants in digital communication. The foundations of digital rights are set out in Article 19 of the Universal Declaration of Human Rights, according to which everyone has the right to freedom of opinion and expression. This right includes freedom to hold opinions without interference and to seek, receive, and impart information and ideas through any media and regardless of frontiers.

The Okinawa Charter on the Global Information Society enshrines the right to exchange information, freely communicate, and express opinions on communications networks, the right to access and use telecommunications networks, and the right to digital technologies and artificial intelligence. This list includes block chain technology, the Internet of Things, cloud services, augmented reality, and the right to use digital services.

REFERENCES

1. Hassan, R., Sutherland, T. (2017). *Philosophy of Media A Short History of Ideas and Innovations from Socrates to Social Media*. Routledge, London – New York.
2. Derrida, Jacques (1997). *On Grammatology*, (Baltimore/London, John Hopkins University Press).
3. Rorty, Richard, (ed.), (1967) *The Linguistic Turn. Essays in Philosophical Method*, Chicago, University of Chicago Press.
4. Saarinen, E, Taylor, M.C. (1994). *Imalogies: media philosophy*. Routledge, London – New York.
5. Rorty, Richard (1993), Putnam and the Relativist Menace // *The Journal of Philosophy*, vol. XC, 9, September, pp. 443-461.
6. Wittgenstein, L. (1958). *Philosophical Investigations*. Oxford: Blackwell.

7. Friedrich Nietzsche (1995). *The Utility and Liability of History for Life // Unfashionable Observations*, *The Complete Works of Friedrich Nietzsche*, vol. 2. Stanford: Stanford University Press, trans. Richard T. Grey, pp.85-167
8. Flusser, V. (2013). [Post-History](#), trans. Rodrigo Maltez Novaes, Minneapolis, MN: Univocal.
9. Flusser, V. (2000). [Towards A Philosophy of Photography](#), trans. Anthony Mathews, London: Reaktion Books.
10. Hartmann, F. (2000). *Medienphilosophie*. Wien: WUV.
11. Innis, H. A. (1972). *Empire and communication*. Toronto: University of Toronto Press.
12. Virilio, Paul. (2006). *Speed and Politics: An Essay on Dromology*. Translated by Mark Polizzotti. Semiotext(e) Foreign Agents.
13. [Jason Barker](#) (2002). *Alain Badiou: A Critical Introduction*, London, Pluto Press.
14. McLuhan, M. (2005). *Understanding Me*; The MIT Press: Cambridge, MA, USA.
15. Debray, R. (1996). *Media Manifestos: On the Technological Transmission of Cultural Forms*. London, Verso.
16. Bolter, J.D., Grusin, R. (1999). *Remediation: Understanding New Media*. Cambridge: MIT Press.
17. George Steiner (1991). *Real Presences*. University of Chicago Press
18. Stiegler, B. (2010). *Taking care of youth and generations*. Stanford University Press, Stanford.
19. Serres, M. (2020). *Branches*; Bloomsbury Academic: London, UK.
20. Matthias Vogel (2012). *Media of Reason: A Theory of Rationality*, Darrell Arnold (tr.), Columbia University Press.
21. Virilio, P.; Lotringer, S. (2002). *Crepuscular Dawn*; Semiotext(e): Los Angeles, CA, USA.
22. Barthes, R. (1980). *La chambre claire. Note sur la photographie*. Paris: Gallimard.

23. Kittler, Friedrich, (1999). Gramophone, Film, Typewriter, Stanford, CA: Stanford University Press.

24. Goodman, Nelson (1976). Languages of Art: An Approach to a Theory of Symbols, Indianapolis/Cambridge: Hackett Publishing Co.

25. International Scientific Report on the Safety of Advanced AI. – London: UK Government, Frontier AI Taskforce, January 2025. – 132 p. URL: <https://www.gov.uk/government/publications/international-scientific-report-on-the-safety-of-advanced-ai>

TABLE OF CONTENTS

INTRODUCTION.....	3
The Concept of Multimodality	4
Multimodal System	5
Technical Principles of Multimodal Models.....	9
Multimodal Assistants and Agents.....	12
Ethical Risks of Substituting Perception and Trust	13
Technological Upgrade	16
Content Creation and Media Production.....	16
Benefits of Multimodal AI	17
Challenges of Technological Upgrades	18
Agent-Based AI and Autonomous Systems.....	18
Ethical AI Deployment.....	20
Multimodal Language Models	21
Media Philosophy of Digital Ecosystems	21
Digital Epistemology.....	31
Cyber security	34
Legal Philosophy and Intelligent Automation	41
Digital Law.....	53
Cyber Law	56
TABLE OF CONTENTS	62