Journal of Digital Economy Research. 2023. 1. P. 6–60
DOI: 10.24833/14511791-2023-1-6-60

Received: 1 October 2022
Revised: 7 November 2022
Accepted: 15 December 2022

SOCIO-ECONOMIC ASPECTS OF THE INTRODUCTION OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES

by Maxim V. Fedorov

Abstract. The main objective of the paper is to give an overview of global effects of AI technologies, including socio-ethical principles, direct and non-direct economic impact and regulatory frameworks for developing strategies of sustainable development based on AI technologies. We will discuss these problems considering AI as a part of the global process of technological development, and, therefore, will briefly overview relationships between AI and other close fields (computational technologies, data acquisition techniques etc). A particular focus will be on global risks associated with the intensive use of AI technologies. Special attention will be given to the issues of international standardization of AI and related technologies. A section on AI-based social ranking will discuss fundamental problems inherent for such systems (biases, non-transparency etc). That section will be followed by a section on deepfakes which will be discussed in view of their dramatic effect on the conception of trust, both on individual and population/state levels. The paper will also discuss effects of widespread introduction of AI on other fields of research, such as chemical sciences and molecular biology. We will discuss pathways for sustainable development of "Trustworthy AI" which may achieve the desired balance between the benefits and risks of using these technologies and a global scale. We will discuss approaches that may lead to development of strategic principles for accessing long term effects of AI followed by relevant regulatory approaches.

Keywords: Artificial Intelligence (AI), Machine Learning, Big Data, Trustworthy AI, Deepfakes, Natural Language Processing, Information&Communication Technology (ICT), Social Ranking, Global Aspects of implementation of AI technologies (ethical, social, economical etc.); AI-powered Digital Pharma

For citation: Fedorov Maxim V. (2023) Socio-economic aspects of the introduction of artificial intelligence technologies. *Journal of Digital Economy Research,* vol. 1, no 1, pp. 6–60. (in English). DOI: 10.24833/14511791-2023-1-6-60

Information about the author:

Maxim V. Fedorov – Corresponding Member of the Russian Academy of Sciences, Doctor of Chemical Sciences, Candidate of Physical and Mathematical Sciences, Rector Sirius University of Science and Technology.
354340 Township Sirius, Olimpiyskiy avenue 1.
Ph.: 8 862 241 98 44, E-mail: info@siriusuniversity.ru
ResearcherID: C-7458-2011
ORCID: 0000-0003-3901-3565

Introduction – economic and social impact of AI

The widespread use of artificial intelligence (AI) over the whole planet as a consequence of the accelerated rate of digitalization of society, is leading to the 'smearing' of national borders (European Commission, 2021; Fedorov, Tsvetkov, 2020a; Fedorov, Tsvetkov, 2020b). The huge technological capabilities of AI systems have been implemented in many digital services (social networks, search engines, news aggregators, marketplaces, etc.), significantly expanding capabilities of 'traditional' information&communication technology (ICT) tools (Kurzweil, 2006; Yukhno, 2022a). Today, AI is an important part of the global digital space (Lindre, 2022; Mironova, 2021; Yukhno, 2022b; Yukhno, Umarov, 2022). However, in fact, it is still an unregulated (albeit extremely effective) tool which can provide significant geopolitical and economic advances to owners of global digital platforms and services and change directions of modern consumer, socio-political and socio-cultural trends (Fedorov, Tsvetkov, 2020b; Fedorov, Tsvetkov, 2021a; Lindre, Kapitanov, 2022. In this paper, we will overview potential risks associated with AI and discuss pathways for sustainable development of these technologies, leading to "Trustworthy AI". Analysis of trends in patents and scientific publications will be provided mostly based on analytics by Statista (www.statista.com) and Dimensions (www.dimensions.ai) as well as other sources (will be referred later).

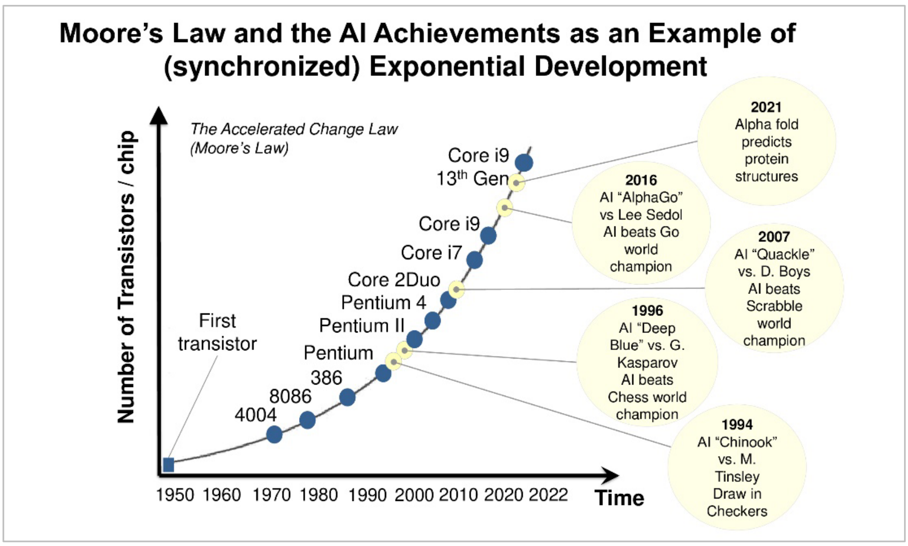

Fig. 1 Moor's law vs AI challenges

One has to note that the term "Artificial Intelligence" is somewhat ambiguous; more than 100 definitions of this term exist in expert communities, and the spectrum of understanding of this term by public is even wider (Miao Fengchun, et al., 2021). Therefore, to avoid confusion, we will follow ideas presented by Kurzweil (Kurzweil,

2006) and make a distinction between "Weak AI" and "Strong AI". By Weak AI (also called Narrow AI) we mean an AI technology limited to a specific task (e.g. face recognition or playing chess etc). We note that weak AI can mimic some particular aspects of human behavior and automate repetitive routine tasks. By Strong AI we will mean the creation of an artificial intelligence that outperforms the capacity that of a regular human being in a number of different tasks simultaneously. We note that it is hard to say whether we are going to achieve Strong AI in the near future and what would be the base of this technology, if developed. In fact, all AI technologies that are known these days can be qualified as Weak/Narrow AI. Therefore, through the manuscript, we will use the term AI meaning that it is the Weak AI if not mentioned otherwise.

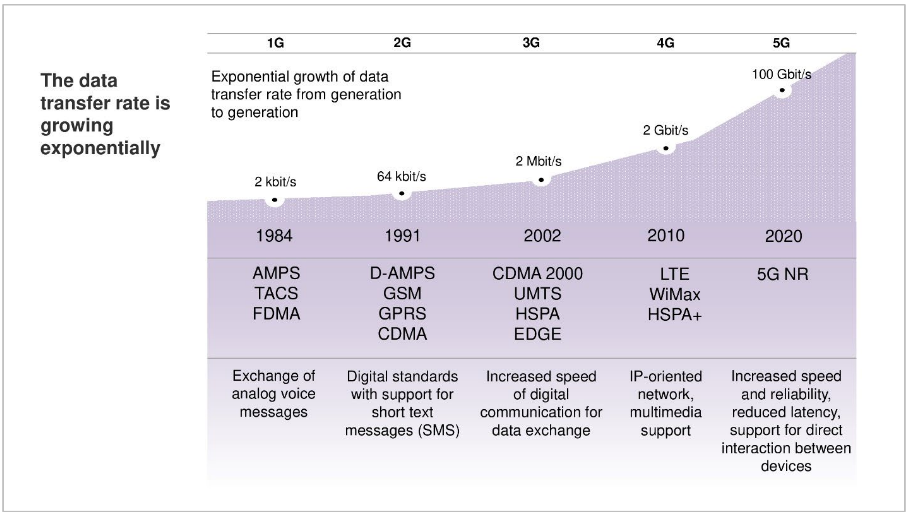

Fig. 2 Illustration of exponential increase of data transfer rate

Since the 1950s, when the term AI for the first time was introduced and became an independent object for research, hundreds thousand inventions have been patented over the world and more than 1.6 million scientific materials on the subject of AI have been published. At the same time, judging by the analytics provided by Statista (www.statista.com) and Dimensions (www.dimensions.ai) more than a half of all these AI-related patents have been published over the past decade. This exponential growth of AI applications strongly correlates with exponential increase of computational power of microprocessors (so-called Moor' law, see Fig. 1) and exponential increase of the rate of data transfer (see Fig. 2). Indeed, AI is strongly interconnected with computational technologies (Pavlov, 2016; Materials, 2020, 2021;) as well as with technologies for processing and transmission of information (see Fig. 3). These three pillars make a synergy together and thus, provide a solid background for a number of applications (Tejal, et al., 2016; Pavlov, 2016;Biamonte, et al., 2017; Sharaev, et al., 2019; Andronov, et al., 2021;Babakov, et al., 2021; Kozlovskii, Popov, 2020; Miao Fengchun, et al., 2021; Morozov, et al., 2021; Khokhlov, et al., 2022) such as E-sport, Chemical Informatics, Robotics, Smart Manufacturing etc (see Fig. 3).

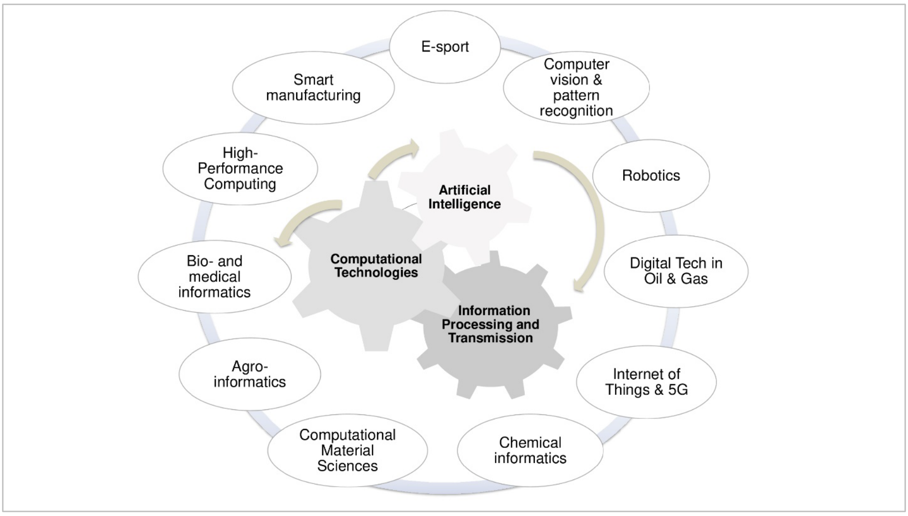

Fig.3 Areas related to AI technology

Indeed (Weak) AI itself strongly relates with data analysis and data processing techniques as well as with supporting technologies such as high-performance computing, sensors etc. Therefore, one may present the modern AI as an assembly of three components: (i) Data, (ii) AI Software (mostly Machine Learning Algorithms and corresponding libraries and frameworks) and (iii) Supporting Technologies (computational and data acquisition and transfer infrastructure etc) – see Fig. 4. The most "intelligent" part of this (arguably) is the Machine Learning frameworks; an example of a typical Machine Learning pipeline is schematically presented on Fig. 5.

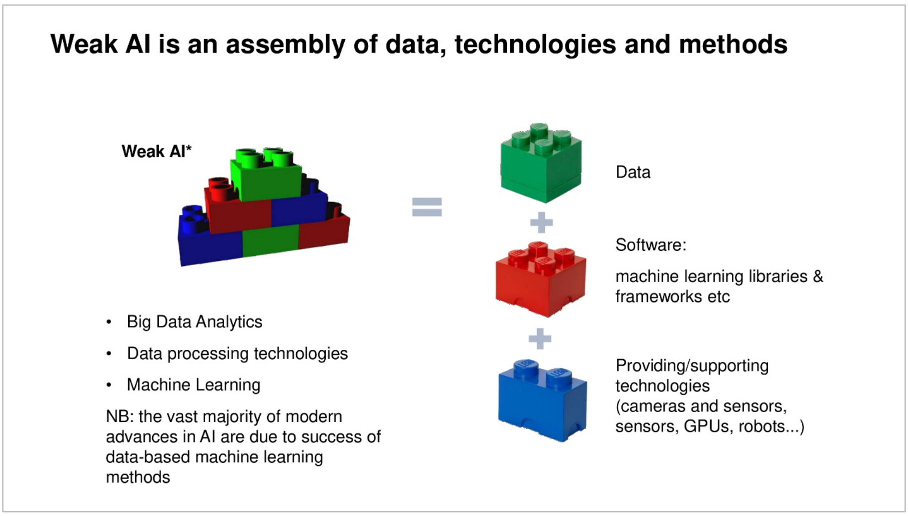

Fig.4 Main components of weak AI

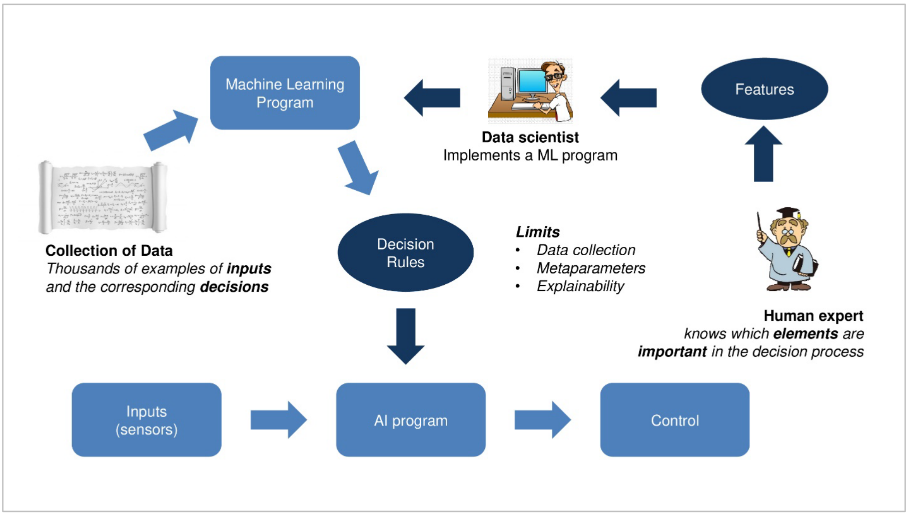

Fig.5 Overview of a Machine Learning pipeline for decision making

Analysis of patents and scientific publications in the filed also demonstrate a steady trend of shifting emphasis from theoretical research to the practical development of AI products and services for commercial purposes. Machine learning, deep learning and new neural network architectures are the most dynamically developing areas of AI technologies in recent years (Miao Fengchun, et al., 2021). Achievements in this area lead to the emergence of automated systems that can learn and perform cognitive tasks previously accessible only to people. In the field of practical applications of AI, computer vision and natural language processing demonstrate most significant progress. However, as presented by Fig. 3, there is a number of promising applications in many other areas as well.

Obviously, such technological development will have far-reaching social and cultural implications (Kurzweil, 2006; UNESCO, 2019; Gurria, 2020; Fedorov, Tsvetkov, 2021a; Fedorov, Tsvetkov, 2021b; UNESCO, 2021;European Commission, 2021). It is necessary to elaborate a system of criteria and a methodological basis for creating strategies of sustainable development for countries and all of humanity in the context of global application and development of AI technologies (UNESCO, 2019; UNESCO, 2021). We note that AI is a General Purpose technology: it transforms all economic sectors and social activities in such a way that:

- AI will boost productivity in most activities it touches;
- AI, like traditional ICT but to a larger extent, will affect income distribution, hence social equilibria;
- AI challenges the very identity of humanity, which is based on intelligence: hence it raises complex societal and ethical issues.

Following ideas of Mironova (Mironova, 2021) on classification of global risks, we may select these main groups of global risks of widespread implementation of AI.

The group of global *geopolitical* risks includes: cyberterrorism using AI, instruments of influence on the domestic and foreign policy of states, violation of the digital sovereignty of states, ever increasing threat of the use of AI tools for the creation of weapons of mass destruction (mainly chemical and biological), etc.

The group of global *economical* risks includes: the use of AI tools to provoke a financial crisis (through manipulation on the stock exchange, cryptocurrencies, etc.), high price instability in the market of computing components, monopolization of the market of computing infrastructure, instability in the labor market and shortage of personnel.

The group of global *technological* risks includes: negative consequences of scientific and technological progress, disruption of leading information systems, exponential complexity of infrastructure, accumulation of errors and accidents, etc.

The group of global *social* risks includes: potential food crisis provoked by (misused) AI-based logistics tools, the possibility of a "digital pandemic", potential real pandemics provoked by new infectious diseases "engineered" by AI, increased migration activity of the population due to remote work, social inequality generated by technology, a huge gap in the standards of living of people in "digital" and "non-digital" regions, distrust of power generated by deepfakes and other AI-based tools for social manipulations.

Existential risks – violation of the balance between management and manipulation in society, the emergence of processes of de-intellectualization and disorientation of people in the information space due to substitution of their cognitive functions by AI, suppression of their biological and psychological needs by AI and related technologies (virtual reality, robots etc).

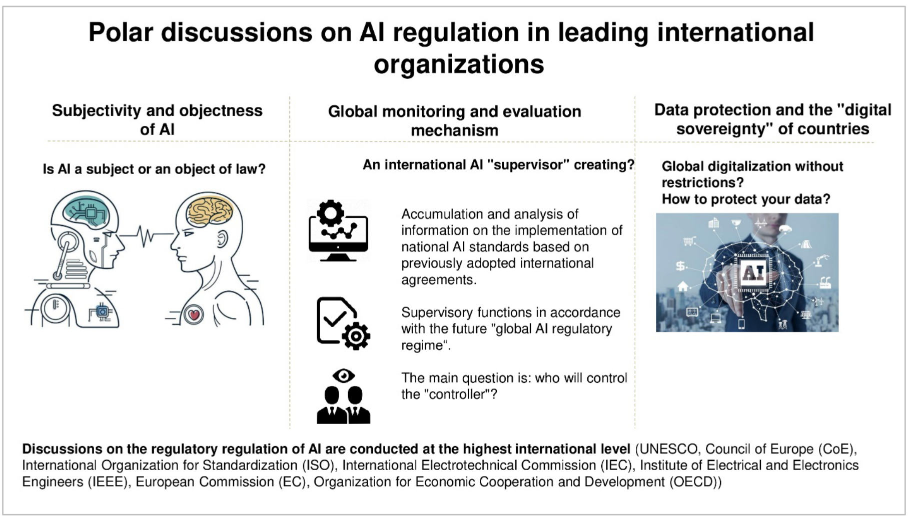

Fig.6 Discussions on the regulation of AI

To summarize, one of the main existential threats is the phenomenon of "dangerous knowledge" (Mironova, 2021). Dangerous knowledge can be defined as information obtained in the course of scientific research, development and analysis of big data by AI, which negative consequences our society cannot always effectively control. This is especially true for the use of AI in such fields as genetic engineering, nuclear physics, chemical sciences and humanitarian technologies. Ultimately, this is a problem of the ethics of science, values and the formation of humanistic ideals in the field of AI technologies. The discussion above leads us to a conclusion, that the almost unregulated development of AI, if it continues without being overseen by a proper set of regulatory instruments, may lead to catastrophic global consequences. However there are very polar debates on many national and international regulatory platforms on very basic principles of regulations and there is no clear solution how to reach the desired balance between safety and technological freedom yet (see Fig. 6). Therefore, in the next section we will discuss main international and national activities on development of regulatory frameworks for AI.

Regulation of AI technologies: still a challenge.

One of the most informative indicators of trends in technology is the content of relevant technological standards and regulatory documents (Materials, 2020, 2021). The analysis of documents created for the standardization and regulation of AI and big data technologies allows us to see the priorities of developments in various countries, as well as draw conclusions about the activities of corporations developing AI technologies.

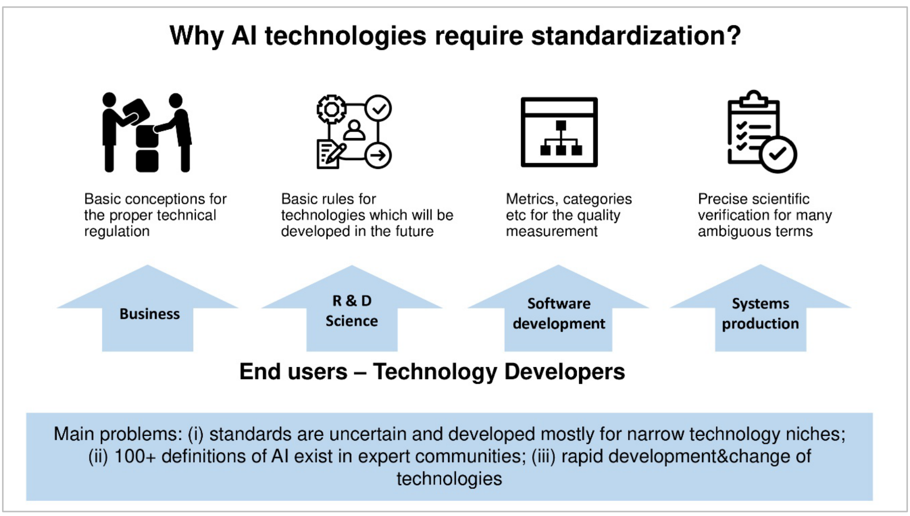

Fig.7 Main barriers between different stakeholders of AI technologies and ways of their settlement through proper standardization mechanisms

The rapid growth of the AI technology market has requested the formation of a regulatory framework for the subsequent regulation of the use of AI-based solutions with the main goal of preventing harmful effects on humans and society as a whole (see Fig. 7). At the same time, a more or less meaningful process of forming the regulatory framework started only in the 2010s, when AI systems had been already introduced in a number of fields.

The key direction in creating an integrated system of international regulation in the field of AI and end-to-end technologies is standardization, the results of which are perceived differently among the vast expert community and various groups of stakeholders (developers, regulators, engineers and users, industrial and academic sectors, government agencies). It is not always possible to maintain a balance of interests, since the technological agenda a priori has many obvious and hidden conflicts of interest that arise when prioritizing most areas of standardization. Large IT companies are essential investors in research in the field of AI standardization, including financing pilot tests and conducting painstaking work on generalizing international expert thought. Almost all editors and developers of AI standards and related documents have affiliations with global multinational corporations such as Microsoft and Alphabet.

The documents required as the output of standardization process mostly have the following format: an international standard, a technical report or a technical specification. In addition, a set of documents may be developed to form a series of standards, if the standardization area covers aspects that are clearly shared by consumers.

The main international initiatives on standardization of AI technologies have been carried out through the Artificial Intelligence Subcommittee (PC 42) and the Joint Technical Committee 1 (OTC 1) "Information Technologies" of the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC)[1] established in 2017. To a much lesser extent, they are discussed and developed in the International Telecommunication Union (ITU). Expert support for the standardization process is also provided by the Institute of Electrical and Electronics Engineers (IEEE).

Currently, the main focus of standards development is shifted to the following areas of application of AI systems: Big Data (DB), Intelligent and Autonomous Vehicles (IATS). A special place in standardization is also occupied by issues of increasing confidence in AI systems, risk management, ethics and problems of AI bias.

Since 2018, separate work has been carried out on the terminological and conceptual standardization of AI systems – the fundamental ISO/IEC document (ISO/IEC 22989 – AI – Concepts and Terminology) for all AI-related standards that have already been adopted or are planned to be adopted.[2]

---

1 https://www.iso.org/committee/6794475.html

2 ISO/IEC DIS 22989 Information Technology — Artificial Intelligence — Artificial intelligence concepts and terminology

The formation of a series of international standards in the field of *databases* is proceeding in leaps and bounds. Despite the long overdue need to create a fundamental terminological database framework that would establish the basis for subsequent documents in this area, other technical areas and related areas of standardization, there is still no comprehensive terminological and conceptual framework. A fairly impressive group of countries (USA, China, Ireland, Korea, Japan, etc.) represented by their scientific expert and business community with objectively different starting approaches to understanding the database are extremely slow to reach some consensus decisions. Standards appear on certain aspects of the database, lagging behind the development time, for example, from documents in the field of quality of AI systems.

On the national arena, in the context of big data, Russia has adopted an analogue of the basic international standard – the preliminary national standard (PNST) – "Information technology. Big data. Typical architecture". This document focuses on the description of the most important terminology and indicates the relevant areas of application.

The technical report, which is essentially an overview of the *parameters of human "trust"* in the field of AI technology applications, was published by ISO/IEC in April 2020. The emergence of human "trust" in AI technologies is linked to ensuring their transparency, explainability, manageability, etc. at three levels: physical, cybernetic and social. As part of the international discussion on standardization, it is assumed that "trust" should be ensured at all stages of interaction and use of AI systems by humans leading to the so-called conception of "Trustworthy AI" (see Fig. 8).

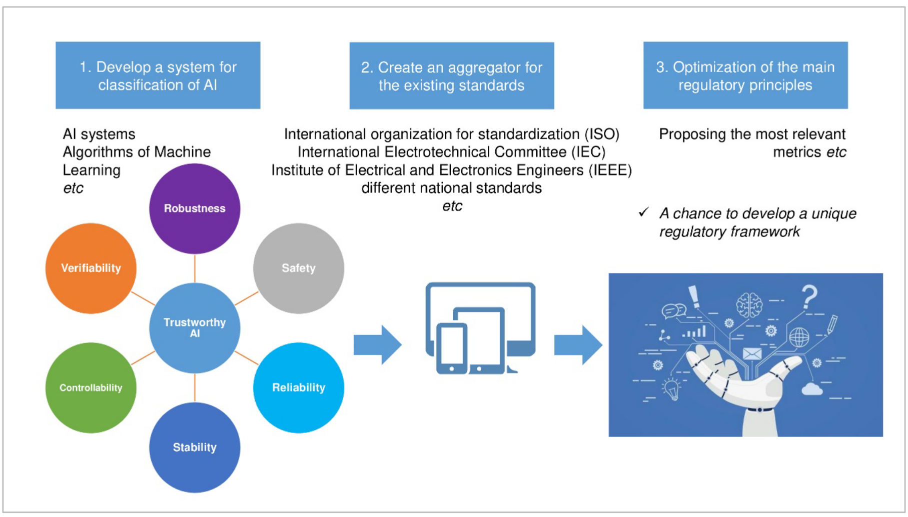

Fig.8 Components of "Trustworthy AI" and a pathway to an optimal regulatory framework
(UNESCO, 2019; UNESCO, 2021; Fedorov, Tsvetkov, 2020)

For complex systems based on machine learning, trustworthiness is generally understood to mean absolutely clear formalizable properties of the systems, such as stability, reliability (aka robustness), guaranteed convergence, security, adversarial robustness – what is known as adversarial robustness, as well as a set of properties like data privacy, that is, differential privacy (Materials, 2020, 2021).

We note that reliability, also known as robustness, is a kind of meta–property that describes the preservation of a certain basic property, such as, for example, the same level of stability in the presence of some uncertainties and disturbances in the system. Another term which is widely used to describe trustworthiness is the adversarial robustness which is resistance to adversarial attacks. The most banal example here is when some minimal distortion of visual information leads to severe misclassification: a car can mistake a pedestrian for a pole or, say, for a fly on the windshield. And it will not be classified as a kind of danger that needs to be avoided, just because the image is a little noisy (Materials, 2020, 2021).

Differential privacy, aka differential privacy, is a property that consists in the fact that the system can provide a certain level of protection of the user's personal data. Fault tolerance means the ability to maintain normal functioning, including stable, performing restrictions, and so on, in the presence of a malfunction in sensors, a malfunction in actuators, that is, control devices, and so on (Materials, 2020, 2021).

Adversarial robustness and differential privacy are relatively new concepts. These are certain properties that can be considered as subtypes of reliability and stability, but they have some specifics. Maybe there will be even more of these indicating properties in the future, taking into account the specifics of AI (Materials, 2020, 2021).

In the near future, flagship documents in the field of AI standardization will be the standard for technical regulation of artificial neural networks and the technical report on their stability characteristics adopted in 2021. The main task of these documents is to regulate the control at the stage of software product development in various industries and the rules for certification of system elements for assessing their stability, with appropriate balance of interests on the part of engineers and manufacturers. For example, the American National Standards Institute (ANSI) made a proposal, to fill in the technical specification of ISO/IEC PC 42 in the field of machine learning goals and methods and AI systems. It lists methods that make it possible to increase the explainability of AI systems for various stakeholders (developers, users, providers, regulators), and the methods differ for each group of stakeholders. For developers, it means strengthening the aspects of security and technical reliability; for users – the degree of trust in the system or method, so that bias does not play a significant role; for providers – compliance with regulatory and technical regulations, and for regulators – understanding the opportunities and limitations when creating innovative frameworks.

Functional safety has become relevant in the course of the development of machine learning methods and related AI technologies, which have built a line of interest in engineering, building a safe world based on innovative solutions. At the 6th plenary session of ISO/IEC SC 42 in April 2020, a project on the functional safety of AI systems was proposed, which, after a short discussion, it was decided to implement in the format of the ISO/IEC SC 42 technical report. Functional security plays an important role in the introduction of mathematical models, with the help of which the application of a particular AI technology is realized. Their properties reflect the compliance of the AI technology and system with peculiar safety rules. If there are explicable and understandable dependencies between the key parameters that determine the behavior of a mathematical function, the correlation between the observed input and output data is obvious, and then the model will be transparent and effective. However, the physical phenomenon described by the model can be very complex, or have a very small scale, or cannot be observed without the influence of experimental data that cannot be described by a scientifically based model. All this causes difficulties in understanding and verifying the model, introduces ambiguities and a certain degree of uncertainty.

The technical report on the functional safety of AI serves as a description of how to control security, security risks and their relation to AI systems and technologies, also giving classes of AI technologies and the level of their possible use. Classes (1–3) reflect the possibility of greater variability when applied in various fields, from the presence of clear recommendations on safety criteria to the inability to create safety requirements at all. Levels (AD) show the degree of influence of the AI component of the system on security in general. In addition, the document used the concept of the level of automation and control, the degree of transparency and explainability of solutions (from a completely transparent system to a "black box"), as well as other aspects related, including, to the physical components of systems.

Thus, functional security can be specific and depend on the application of the AI system. This opens up the possibility of implementing its aspects using different approaches at one or another stage of the life cycle of AI systems.

The above documents were created in ISO/IEC in order to form an extensive and exhaustive base for a standard or specification for testing AI systems.

AI and Social Ranking: main risks.

Perhaps one of the most controversial fields of AI applications is the usage of AI for development of social rating systems (Fedorov, Tsvetkov, 2020a ;Fedorov, Tsvetkov, 2021b). In short, a social rating system can be defined as, a set of *social indicators evaluated by an automated system* (Borodkin, 2004; Meadows, 1998). Based on the output data (ratings or individual rating), state institutions and public institutions form a model of interaction with a particular citizen. We note that many such systems have been developed in education-related domains (Miao Fengchun, et al., 2021). Naturally, as an effective tool for data analysis, AI became a popular instrument for development of different regional and national social ranking systems. However, recent studies (Fedorov, Tsvetkov, 2020a ;Fedorov, Tsvetkov, 2021b; Lindre, 2022) which analyze effects of social ranking systems introduced in different parts of the planet revealed several fundamental problems of such systems which we will briefly overview below.

The main problems of the concept of social ratings are listed below.

1. Opacity of rules and algorithms based on which the social rating system will evaluate human behavior. Modern systems of social rating are not transparent as they typically hinder the information about both development process of rules and criteria for evaluating human actions by an intellectual system as well as names of the developers. That raises the question about the degree of responsibility of the developers and owners of such systems for making decisions that may be of crucial importance for millions of fellow citizens? At the same time, it remains unclear how the social rating system, which uses artificial intelligence (AI) to process huge amounts of citizens' data across the country or a single society, is actually making decisions as most of AI algorithms operate on the principle of a classic "black box". This means that the output results do not imply an explanation of the reasons and circumstances for which the AI system draws conclusions and makes decisions. People can only take for granted the fact that an "intelligent" machine in real time "sorts" them into groups depending on their behavior and committed actions, even if such actions and decisions are not contrary to the law.

2. No one is immune from social deranking. Some lobbyists of the social rating system (naively) believe that following certain patterns of a "good citizen" is a guarantee of finding a person in the upper privileged group. However, like any other complex technology, an AI-based social rating is able not only to track the general background around a person based on a certain system of indicators, but also to go further – to form an "opinion" about such a person, taking into account the multiplicative effect of the perception of human behavior by the whole society. For example, a popular media personality (politician, artist or athlete), on the one hand, has more opportunities to maintain their social rating at a high level due to the cumulative effect of his positive image. On the other hand, with a planned information campaign to discredit such a person, the "intellectual" social rating system will take into account the perception of this person in society that is changing for the worse, and it will be more difficult for such a person to "correct" his rating with "good deeds" later. The system will give an average assessment of a person taking into account "good deeds" and "bad" image in the media space.

3. The social rating system may have a large number of hindered outcomes. The distribution of people into categories depending on the current social rating may have far-reaching consequences in all spheres of life. For example, a person with an insufficiently high rating is unlikely to be hired for a prestigious job, and in the end this person will not even understand why their job application was refused. The fundamental right of every person – the right of freedom to choose one's life trajectory will be proportionally limited to the designated part of the 'phase space' occupied in the person's social rating. An unpaid fine for traffic violation or an unfulfilled "children's program" may manifest itself in the most unexpected way in one's future and play a decisive role in the fate of the person in the implementation of their plans for life.

4. Digital discrimination of people. The adoption of a social rating system at a state level in a country would actually mean the emergence of a "digital code of law" and a "digital procedural code" in this country which will be parallel to the Constitution of the country and other regulatory legal acts (a set of social indicators and algorithms for collecting, analyzing and evaluating information about a person included in the AI system). Therefore, instead of giving people equal rights and responsibilities, such a system would introduce a full-fledged mechanism of discrimination and restriction of human rights which may arise without committing illegal acts, which are such under the current legislation.

There is a direct threat that the social rating will turn into a dictatorship of a certain system of values introduced by unknown programmers into an automated system that works according to some criteria. This, in particular, is a direct violation of the seventh article of the Constitution of the Russian Federation, which guarantees the free development of a person.

5. The social rating system may lead to social destabilization. The introduction of even individual elements of social rating can be very criminogenic – because such a system will automatically provide illegal ways of "tweaking" the rating for a moderate fee and ending up in a category of higher level. In addition, the social rating is a clear provocation of public tension. It will obviously cause a protest mood in society, which directly contradicts the modern understating of the function of a democratic state where one of the most important tasks is to create conditions for mutual trust between it and society.

In this context, the question of a person's "transit" from one category of social rating to another remains open. Such systems provide an opportunity for a direct threat of legal "enslavement" of a person, i.e. a lifetime at the bottom of the rating pyramid due to an objective narrowing of opportunities for the latter to improve their performance. Initially, the discrimination-oriented concept of social rating will infringe on the rights of persons who, although they do not violate the law, do not "earn" enough points of a "good citizen".

It is noteworthy that the risks of social rating systems were identified and realized by the governing structures of the European Union at an early stage. In April 2021, the European Commission proposed to ban the introduction of AI technologies (resolution "On the European Approach to artificial Intelligence"), which are used for "mass surveillance, applied in a generalized form to all individuals without any differences." Surveillance methods such as "monitoring and tracking of individuals in a digital or physical environment, as well as automatic aggregation and analysis of personal data from various sources" will become illegal[3].

[3] https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

If this initiative is approved by the European Parliament and the European Council, the European Union will completely ban the use of "high-risk" AI systems and restrict the use of others if they do not meet the new standards. It is noteworthy that the "high risk" category includes, among other things, AI technologies in robotic surgery, AI in software for hiring employees, AI for checking documents and verifying evidence (in court proceedings), as well as AI for assessing the *credit rating* of citizens – the primary element of the entire social rating system. Moreover, the explanation to the European Commission document notes that the very essence of social rating and the use of AI in applications that manipulate human behavior to circumvent their will are unacceptable (European Commission, 2021). Business representatives who ignore these new rules may be fined up to 20 million euros or 4% of the annual turnover.[4]

We believe that it is time to initiate full-fledged scientific research on the phenomenon of social rating on global scale. The results obtained should be submitted for a wide public discussion with the participation of representatives of all sectors of society (scientific and expert community, human rights organizations, legislators, business circles, political parties, civil society, etc.).

Deepfakes – what is truth?

Deepfakes can be defined as synthetic auditory or visual media developed using deep machine learning methods that create such a realistic impression that they can deceive the audience (Hutchinson, 2020; Jaiman, 2020; Lindre, Kapitanov, 2022; Prajakta, 2020). This is the main task of the authors of deepfakes – to make as realistic illusion as possible. Synthetically created media vary widely in technical complexity and application, ranging from low-quality "cheap deepfakes" to high-quality products that can influence a person's "perception of reality", as well as in a certain sense alter the decision making process of the person

Of course, the overall idea of creating of synthetic audiovisual content is not new: Filmmakers have been using computer-generated images (CGI) since the 1970s in order to enhance. However, these methods were very expensive and only a few companies had access to such technologies. These days, the development of digital technologies has made complex synthetic media inexpensive and easy to produce (especially thanks to the proliferation of free and open source software). As a consequence, millions of people create such content every day using their home computers and even smartphone (Davis, 2020).

4 https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

Deepfakes can be successfully and usefully used in such areas as education and art, but it turned out that deepfakes are more likely to harm individuals, business, society and democracy as a whole, and also accelerate the process of falling public confidence in the media (Jaiman, 2020;Logacheva, 2021; Babakov, et al., 2021; Dementieva, Panchenko, 2021). Such an erosion of trust will contribute to the flourishing of de facto relativism, destroying the structure of democracy and civil society under strong pressure of false information.

The grand scale of using the deepfakes by many stakeholders (including state agencies, companies and individuals) pose a number of challenges to modern society. The most growing general concern is about how deepfakes contribute to the evolution of the era of Internet disinformation (Perez, 2019). Indeed, as tools for creating synthetic information for a mass audience are being constantly improved at the technical level, regulatory agencies across the globe are trying to solve simultaneously emerging socially significant problems associated with an adequate response to new risks and threat provide by deepfakes (mainly the high rate of distrust to public media).

For instance, there is an important question on protecting the rights of individuals in the sphere of control over the commercial use or misuse of their images and other aspects of "digital" identity (Jaiman, 2020). The methodology of creating deepfakes has already developed to such an extent that many people may fully recreate their image or somebody's else image in digital form in order to use these 'updated' images in various kinds of virtual projects in which they do not take direct "personal" participation (Davis, 2020). In this case what would be the fee for the commercial use of their images and the image in general, who will posses the rights of using the new image (which can be of considerable value)?

Even more importantly, in the framework of legal proceedings, deepfakes can pose a serious threat to determining the authenticity of the most important evidence, since the continuous development of technology significantly complicates the process of separating real digital evidences (images, video, audio etc) from forgery and, therefore compromises the use of digital evidences (Kietzmann, et al., 2019; Prajakta, 2020). For example, some courts in the United States have allowed taking photographic evidence without verifying the reliability of eyewitness testimony in accordance with the so-called the "silent witness rule" (the use of "substitutions" when accessing confidential information in the United States in the open jury trial system). While maintaining the current dynamics of the spread of deepfakes, it may become a common practice for lawyers to declare that the digital evidence against their client is a fake. This circumstance may cause the jury to doubt the authenticity of real evidences (Prajakta, 2020).

Another problematic area of using deepfakes is the field of cybersecurity (Gurria, 2020). Indeed, falsified data and digital trace of a person provide unprecedented risks. For example, sensor data may be falsified and then transferred to an AI decision making systems, which will lead to the deception of the artificial intelligence system with the subsequent adoption of incorrect decisions by it (Davis, 2020).

A comprehensive system of protecting society from the destructive use of deepfakes begins with software designed to detect deepfakes or increase the technical ability to distinguish between real and fake content. Such technical solutions include both AI systems trained to recognize anomalies characteristic of deepfakes, and cryptographic methods that can be integrated into video and audio recording equipment in order to detect unauthorized access. However, the means of detecting and countering deepfakes will always lag behind the pace of development of the deepfake technologies themselves. In addition, the solution to the following problem is not yet visible: simply knowing that a certain video file has been processed does not provide information on how to identify the creators and hold them accountable (Davis, 2020; Jaiman, 2020).

The second level of counteraction to destructive deepfakes is legislative regulation. For instance, a number of states in the United States, for example, Virginia, have adopted local laws criminalizing deepfakes used in pornography; and the state of Texas has additionally criminalized deepfakes created to influence the electoral process. From another side, the states of Massachusetts and California have not been able to approve bills aimed at countering deepfakes, due to concerns about excessive "overregulation" of the IT industry (Davis, 2020; Jaiman, 2020).

Perhaps, the best result would be the development of technical and legal tools for managing the risks of harm from deepfakes that do not suppress innovation and business activity, as well as do not encroach on freedom of expression. Responsibility should come for individuals without the introduction of universal bans on new technology.

An urgent and practically unsolvable problem in most countries of the world today is that state, primarily regulatory authorities, do not take sufficient actions to ensure the authenticity of audiovisual content distributed in the public space. Encouraging digital platforms to self-censor and self-identify destructive content using deepfake technologies has a very limited effect. For example, if the origin of the video cannot be traced, national governments are recommended to create a body that can detect deepfakes using blockchain technology. The essence of the experts' proposal is that block chains store data in a decentralized network in which anyone can verify the originality of the information by comparing it with a certain unique and unchangeable key. Even the slightest manipulation of data will lead to an easily detectable discrepancy.

The technological approach to managing and limiting malicious deepfakes is to develop methods and methods for authenticating authentic content. An example is a technology called AmberAuthenticate, which works on devices that reproduce original, i.e. authentic photos, audio and video content in real time as content is recorded. The program creates the so-called "truth layer", which is the original content cryptographically marked with numerous digital fingerprints, and then archived in a publicly accessible block chain. This fingerprinting of digital content is used to track its origin as it spreads in the digital space and to help detect and respond to attempts to manipulate the original content.

In another part of the planet, Cyberspace Administration of China has developed rules (enforced on January 1, 2020) prohibiting any publication in the digital information space with modified content created using AI technologies without appropriate warning. These measures were taken for reasons of national security in order to prevent the replication of fake news and the spread of political disinformation at an early stage (Jaiman, 2020).

The new rules allow law enforcement agencies to prosecute individuals who create deepfakes and video platforms on which they are posted. The document does not specifically mention the term deepfake — instead, it is prohibited to publish generally misleading content that could damage the reputation, for example, of a political figure or deceive voters (Babakov, et al., 2021). At the same time, the rules for regulating deepfakes are lengthy in nature, for violation of which no specific administrative or criminal measures are formally provided (Jaiman, 2020).

At the end of the section we would like to cite the proposals from the study of Indiana University "Deepfakes: Trick or treat?" on how to regulate deepfakes and combat malicious content based on this technology (Kietzmann, et al., 2019).

1. Fixing the source content that guarantees the detection of the unreliability of the deepfake. An effective fight against deepfakes created for the purpose of deception or manipulative influence on a person involves a system of early detection and blocking of such content. In this regard, technologies that track and record a person's life, including his geolocation, interaction with other people and institutions, as well as information about any other "external" activity, are becoming in demand. Despite the direct threat to privacy, privacy and confidentiality of human actions, from a technical point of view, collecting such information from modern mobile devices seems quite simple and convenient. However, such personal and sensitive data should be encrypted, stored and used for comparative analysis only if necessary in order to "expose" deepfakes.

2. Exposing malicious deep fakes. Along with the development of AI technologies that simplify the process of creating and improving the final result of audiovisual deepfakes, there are also technological innovations being developed for the detection and classification of deepfakes. The leaders in this area are the United States. For example, the Agency for Advanced Research Projects of the US Department of Defense (DARPA) has innovative means of detecting deepfakes, in particular, the program for forensic examination of digital media content, as well as deepfake detection services such as Truepic. Global corporations and digital platforms are also investing serious resources in the identification and detection of deepfakes.

3. Protection of rights within the framework of the law. Victims of deepfakes must have a system of protection provided by law against material or moral harm caused to them in the event of defamation, malicious intent, violation of privacy or emotional distress caused by deepfakes, as well as in cases of copyright infringement, imitation and fraud related to deepfakes.

4. Taking confidence-building measures. Global companies that claim to adhere to universal moral principles and human rights, in the context of deepfakes, should redouble their efforts to build trust with their customers and establish strong emotional ties with them. In this case, if companies or their brands become involved in scandals on the basis of diplomatic threats, traditional customers and partners are likely to take the information noise critically and maintain their loyalty.

AI in Digital Pharma and Molecular Biology

Pharmaceutical sciences and molecular biology undergo a “phase transition” due to the widespread usage of data-intensive machine learning techniques in these fields (Sosnin, et al., 2018b; Yang, 2018; Kozlovskii, Popov, 2020; Sosnina, et al., 2020; Zaretckii, et al., 2022). From a distant point of view, it might seem strange because these fields have been traditionally associated with heavy usage of experimental techniques and clinical trials. However, modern research programs in biomolecular sciences have at their cores experimental (e.g. data from clinical trials) as well as computational data-generators (e.g. data on molecular modelling); thus data analysis and machine learning research based on the generated data may greatly contribute to the progress in these fields. Digital Pharma is one of the remarkable examples, where in vitro or in silico High-Throughput-Screening platforms are used to derive predictive models which can be applied to solve fundamental and practical problems in modern drug discovery and molecular design (Sosnin, et al., 2018b; Sosnina, et al., 2020; Andronov, et al., 2021; Khokhlov, et al., 2022).

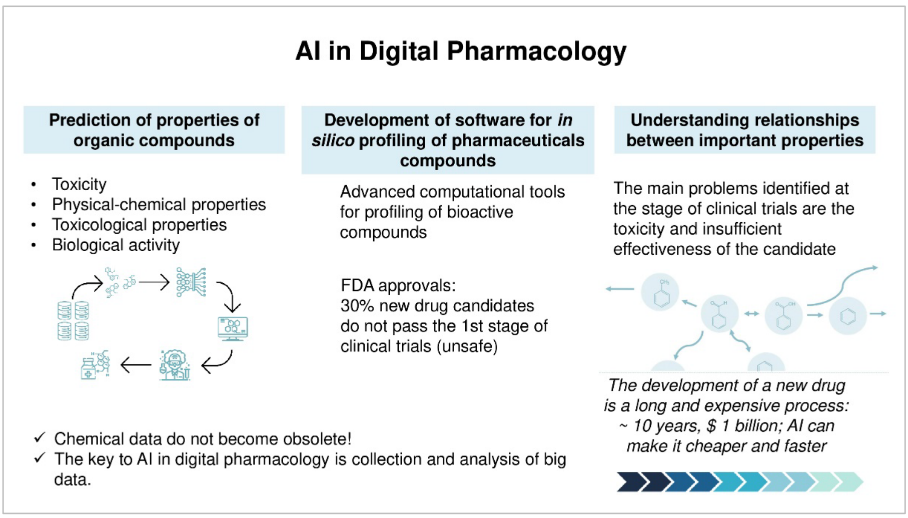

Fig.9 Digital pharmacology problems
(Andronov, et al., 2021; Karlov, et al., 2019; Karlov, et al., 2020; Khokhlov, et al., 2022; Krasnov, et al., 2021; Sosnin, et al., 2018a; Sosnin, et al., 2018b; Sosnin, et al., 2018c; Sosnina, et al., 2020d)

Every year pharmacological companies across the globe spend hundreds billions of US dollars on drug development and design. To cope with the crisis related to the exponential increase of the costs of R&D in these areas, one needs to develop approaches, and data-intensive molecular technologies hold a great promise for this. Indeed, recently a number of works on applications of AI in digital pharma and molecular biology have been published in the high-ranking journals which provide excellent examples of a cross-over between state-of-the-art data-intensive computational methods and "wet" biology (Young, et al, 2018; Popov, et al., 2019a; Popov, et al., 2019b; Popov, et al., 2019c; Karlov, et al., 2020; Kozlovskii, Popov, 2020; Kozlovskii, Popov, 2021a; Kozlovskii, Popov, 2021b; Morozov, et al., 2021; Zaretckii, et al., 2022).

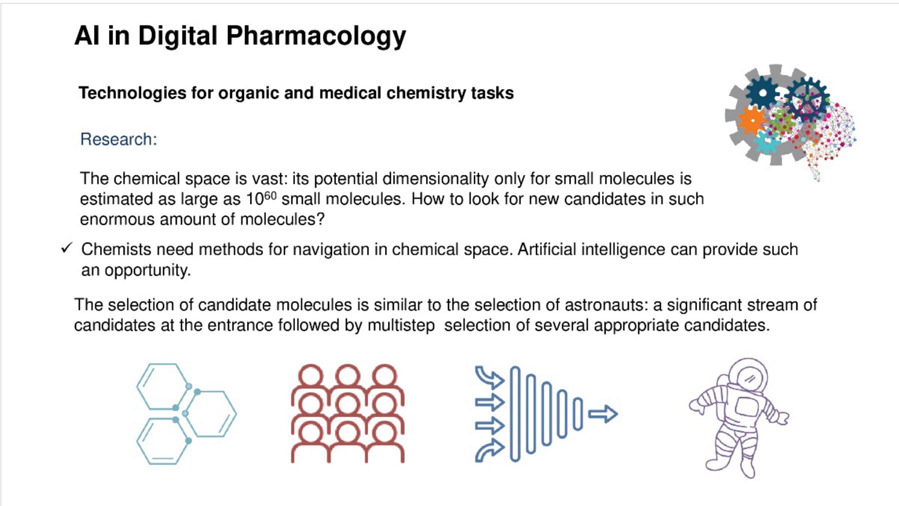

Fig.10 Tasks for AI in digital pharmacology

Progress in experimental techniques have already resulted in accumulation of a large amount of corresponding data and the speed of generation of new data is still growing (Sosnin, et al., 2018b). Effective integration of large-scale data analysis with "wet" biomolecular sciences require development of new approaches and new tools, as well as qualified researchers and engineers with both, computer science and chemical/biological backgrounds, to integrate AI-based tools with experimental research programs. Moreover, the chemical space has a complex structure and vast dimensionality; therefore, straightforward machine learning pipelines typically have a very limited applicability domain, making them useless in practice. Using state of the art data-intensive statistical analysis techniques along with molecular modelling and machine learning approaches allows one to expand the applicability domain of 'pure' machine learning methods and develop more robust solutions (Sosnin, et al., 2018a). This strategic area aims to develop the end-to-end technologies using knowledge-, sequence-, graph-, and structure- and multidimensional-based representations of intensive molecular data (Andronov, et al., 2021).

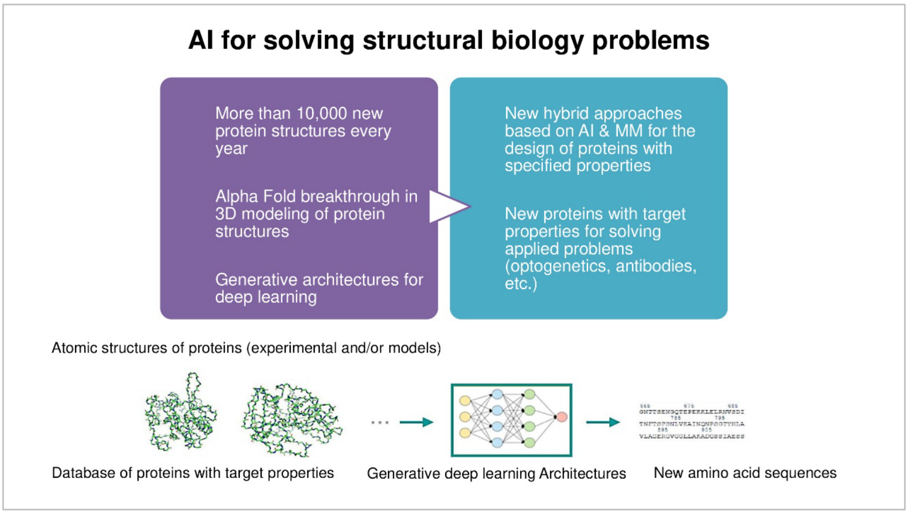

Fig.11 Structural biology problems

(Popov, et al., 2019a; Popov, et al., 2019b; Popov, et al., 2019c; Karlov, et al., 2020; Kozlovskii, Popov, 2020; Kozlovskii, Popov, 2021a; Kozlovskii, Popov, 2021b; Morozov, et al., 2021; Zaretckii, et al., 2022)

The two key strategic directions in the field are:

- develop and apply AI-based tools to explore target regions in chemical space and generate new chemicals with desired properties;
- develop and apply AI-based tools to explore target regions in biological space with a focus on molecular biology and molecular mechanisms of drug-macromolecule and macromolecule-macromolecule interactions.

Figures 9–10 illustrate the first approach. Figures 11–12 illustrate the second approach.

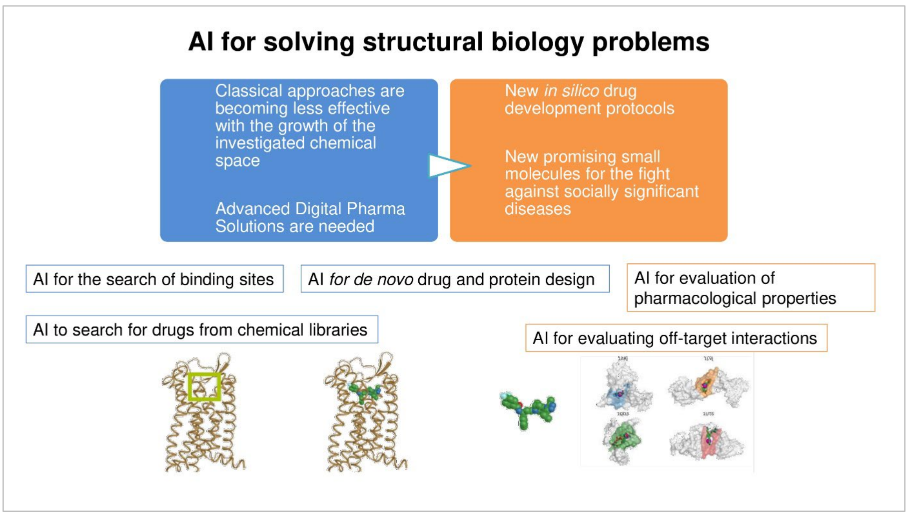

Fig.12 Tasks for AI in structural biology

Potentially, one may estimate that AI applications in bimolecular sciences may (i) reduce the costs of R&D for development and design by 30–50 percent; (ii) reduce the time for development of new drugs and vaccines by 20 (pessimistic) to 70 (optimistic) percent; (iii) greatly reduce potential risks of new drugs (side effects, stability issues etc) due to the thorough in silico validation of drug candidates. Therefore, if widely introduced, these tools can potentially save equivalents of hundreds billion of US dollars per year and, more importantly, save millions of human lives and greatly contribute to the improvement of human health on the whole planet. However, despite of the positive effects of the already many applications of AI-based techniques in digital molecular sciences, there are raising concerns about potential risks related with the implementation of these tools. For instance, recent study has shown that AI-based methods can be used for rapid generation of highly poisonous compounds (Urbina, et al., 2022). Therefore, the dual nature of AI again reveals itself in these areas as well, which raise questions about proper regulations of applications of AI methods in interdisciplinary domains.

Discussion

Due to the many social and economic effects of these technologies, a number of AI-related topics been included in the political and humanitarian agenda of the UN in recent years (UNESCO, 2019; UNESCO, 2021). A significant role in this was played by the Sustainable Development Goals (SDGs) adopted at the summit of the World Organization in 2015, which are a list of 17 global goals and 169 tasks (United Nations, 2015). The successful implementation of most of them is connected with the need to consolidate the efforts of mankind to achieve the new level of development of the society. It is assumed that the transition from the third to the fourth industrial revolution (characterized by the development and mass introduction of cyber-physical systems into global production) will allow one to accumulate and use additional resources in order to stimulate the growth of the world economy, the widespread elimination of poverty, reducing the negative human impact on the environment, etc. The prerequisites for such a transformation are the introduction of a fundamentally new technological base into all key spheres of society's life that can accelerate the evolution of the main socio-economic processes, as well as create new ones and increase the efficiency of existing production chains. The leading role in this is given to "smart" high-tech systems that use AI or its individual components in their work. However, as stated in the introduction, mass-scale introduction of AI poses a number of existential risks, which may affect the fate of the whole civilization. At the same time, in the area of assessment of risks associated with AI we need to transfer from discussions and postfactum observations to fundamental research aimed to development of a set strategic principles for proactive risk assessment and modelling/evaluating corresponding measures to mitigate these risks. The tasks is challenging indeed because of the exponential speed of development of AI and Big Data technologies (see Fig. 1 and Fig. 2 for illustration).

However, one may learn from the history of pharmaceutical sciences which during the 20–21st centuries have evolved from an unregulated wild field to the stage where every steps of the drug development process are strictly regulated based on both, science&technology as well as ethics. It is timely to remember the oldest principle of biomedical ethics, *primum non nocere* ('first, do no harm'), which can be applied as the main approach for regulating development of AI technologies. Overall, despite of the many distinctive features of AI, the overall problem of regulating potentially dangerous technologies (e.g. nuclear tech, cloning, development of bioactive compounds etc) is not something unique for the humankind. Many useful practices and policies from other fields can be adopted which can optimize resources spent on development of regulatory framework for AI. Such an approach may lead to the development of an adaptive regulatory framework which will be aligned with the principles of transparency, cooperation, accountability and ethics in order to facilitate safe and sustainable development of AI-based innovative products.

Gratitude

The author acknowledges the great help of Ksenia Polyakova (Sirius University) for technical assistance with preparation of the manuscript. Numerous discussions with Yuriy Landre (Skoltech) and Anna Zhdanova (Skoltech) on different aspects of global effects of AI are very much appreciated; many ideas presented in the manuscript have been inspired by these discussions. Many thanks are to Natalia Struskevich (Skoltech), Petr Popov (Skoltech) and Sergey Sosnin (Syntelly) for joint works and discussions on AI in biomolecular sciences and digital pharma. The author is thankful to the many of his colleagues from Skoltech and to the members of the Committee on Ethics of Artificial Intelligence at the Commission of the Russian Federation for UNESCO for discussions on different AI-related topics. Special thanks are to Alexander Kuleshov, President of Skoltech (Chair of the Committee). The author appreciates discussions with Igor Ashmanov (Kribrum), Natalia Kasperskay (Infowatch) and Stanislav Ashmanov (Nanosemantica) on deep fakes and risks associated with digital technologies.

References

1. Andronov M., Fedorov M., Sosnin M. "Exploring Chemical Reaction Space with Reaction Difference Fingerprints and Parametric t-SNE". // ACS Omega. – 2021. URL: https://pubs.acs.org/doi/10.1021/acsomega.1c04778

2. Babakov N., Logacheva V., Kozlova O., Semenov N., Panchenko A. "Detecting Inappropriate Messages on Sensitive Topics that Could Harm a Company's Reputation". // Trustworthy AI Conference speech. – 2021. URL: https://www.aclweb.org/anthology/2021.bsnlp-1.4/

3. Biamonte J., Wittek P., Pancotti N., Rebentrost P., Wiebe N. & Lloyd S. "Quantum machine learning". // Nature. – 2017. URL: https://www.nature.com/articles/nature23474?error=cookies_not_supported&code=8a773c58-d06c-4938-a3a4-ea5ef568463a

4. Borodkin F.M., “Social indicators — what are they?”, “Universe of Russia”. // Mir Rossii. – 2004. URL: https://cyberleninka.ru/article/n/sotsialnye-indikatory-chto-eto-takoe/viewer

5. Davis R. “Technology Factsheet: Deepfakes”. // Policy Brief, spring. – 2020. URL: https://www.belfercenter.org/publication/technology-factsheet-deepfakes

6. Dementieva D., Panchenko A. “Cross-lingual Evidence Improves Monolingual Fake News Detection”. // Association for Computational Linguistics. – 2021. URL: https://aclanthology.org/2021.acl-srw.32/

7. European Commission, “Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in AI”. – 2021. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

8. Fedorov M., Tsvetkov Yu. “Ethical issues of artificial intelligence technologies – how to avoid the fate of the Tower of Babel”. // CDO2DAY; Digital Russia. – 2020. URL: https://cdo2day.ru/vid-sverhu/jeticheskie-voprosy-tehnologij-iskusstvennogo-intellekta-kak-izbezhat-sudby-vavilonskoj-bashni/, https://d-russia.ru/jeticheskie-voprosy-tehnologij-iskusstvennogo-intellekta-kak-izbezhat-sudby-vavilonskoj-bashni.html

9. Fedorov M., Tsvetkov Yu. “Ethics of artificial intelligence in UNESCO activities”. // RIAC. – 2020. URL: https://cdo2day.ru/vid-sverhu/jeticheskie-voprosy-tehnologij-iskusstvennogo-intellekta-kak-izbezhat-sudby-vavilonskoj-bashni/

10. Fedorov M., Tsvetkov Yu. “Artificial intelligence as a tool of struggle for people’s consciousness”. // RIAC. – 2021. URL: https://russiancouncil.ru/analytics-and-comments/columns/cybercolumn/iskusstvennyy-intellekt-kak-instrument-borby-za-soznanie-lyudey/?sphrase_id=94902947

11. Fedorov M., Tsvetkov Yu. “Artificial intelligence and social rating: the beginning of the era of the digital concentration camp “in the humanity interests”?”. // RIAC. – 2021. URL: https://russiancouncil.ru/analytics-and-comments/analytics/iskusstvennyy-intellekt-i-sotsialnyy-reyting-nachalo-epokhi-tsifrovogo-kontsentratsionnogo-lagerya-v/?sphrase_id=94902947

12. Gurria A. speech “Munich Cyber Security Conference (MCSC) Fail Safe-Act Brave: Building a Secure and Resilient Digital Society”. // OECD Secretary-General. – 2020. URL: https://www.oecd.org/about/secretary-general/fail-safe-act-brave-building-a-secure-and-resilient-digital-society-february-2020.htm

13. Hutchinson A. “Snapchat and TikTok are Both Reportedly Working on New ‘Deepfake’ Type Features”. // Content and Social Media Manager. – 2020. URL: https://www.socialmediatoday.com/news/snapchat-and-tiktok-are-both-reportedly-working-on-new-deepfake-type-feat/569792/

14. Jaiman A. “Debating the ethics of deepfakes”. // Digital Frontiers. – 2020. URL: https://www.orfonline.org/expert-speak/debating-the-ethics-of-deepfakes/

15. Jumper J., et al. “Highly accurate protein structure prediction with AlphaFold”. // Nature. – 2021. URL: https://www.nature.com/articles/s41586-021-03819-2

16. Karlov D., Sosnin S., Tetko I., Fedorov M. “Chemical space exploration guided by deep neural networks”. // RSC Advances. – 2019. URL: https://pubs.rsc.org/en/content/articlelanding/2019/RA/C8RA10182E

17. Karlov D., Sosnin S., Fedorov M., Popov P. “GraphDelta: MPNN Scoring Function for the Affinity Prediction of Protein–Ligand Complexes”. // ACS Omega. – 2020. URL: https://pubs.acs.org/doi/10.1021/acsomega.9b04162

18. Khokhlov I., Krasnov L., Fedorov M., Sosnin S. "Image2SMILES: Transformer-Based Molecular Optical Recognition Engine". // Chemistry Europe. – 2022. URL: https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/cmtd.202100069

19. Kietzmann J., Lee Linda W., McCarthy Ian Paul, Kietzmann Tim C. "Deepfakes: Trick or treat?". // Business Horizons. – 2019. URL: https://www.researchgate.net/publication/338144721_Deepfakes_Trick_or_treat

20. Kozlovskii I., Popov P., "Spatiotemporal identification of druggable binding sites using deep learning". // Nature. – 2020. URL: https://www.nature.com/articles/s42003-020-01350-0

21. Kozlovskii I., Popov P., "Protein–Peptide Binding Site Detection Using 3D Convolutional Neural Networks". // J. Chem. Inf. Model. – 2021. URL: https://pubs.acs.org/doi/full/10.1021/acs.jcim.1c00475

22. Kozlovskii I., Popov P, "Structure-based deep learning for binding site detection in nucleic acid macromolecules". // NAR Genomics and Bioinformatics. – 2021. URL: https://academic.oup.com/nargab/article/3/4/lqab111/6441762?login=false

23. Krasnov L., Khokhlov I., Fedorov M., Sosnin S. "Transformer-based artificial neural networks for the conversion between chemical notations". // Scientific Reports. – 2021. URL: https://www.nature.com/articles/s41598-021-94082-y

24. Kurzweil R. "The Singularity Is Near: When Humans Transcend Biology". // Penguin Books. – 2006. URL: https://www.amazon.com/Singularity-Near-Humans-Transcend-Biology/dp/0143037889

25. Lindre Yu., Kapitanov A. "Deepfake: an innocent technology for entertainment or a threat to modern society?". // RIAC. – 2022. URL: https://russiancouncil.ru/analytics-and-comments/analytics/dipfeyk-nevinnaya-tekhnologiya-dlya-razvlecheniya-ili-ugroza-sovremennomu-obshchestvu/

26. Lindre Yu. Collection of articles for the 12th Russian Internet Governance Forum RIGF 2022: "Fragmentation of the Internet space and control of information flows – necessary conditions for effective manipulation of public consciousness in the interests of global IT giants". – 2022.

27. Logacheva V., "Detoxification: NLP task showcase". // Trustworthy AI Conference speech. – 2021.

28. Materials of the "Trustworthy AI" conference. – 2020, 2021. URL: https://events.skoltech.ru/ai-trustworthy

29. Meadows D., "Indicators and Information Systems for Sustainable Development". // The Sustainability Institute. – September 1998. URL: https://edisciplinas.usp.br/pluginfile.php/106023/mod_resource/content/2/texto_6.pdf

30. Miao Fengchun, Holmes, Wayne, Ronghuai Huang, Hui Zhang "AI and education: guidance for policy-makers". // UNESCO. – 2021. URL: https://unesdoc.unesco.org/ark:/48223/pf0000376709

31. Mironova N.V. "Specificity of Civilizational Risks and Mechanisms to Overcome Them". // Manuscript. – 2021. URL: https://cyberleninka.ru/article/n/spetsifika-tsivilizatsionnyh-riskov-i-mehanizmy-ih-preodoleniya/viewer

32. Morozov A., Zgyatti D., Popov P. "Equidistant and Uniform Data Augmentation for 3D Objects". // IEEE Access. – 2021. URL: https://ieeexplore.ieee.org/document/9662385

33. Pavlov A. "Architecture of Computing Systems". // ITMO University. – 2016. URL: https://books.ifmo.ru/file/pdf/2074.pdf

34. Perez S. "Twitter drafts a deepfake policy that would label and warn, but not always remove, manipulated media". // TechCrunch. – 2019. URL: https://techcrunch.com/2019/11/11/twitter-drafts-a-deepfake-policy-that-would-label-and-warn-but-not-remove-manipulated-media/

35. Popov P., Kozlovskii I., Katritch V. "Computational design for thermostabilization of GPCRs". // Current opinion in structural biology. – 2019. URL: https://www.sciencedirect.com/science/article/pii/S0959440X18301374

36. Popov P., Grudinin S., Kurdiuk A., Buslaev P., Redon S. "Controlled-advancement rigid-body optimization of nanosystems". // Journal of computational chemistry. – 2019. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.26016

37. Popov P., Bizin Ilya, Gromiha Michael, Kulandaisamy A., Frishman Dmitrij "Prediction of disease-associated mutations in the transmembrane regions of proteins with known 3D structure". // PloS one. – 2019. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0219452

38. Prajakta P. "AI Deepfakes. The Goose Is Cooked?". // University of Illinois Law Review. – 2020. URL: https://illinoislawreview.org/blog/ai-deepfakes/

39. Sharaev M., Smirnov A., Melnikova-Pitskhe T. "Functional Brain Areas Mapping in Patients with Glioma Based on Resting-State fMRI Data Decomposition". // IEEE. – 2019. URL: https://ieeexplore.ieee.org/abstract/document/8637450

40. Sosnin S., Misin M., Palmer D., Fedorov M. "3D matters! 3D-RISM and 3D convolutional neural network for accurate bioaccumulation prediction". // Journal of Physics: Condensed Matter. – 2018. URL: https://iopscience.iop.org/article/10.1088/1361-648X/aad076

41. Sosnin S., Vashurina M., Withnall M., Karpov P., Fedorov M., Tetko I. "A Survey of Multi-task Learning Methods in Chemoinformatics". // Molecular Informatics. – 2018. URL: https://onlinelibrary.wiley.com/doi/10.1002/minf.201800108

42. Sosnin S., Karlov D., Tetko I., Fedorov M. "Comparative Study of Multitask Toxicity Modeling on a Broad Chemical Space. // J. Chem. Inf. Model. – 2018. URL: https://pubs.acs.org/doi/10.1021/acs.jcim.8b00685

43. Sosnina E.A., Sonin S., Nikitina A., Nazarov I., Osolodkin D., Fedorov M. "Recommender Systems in Antiviral Drug Discovery". // ACS Omega. – 2020. URL: https://pubs.acs.org/doi/10.1021/acsomega.0c00857

44. Tejal A. Patel MD, et al. "Correlating mammographic and pathologic findings in clinical decision support using natural language processing and data mining methods". // ACS Journals. – 2016. URL: https://acsjournals.onlinelibrary.wiley.com/doi/full/10.1002/cncr.30245

45. UNESCO "Beijing Consensus on Artificial Intelligence and Education". – 2019. URL: https://unesdoc.unesco.org/ark:/48223/pf0000368303

46. UNESCO "Recommendation on the Ethics of Artificial Intelligence". – 2021. URL: https://unesdoc.unesco.org/ark:/48223/pf0000380455

47. United Nations "Transforming Our World: The 2030 Agenda for Sustainable Development". – 2015. URL: https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf

48. Urbina Fabio, Lentzos Filippa, Invernizzi C., Ekins Sean "Dual use of artificial-intelligence-powered drug discovery". // Nature Machine Intelligence. – 2022. URL: https://www.nature.com/articles/s42256-022-00465-9?fbclid=IwAR11_V1cd9SUxEvUfwrWMA7TU-croyYIY1nBDUL3KaS-8B4rG5MIqZCmjm0M&utm_medium=affiliate&utm_source=-commission_junction&utm_campaign=CONR_PF018_ECOM_GL_PHSS_ALWYS_PRO-

DUCT&utm_content=productdatafeed&utm_term=PID100032693&CJEVENT=25bc2722a-9b311ec814303660a18050f#Fig1

49. Yang Sh. "Crystal structure of the Frizzled 4 receptor in a ligand-free state". // Nature. – 2018. URL: https://www.nature.com/articles/s41586-018-0447-x

50. Yukhno A.S. "Overview of trends in the development of the metaverse market". // Vestnik of the Institute of Economics No. 6/2022. URL: https://vestnik-ieran.ru/index.php/mh-currentissue/22-stati/ekonomika-i-upravlenie/151-yukhno-a-s-obzor-tendentsij-razvitiya-rynka-metavselennoj

51. Yukhno A. "Digital Transformation: Exploring big data Governance in Public Administration". // Public Organization Review. – 2022. URL: https://link.springer.com/article/10.1007/s11115-022-00694-x

52. Yukhno A.S., Umarov H.S. "Prospects of Development of the Metaverse: Empirical Evidence". // Administrative Consulting. – 2022. URL: https://www.acjournal.ru/jour/article/view/2097

53. Zaretckii M., Bashkirova I., Osipenko S., Kostyukevich Y., Nikolaev E., Popov P. "3D chemical structures allow robust deep learning models for retention time prediction". // Digital Discovery. – 2022. URL: https://pubs.rsc.org/en/content/articlelanding/2022/dd/d2dd00021k