Artificial Intelligence at Present and Tomorrow

V Gorodetsky and R Yusupov

St Petersburg Electrotechnical University LETI, Saint Petersburg, Russia

E-mails: vladim.gorodetsky@gmail.com

Abstract. Artificial Intelligence, part of Information Technology, is currently the hottest spot in Information technology the main research subject of which is knowledge and the main research directions are methodologies, methods, and algorithms for knowledge acquisition, knowledge representations, and knowledge usage. The paper describes the core scientific content of these directions, outlines their state-of-the-arts, and analyzes briefly new trends and prospects for research and development in Artificial Intelligence for the forthcoming years.

1. Introduction

Artificial Intelligence (AI) is currently the hottest spot in Information technology (IT). Since the early 2000s, AI has become the main driving force of IT providing speed-up of scientific and technical progress in the digital world, i.e. in industry, economy, etc. AI applications usually called also intelligent products have practically unlimited application areas that range from global manufacturing and transportation systems to small-scale companies, from large-scale communication and control systems to household devices, from sophisticated systems of military destination to medicine, healthcare, sociology, and culture.

The world economy leaders like the USA, China, the European Union, Japan, etc., have claimed the national strategic programs aimed at ever-increasing developments of AI technologies during the forthcoming period until 2035. Russia has also recently adopted such Program [1].

AI science and technology have already more than 60-year history, and during this history, the periods of unreasonable expectations have been interspersed, from time to time, with periods of undeserved disappointments called AI winters. Coherently, one can observe the change of priorities in AI research, re-evaluations of its methodologies, conceptual architectures, and software implementations. However, what is the subject of AI science research, and what is its structure? What chapters constitute AI? What are their current and forthcoming priorities? These questions, along with many others, regarding AI science remain discussable so far. There is no common point of view on this matter in both communities, scientific and industrial, although it is extremely needed due to their ever-increasing interactions nowadays.

Indeed, there exists a natural gap between the AI science and AI industry. This gap is especially notable, e.g., in the Russian AI Strategy [1] that looks too selective, referring to only a very short list of technologies and scientific chapters of AI. It considers AI as a set of particular technologies and supporting software and hardware tools. However, there is no unique view on AI even inside the AI scientific community in the world, including Russia. For example, it is difficult to agree with the definition of AI itself given in the Russian AI Strategy as “a set of technologies allowing imitating
some cognitive functions of human to get the results that are, at least, comparable with the results of human’s intellectual activity” [1]. It is too poor and superficial, incomplete and too vague, emphasizing too few parts of the particular objectives of AI and missing too many of them. In some sense, a contrast but much more narrow opinion is presented in [2], defining AI as a kind of intelligent architecture and software of a computer.

In fact, it is impossible to strictly define such an intuitive concept as AI; this is the reason is why most of the existing definitions of AI are given in descriptive forms. J. McCarthy formulated one of such descriptions in [3, 4]: “Artificial intelligence is a multidisciplinary field of science and engineering whose goal is to create intelligent machines”. In this respect, we should mention that it is the multidisciplinary nature of AI that is the main source of multiple views on what AI is, and what the word intelligence within the term Artificial Intelligence means. However, this discussion is beyond the subject of this paper. The paper briefly refers to the AI multidisciplinary nature explaining a lot about the sources and causes of AI multiple interpretations (section 2). The main paper objectives are to outline the essence of the research subject of AI science and technology and to present its basic scientific directions (section 3), to outline the current structure of the researches and developments in AI and state-of-the-art from both science and technology points of views (section 4). An important objective is also to outline a high-level view of the main present AI trends and perspectives over the short and medium terms (section 5).

2. Artificial Intelligence: Multidisciplinary View

It is not simple to strictly frame any science, and this is much more difficult concerning AI, a multidisciplinary and multi-aspect science intersecting with many disciplines including those exploiting the word intelligence in very different senses. Let us recall that the term Artificial Intelligence was introduced in 1956 by the organizers of the Dartmouth summer workshop that triggered the era of AI. J. McCarthy, M. Minsky, C. Shannon, and other workshop organizers were professionals in cybernetics, and they did not directly associate AI with human intelligence. They considered AI in the context of cybernetics and as a part of cybernetics. While answering the question Isn’t AI about simulating human intelligence? J. McCarthy said: “Sometimes but not always or even usually. On the one hand, we can learn something about how to make machines solve problems by observing other people or just by observing our methods. On the other hand, most work in AI involves studying the problems the world presents to intelligence rather than studying people or animals. AI researchers are free to use methods that are not observed in people or that involve much more computing than people can do” [3]. By these words, he clearly emphasized the prevalent role of cybernetics basics of AI and divergence between AI and human intelligence.

In fact, from the very beginning, in 1956 and up to now, AI has a remarkable “cybernetic face”. Indeed, modern professionals in cybernetics refer to AI as a science having intersections with cybernetics and many of its constituents [4] (Fig. 1). Let us note that the above workshop gave birth to two directions in AI. The founder of the first one was J. McCarthy, who made an emphasis on precise AI knowledge models formalized in terms of various logics and not only logics. M. Minsky became the founder of AI systems with intensive use of data as a source of approximate knowledge. Later on, data-based knowledge discovery gave birth to data mining and machine learning (ML) and a recent notable breakthrough called Big Data.

However, one more event occurred close to the time of the Dartmouth workshop. In 1957, American psychologist Frank Rosenblatt proposed a model and an electronic device constructed on biological principles, and this model showed its ability to learn. This model is a prototype of a modern
neural network. The latter initiated the AI direction called now connectionism. It is more oriented to biology with an implicit objective to model the human brain’s decision-making processes. This model and associated research defined the second “face” of AI, which is the “biological face”. It also has a multidisciplinary nature with many intersections with the sciences of the biological cluster like neuroscience, psychology, cognitive psychology, brain informatics, etc., as it is shown in Fig. 2. Let us note that these sciences use the term intelligence in the sense of an explicit reference to human intelligence. It is these sciences intersecting with AI that gave rise to the main misunderstanding between AI scientists, i.e. between those accepting either cybernetics or biological viewpoint to the AI depending on their professional interests.

Figure 3 illustrates, figuratively, interactions among different directions, methodologies, tasks, algorithms, and technologies of AI in the context of their prevalent closeness to the sciences of the cybernetic or biology clusters. In the very right and very left parts of the figure, the sciences of purely cybernetic and biology clusters are located, correspondingly. These clusters depict the two poles of AI. Let the ellipsis located in between these poles with the vertical dotted line in its middle depict a set of various directions, methodologies, tasks, algorithms, and technologies of AI. The rule determining the place of each such element belonging to the ellipsis is as follows: the closer is an element of the ellipsis to the cybernetic pole, the more to the right is its location inside the ellipsis, and vice versa as regards the biology pole.

Several instances of such locations for various elements of AI directions, methodologies, tasks, algorithms, and technologies belonging to the ellipsis are depicted in Fig. 3. For example, the general random search is located closer to the right pole than the bio-inspired random search. Data mining methodology and algorithms are more formal than machine learning that can use some human-inspired algorithms. Expert-based knowledge acquisition procedures are located closer to the biology pole as compared to ML. In the left pole, one can see the overlap of AI components (neural network and deep
learning, e.g.) on the one hand, and neural network models proposed in neurobiology, cognitive sciences, and brain informatics, on the other one. The latter contains a notable overlap among each other and with an AI chapter called connectionism. The same situation can be observed close to the cybernetic pole where one can observe significant overlapping of computing sciences and technologies with the AI direction called computational intelligence [2]. A convincible example of such a close intersection is FPGA, which in the Gartner hype cycle as of August 2019 was indicated as an important AI technology that would achieve the productivity plateau in a couple of years since that time. Another example of achievements of computational intelligence is the AlfaGo program usually evaluated as a great success of AI.

3. Features and Structure of AI Research and Developments

From the very beginning, AI systems are referred to as knowledge-based systems, and the main efforts of AI scientists were focused on various tasks associated with knowledge. The most respected Russian scientist D. Pospelov wrote at that time that “artificial intelligence is a technology of problem-solving based on the idea of exploiting subject domain knowledge” [6]. Indeed, it is knowledge that constitutes the core of any intelligent system, its properties like soundness, completeness, consistency, validity, as well as efficiency, and performance and productivity of any intelligent application.

In contrast to other classes of IT systems, a knowledge base is a mandatory component of any AI system. The key role of a knowledge base in IT concerns many respects, and the most important one is that using human cognitive heuristics represented, e.g. as production rules, many tasks of exponential complexity became efficiently solvable with appropriate quality. Knowledge is the main subject of research and development in AI and this is the first important fact about AI. The second fact is that, from a high-level point of view, all methodologies, methods, algorithms, and technologies of AI can be divided into three groups: (i) knowledge acquiring, (ii) knowledge representation, and (iii) knowledge usage.

It is important to mention that these components of knowledge processing ordered in the indicated way constitute the high-level life cycle of any intelligent tool, system, or application (Fig. 4). The same position is expressed also in [6]: “AI is a part of informatics the essence of which is knowledge representation, inference based on knowledge, learning, expert systems, etc.”

The next section discusses the current state of the art of the aforementioned (i) – (iii) areas of AI in more detail.

4. AI State-of-the-Art

The directions of AI mentioned above cover practically all modern AI problems and tasks. Let us outline their state-of-the-arts in more detail.

4.1. Knowledge acquisition

The set of constituents of this part of AI includes available sources of knowledge, methods, and algorithms designed to extract knowledge hidden in the sources, and software tools supporting knowledge extraction. The input of acquisition technology corresponds to the knowledge source(s) whereas its output is a conceptual model of extracted knowledge of conceptual level ready for the next step of processing (Fig. 4).

There are two main classes of knowledge sources as shown in Fig. 5. The first of them is semantic sources as indicated in Fig. 5 in its lowest line. The second basic class of knowledge sources is data of various types obtained from various sources as indicated at the top of Fig. 5.
Domain experts usually create semantic knowledge sources. The term *semantic* underlines that these sources contain embedded domain semantics, either explicitly or implicitly, where the term *semantics* is understood as Natural Language (NL) semantics.

Fig. 5. Knowledge acquisition: sources, software tools and outputs

There are three types of semantic knowledge sources:

i. elicited directly from domain experts: the semantically labeled decision making (production) rules, set of concepts, attributes of concepts, relations, and dependencies given over them;

ii. elicited from textbooks, scientific texts, instructions, guidelines, etc. prepared by domain experts as NL texts, figures, etc.;

iii. Semantic Web resources containing high-level and domain RDF-ontologies, databases of RDF-triplets, lexical databases (WordNet databases for multiple natural languages, Wikipedia semantic resources, resources of Linked Data Web, etc.). At present, this type of semantic knowledge source is the richest one. It contains thousands of ontologies, databases of RDF-triplets, and other Web-based resources.

Let us note that the theory and practice of extraction and structuring of knowledge directly from experts constitute the AI section called *knowledge engineering* [7].

The second ingredient of knowledge sources is data that has recently become the first-class citizen in the AI world, especially since about 2008, which is the official birth year of the Big Data era. Offline historical data are usually stored in data warehouses representing, for instance, statistical data about weather, crop yield, population, etc. Data can come from sensors, devices, from the context where the data are born. Data can also be generated by equipment, collected through logging of operating software in real-time mode. Another important data source is computer simulation using mathematical models of objects, systems, processes, phenomena, etc., or using digital twins operating in parallel with the real physical entity in real-time. Other still unknown data sources can appear too.
Figure 6 (upper part) illustrates some frequently used types of data serving as inputs for knowledge acquisition technologies.

Knowledge acquisition from both sources, expert-based and data-based, is supported by a formidable number of methods, algorithms, and software tools developed specifically for AI during its 60-year-long history or, otherwise, borrowed mainly from mathematical statistics, discrete mathematics, mathematical logic, and optimization scopes. Scientific and practical literature on this matter is very vast; it is available in bibliographical and survey literature. Figure 6 (in its lower part) illustrates the classes of methods and algorithms most frequently used for knowledge extraction.

The outputs of these processes are particular domain- or application-oriented conceptual models of knowledge represented in some formal language that should be further transformed in machine-readable and, possibly, in machine-understandable structure ready for the next step of the knowledge life cycle.

4.2. Knowledge representation

Computers cannot directly use a conceptual model of knowledge obtained through acquisition technologies, since the computer language is a language of binary sequences. Therefore, the conceptual knowledge model has further to be transformed into a machine-readable form. Such a transformation is the essence of the process called knowledge representation (KR). The resulting knowledge structure is called KR too. In other words, KR as an area of AI comprises a set of methods, algorithms, and technologies intended to transform knowledge specified in a language at the conceptual level into a machine-readable form. The KR technology result is called the KR structure. As a rule, KR technology includes two steps: (i) specification of conceptual knowledge in terms of particular KR language usually performed by humans and (ii) automatic translation of this specification into machine-readable form. Let us note that semantic technologies of KR assume the use of ontology. In this situation, the resulting machine-readable form of knowledge will be machine-understandable too.

The history of the KR area of AI is very rich, and many models, algorithms, and technologies have been proposed. It is worth mentioning that this history began with semantic network structures for KR on the conceptual and formal levels, and translation of these formal structures into machine-readable form. Later on, many formal structures and formal representation languages were proposed and tested, and, surprisingly, recently, the semantic network has become again the most popular KR structure.
However, the novel quality of this structure is that, on top of the KR, the ontology concept structure enriched with multiple relations is used.

Descriptive logics that are specially designed decidable fragments of the first-order logic sometimes enriched with modal operators constitute the formal basis of modern KR models. These logics use unary predicates called concepts of ontology and binary predicates called roles corresponding to the relations given over the ontology concepts. These assumptions result in the decidability of the family of descriptive logics. The set of unary predicates \( C = \{ C_1, C_2, ..., C_n \} \) structured according to the set of binary predicates representing atomic roles \( R = \{ r_1, r_2, ..., r_m \} \) subject to the standard syntax of the first-order logic constitutes ontology scheme, Tbox.

Another ontology component, Abox, represents the set of instances (facts) of the atomic and, possibly, other ontology concepts. Syntax of Abox comprises the statements of two types:

i. \( a: A \) (“\( a \) is an instance of the concept \( A \)”), and

ii. \( aRb \) (“instances \( a \) and \( b \) of the concepts \( A \) and \( B \) are connected with the role \( R \)”).

In these terms, the knowledge base is a pair \( Kb=< T, A > \), where \( T \) and \( A \) are arbitrary Tbox and Abox.

At present, two main types of structures are used to represent a semantic network of the knowledge base scheme. The first one is RDF-triplet structures [8] supported by several formal specification languages, e.g. RDF, RDF-scheme, and OWL [9]. The second one is LPG structures (Labelled Property Graph) widely implemented in graph databases like Neo4j [10], Orient DB [11], etc. (about 10, in total).

Although both types of structures are designed to represent ontology-based knowledge bases with semantic networks as their formal basis, they differ in many respects. RDF model, which is the basis of Semantic Web, has many advantages, e.g., simplicity and comprehensibility of the RDF model conception, the possibility of using unique global ID, and compatibility with SQL. In addition, it is supported by practically unlimited Semantic Web resources of RDF-triple stores and hundreds of standard domain ontologies standardized by the W3C Consortium. RDF-conception has a strict underlining formal basis. However, RDF-based languages possess very low expressive capabilities to represent tightly connected data and knowledge and computationally intensive in the question-answering component.

In graph databases, the traditional operations like Create, Read, Update, and Delete (CRUD-operations) are implemented through the graph model. Graph databases use an index-free scheme of relation representation possessing high performance of the question-answering component. Additional motivation to use LPG knowledge structures is their extra high expressiveness, flexibility with regard to both the knowledge graph structure and its content, update. However, the technologies supporting knowledge graph databases are not mature so far and require rather highly-qualified personnel involved.

Let us also add that neural networks-based AI systems represent semantics free hidden knowledge in actionable mode.

### 4.3. Knowledge usage

This AI scope refers to the sets of methods, algorithms, and technologies of knowledge processing developed for the design, implementation, and deployment of specific classes of AI applications. There exist many particular technologies, for instance, multi-agent technologies, distributed and parallel technologies, self-organizing intelligent technologies, specific technologies for cyber-physical systems, digital twins, Internet of Things, cloud, and edge systems. This scope of AI includes also problem-oriented technologies, e.g., planning technologies, Semantic web and social network technologies, speech recognition, and synthesis, computer vision, and NL-processing technologies, as well as a broad class of decision-making supporting technologies, intelligent robotics, group control, behavior-based, deepfake, and other classes of technologies. The literature on all these AI-system technologies is widely presented in the literature and Internet-resources.
5. Artificial Intelligence Trends and Perspectives
The sketch of basic trends in AI during the forthcoming period for the short and medium terms perspectives given below is structured according to the same high-level structure of AI, i.e. it is outlined for knowledge acquisition, representation, and usage scopes of AI, sequentially. In this survey, the Gartner Hype Cycle for AI as of July 2020 is used [12] as an important information source. Own experience is involved too.

5.1. Knowledge acquisition
Although knowledge acquisition has been in the focus of AI research and development for several decades, emerging classes of applications are putting forward new challenges. Among them, the most challenging one is distributed intelligent cyber-physical systems (CPS) combining numerous cloud and edge AI components distributed over the network [13]. These applications emphasize the role of the privacy-preserving distributed data mining and learning intended, in particular, to design peer-to-peer coordination of the edge components’ behavior and distributed decision-making. The hot problems include the search for the knowledge needed to design explainable intelligence; mining cause-effect dependencies, development methods for acquiring knowledge from small data. The urgent problems are automatic labeling of heterogeneous big data and semi-supervised ML, knowledge graph mining, automated large-scale ontology design. Semantic computing needs significant enrichment of web-based and other semantic resources. CPS applications put forward several new knowledge acquisition tasks, e.g. digital twin-related mining and ML for biological objects, e.g. wheat, rice, etc. High dimensional time series mining remains to be a challenging task for ages and the most promised model, Gaussian Process Latent Variable is still non-mature due to computational complexity.

Another group of challenges is caused by low performance, instability, lack of robustness, and weak scalability of many existing techniques of knowledge acquiring in the context of the Big Data challenge. They are fundamental. In [14], they are united under the problem called “Robust AI”. This paper claims that the absence of robustness is an intrinsic property of the developed knowledge acquiring techniques practically preventing its use in “new high-stakes AI applications”. This notable drawback is due to several commonly used accustomed problem statements resulting, as a rule, in unstable solutions. Among such problems and potential solutions, the author of [14] indicates:

Robust optimization: the worst-case problem statement leading to the minimax optimization in which uncertainty region for each parameter should be defined; such statement of the problem incorporates robustness in the resulting solution, and it should substitute traditional optimization problem statements resulting in highly unstable solutions.

Regularization in ML: Traditional optimization of ML-techniques should be substituted by the empirical risk ML minimization that should result in a finding of a trade-off between an achievable maximum of the objective function (precision, for instance) and robustness. This is achieved through bounding a resource (“budget”) of the virtual “opponent” generating disturbance. Formally, the “budget” is bounded through optimization of the regularization parameter in empirical risk given by the objective function.

Use of risk-sensitive objectives: It is applicable if someone can organize a series of experiments with multistep decision-making (in the Markov chain model, for instance) for given probabilities of transitions and make the decision later. The experimental results obtained for the objective function in a series of experiments form a random sample whose distribution function, for the selected value of quintile value, can be interpreted as a conditional risk.

Robust inference: The way to provide robust inference is the use of a hierarchical Bayesian model representing uncertainty in terms of a joint probability distribution and treating the parameters of this joint probability distribution as hidden random variables represented as a standard probabilistic graph model [14]. This type of model can be further used as an inference model in various forms.

There are also other less notable trends in the knowledge acquisition.
5.2. Knowledge representation

Although both of the most popular knowledge representation models, RDF-based structure and LPG-based structure, possess their own important advantages and shortcomings, the current trend is in favor of LPG-based graph data and knowledge structures. This model is being rapidly developed due to acceptance of the open-source model and it is practically definite that the latter will be dominant in the nearest future. Nevertheless, extra rich semantic resources of RDF-model should not be ignored and the trend here is to integrate both and, possibly, other existing knowledge structures (production rules-based knowledge structures, for instance) according to a virtualization model like the one depicted in Fig. 7. At present, there is no work on knowledge virtualization. However, the positive practice of the data virtualization approach intended to integrate heterogeneous data from multiple sources is an example of the success story of virtualization technology [15], among the others.

Small Data, Composite AI, Generative AI, Responsible AI, Things as Customers, AI Marketplaces, Neuromorphic Hardware, Augmented Intelligence, AI Governance; note that the first five of these trends are the new entries of 2020.

However, the most notable new trend corresponds to the platform for embedded AI [13] since embedded computing has risen to the forefront in AI for the Internet of Things, CPS, and other classes of distributed applications involving edge computing, along with cloud computing applications.

As a rule, edge computing systems with embedded AI comprises numerous autonomous but simple entities and intelligence of the embedded component emerges due to intensive interactions of these entities implemented, as a rule, as software agents, since multi-agent architecture has a lot of advantages including a standard solution for the agent platform applications. Let us note that FIPA abstract architecture proposes a practically ready solution for the implementation of the digital platform and services supporting the software and communication environment of the digital platform.

An example of digital platform architecture is demonstrated in Fig. 8 [16]. In it, the services of the platform are divided into two groups: (i) system services implementing the standard functionalities of FIPA compliant agent platform, and (ii) application services. They play the roles of interfaces connecting the software and users with an ecosystem of smart services. The system services are standard services of FIPA compliant platform whereas application services should implement the commonly used functionalities required for coordination of operation of the distributed system entities. Some of such services are presented in Fig. 8 (left side) by blocks with self-explaining names.

5.3. Knowledge usage

The perspectives of AI implementation technologies and supporting software tools for novel applications are indicated in the Gartner Hype Cycle curve for AI issued in July 2020 [12]. Gartner’s list of trends retains many of its predictions given in 2019 and has five new entries. The total list of perspective technologies includes...
6. Conclusion

The paper, which presents the authors’ view on modern AI as a part of IT, demonstrates its multidisciplinary nature, defines knowledge as the basic research subject of AI, and describes the essence and contents of the main scientific directions of AI research and development. For each of the directions, which are knowledge acquisition, knowledge representation, and knowledge usage, the paper provides a sketch about its state-of-the-art and outlines the basic trends and forthcoming perspectives in the short and medium terms.

References

[1] The presidential decree of Russian Federation "On the development of artificial intelligence in Russian Federation" 2019 No. 490 http://www.kremlin.ru/acts/news/61785
[2] Kalyaev I and Zaborovsky V 2019 Control Engineering 5 78 – 83
[3] McCarthy J 2007 http://www-formal.stanford.edu/jmc/whatisai/whatisai.html
[4] Benaich N and Hogarth I 2020 State of AI Report https://docs.google.com/presentation/d/1ZUimgfXCBSLsgbad6-a-dqO7yLyzIlI1ZJbiCBUUT4/edit?usp=sharing
[5] Novikov D 2016 Control Problems 1 73–81
[6] Pospelov D 1998 History essay on informatics in Russia Eds D Pospelov and Ja Fet (Scientific Publishing Center OIGGM of the Siberian Branch of RAS) p 7
[7] Gavrilova T A Kudryavtsev D V Muromtsev D I 2016 Knowledge Engineering. Models and Methods (St. Petersburg: Publishing House Lan’) 324 p.
[8] W3C Recommendation 2014 RDF 1.1. Concepts and Abstract Syntax. https://www.w3.org/TR/rdf-concepts/
[9] W3C Recommendation 2012 OWL 2 Web Ontology Language. Document Overview (Second Edition) https://www.w3.org/TR/owl2-overview/
[10] Neo4j database https://neo4j.com/
[11] Orient DB Community Edition https://orientdb.org
[12] Gartner Hype Cycle for Artificial Intelligence https://softwarestrategiesblog.com/2020/10/25/whats-new-in-gartners-hype-cycle-for-ai-2020/
[13] Serpanos D Ferrari G Nikolakopoulos G et al. 2020 Embedded Artificial Intelligence: The ARTEMIS Vision. Computer 11 65 – 69
[14] Ditterich T G 2017 AAAI Presidential Address: Steps towards Robust Artificial Intelligence AI Magazine 3 3–24
[15] Van der Lans R F 2019 Data Virtualization: Selected Writings (Lilu Press Inc.) 342 p
[16] Gorodetsky V Larukhin V Skobelev P 2019 Pros. Int. Conf. on Intelligent Distributed Computing (St. Petersburg) Studies in Computational Intelligence vol 868 Eds Kotenko I Baldica C Desnitsky V et al (Springer Nature Switzerland AG) 35–40.