Scientometric assessment of the development of key artificial intelligence technologies

Andrey Mokhov[a] and Vladimir Tolcheev[a]

[a] *National Research University "Moscow Power Engineering Institute", Krasnokazarmennaya 17, Moscow, 11250, Russian Federation*

Abstract

We conduct a scientometric analysis of the main topics of artificial intelligence (expert systems, fuzzy logic, genetic algorithms, soft computing, machine learning, deep learning) in the period 1990-2018. Research is performed on a dataset of publications obtained from the Web of Science database. Our results indicate that during the period under review there is a steady interest of the scientific community in research in the field of neural networks, while publication activity in other areas changed significantly. We compare our conclusions with the results of other scientometric studies.

Keywords 1

Artificial intelligence, soft computing, artificial neural networks, fuzzy systems, genetic algorithms, machine learning, deep learning, expert systems, scientometric analysis.

1. Introduction

Currently the tools of artificial intelligence (AI) contain a large number of different models and methods. The increasing needs of practice and the complexity of real systems lead to the intellectualization of various spheres of public life and make AI technologies extremely popular in various scientific, technical and socio-economic fields. As a result, the last decade has seen rapid development of AI and new promising areas of research are emerging. Most of them are based on already known approaches that are adapted to modern conditions.

To our mind several paradigms coexist in the field of AI: Soft Computing (Fuzzy Systems, Artificial Neural Networks, Genetic Algorithms), Machine Learning, Neural Networks (including Deep Learning and Deep Neural Networks), Expert Systems. In this article, the term "Fuzzy Systems" is also used to refer to work in the field of Fuzzy Logic, Fuzzy Sets, Fuzzy Rules.

The significant breakthrough in accuracy has created a stable mainstream in the field of AI and attracted a large number of researchers and grant-holders to Deep Learning. However important questions remain: what happens to other (previously very popular) topics of AI, whether they remain effective? The answers to these questions are given in our article based on the scientometric analysis of publications in the Web of Science (WOS). Our research covers the period 1990-2018, because at the time of sampling (March 2020), the process of filling in the WOS database for 2019 had not yet been completed. The main research topics are: Artificial Neural Networks (NN), Fuzzy Systems (FS), Genetic Algorithms (GA), Machine Learning (ML), Deep Learning (DL).

Russian Advances in Fuzzy Systems and Soft Computing: selected contributions to the 8-th International Conference on Fuzzy Systems, Soft Computing and Intelligent Technologies (FSSCIT-2020), June 29 – July 1, 2020, Smolensk, Russia
EMAIL: asmokhov@mail.ru; Tolcheevvo@mail.ru;
ORCID: 0000-0002-1979-6411;

2. The choice of keywords and the sampling

The WOS database contains bibliographic descriptions of scientific documents, which include the following sections: title, abstract, keywords, first and last name of the author, the place of work, and the place of publication. The formed dataset has included documents with following key words in the bibliographic descriptions: Fuzzy Systems (Fuzzy Logic, Fuzzy Sets, Fuzzy Rules), Neural Network (Artificial Neural Net, NeuroNet), Genetic Algorithms, Machine Learning (Unsupervised Learning, Supervised Learning), Deep Learning (Deep Neural Network).

Our research also included Expert Systems (ES) and Soft Computing (SC). It necessary to say that term "Soft Computing", introduced by L. Zadeh in 1994, is used in publications much less often than individual technologies included in the SC. In 2018, the keyword "Soft Computing" was found only in 475 publications (the number of printed works on NN, FS and GA is shown in table 1). Low publication activity was also noted for ES (in 1990 there were 370 articles, in 2018 – 525). Thus, Soft Computing and Expert Systems correspond to significantly fewer publications (in comparison with other topics). That is way we do not use them in our analysis.

3. The choice of keywords and the sampling

The analysis of the generated dataset showed that in the period 1990-2007 the number of publications on the selected topics did not change significantly, so Table 1 shows data for a shorter period of time (2008-2018). In our opinion, the main interest for scientometric research is the analysis of trends of artificial intelligence that took place in the last decade

Table 1
Number of publications by topics of AI in the period 2008-2018

	2008	2009	2010	2011	2012
Fuzzy Systems	3546	4082	3311	3358	3654
Neural Networks	12464	13619	11457	11457	12571
Genetic Algorithms	7646	8413	7377	7750	8591
Machine Learning	3553	4263	4063	4415	4879
Deep Learning	51	39	60	92	133

Table 1
(continued)

	2013	2014	2015	2016	2017	2018
Fuzzy Systems	4165	4488	5307	5606	5666	5653
Neural Networks	13027	14509	17524	21706	28092	36486
Genetic Algorithms	8790	9644	10673	11588	11620	11575
Machine Learning	5941	7874	10684	13316	17812	24963
Deep Learning	296	706	1791	3474	7592	14088

The changes in the number of publications in the considered time period is shown in Figure 1. Blue color corresponds to Fuzzy Systems, red – Neural Networks, green – Genetic Algorithms, purple – Machine Learning, black –Deep Learning.

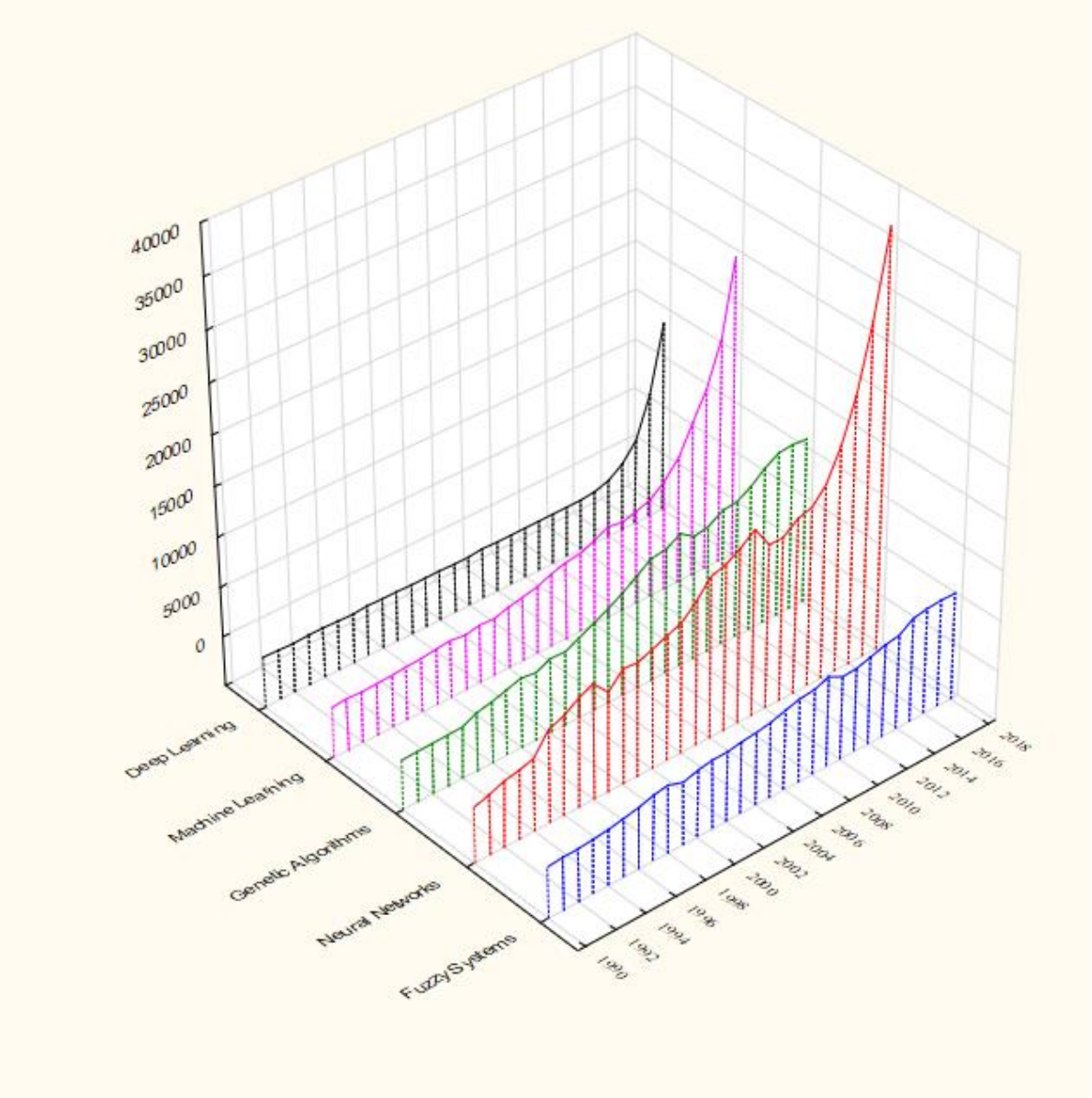

Figure 1: The number of publications on topics AI in 1990-2018 years

Analysis of the results presented in table 1 and figure 1 allows us to draw the following conclusions.

1. In the period 1990-2008, there was a steady development of all technologies (the number of publications was increased). Nevertheless the research activity on artificial neural networks in this time period significantly outstripped the work on other topics (in 2008 the number of publications on NN exceeds GA and FS in 1.6 and 3.5 times respectively, see Table 1). During this time period, there was a linear growth of scientific papers on all topics.

2. In the period 2009-2012, the number of publications on the topics slightly fluctuated. In 2012, the ratio between publications on NN with GA and FS (compared to the period 1990-2008) practically does not change. It seems that during this period there is a rethinking and reassessment of existing approaches, their adaptation to the changed computing capabilities and needs of practice.

3. The period 2013-2018 is characterized by significantly different trends. There is a breakthrough in Neural Networks, Machine Learning, and Deep Learning – an exponential increase in the number of publications. At the same time a "plateau" is formed on Fuzzy Systems and Genetic Algorithms, which indicates a steady (but not growing) number of publications. The number of works on NN in 2018 outstripped exceeds the number of articles on GA and FS by 3 and 6.5 times, respectively. The second and third places in the number of scientific articles (after NN) are Machine Learning and Deep Learning,

which are closely related to the neural network approach. This triumvirate (Neural Networks, Machine Learning, and Deep Learning) occupies the dominant positions among publication on AI.

Special attention should be paid to the interpretation of the boom in publications in the field of Deep Learning. The number of printed works on DL from 2013 to 2018 increases almost 50 times (from 296 to 14088 articles). Among the events that led to the reorientation of the scientific community to Deep Learning should be noted ImageNet, ”Google cats” and opening of free access to specialized libraries (Caffe, TensorFlow) [1,2].

In scientometrics much attention is paid to the calculation of the doubling of the volume of information, which shows over what period of time the annual volume of profile documents is doubled. This indicator is used to detect the points of growth of thsve scientific direction. Since 2013 for Deep Learning we have seen an almost twofold annual growth in the number of research papers.

Figure **2**

Figure 2: The number of publications shows a visualization of the last column of table 1, reflecting the number of publications on each of the AI topics in 2018.

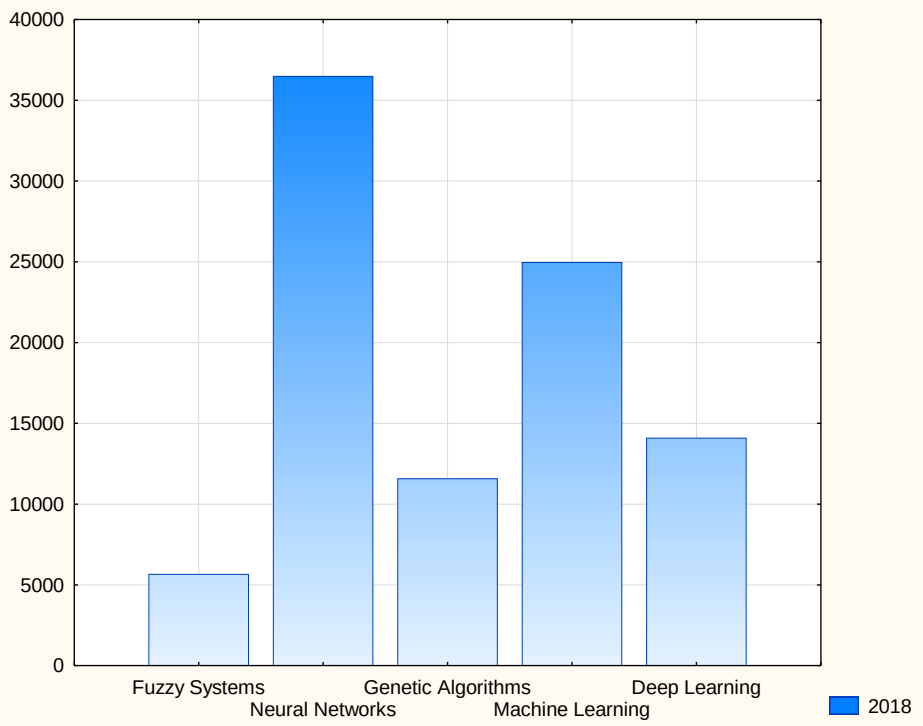

Figure 2: The number of publications of the AI topics in 2018

The obtained scientometric estimates showing the leadership of Neural Network, Machine Learning and Deep Learning. There are the expected results. However estimates for Fuzzy Systems are much less obvious. They show a significant decrease in interest in this area. At the same time in the Russian scientific community Fuzzy Systems is still one of the actively developed topics. On this issue there are a lot of dissertations, new monographs, textbooks and articles.

As an explanation of this situation, we note the following. Fuzzy Systems were originally developed to describe the inaccuracy of the real world problems by making subjective (expert) assessments and rules. Often these rules make easier to solve the problem and provide interpreted results. The success of the Fuzzy Systems significantly depends on the subject area, the “art” (qualification) of the researcher, and the reliability of the expert knowledge. FS were originally designed to control complex industrial objects in an uncertain environment. However, recently, real problems with uncertainty (when large data sets are available) have been solved using machine learning methods (especially in non-engineering fields - financial, economic and biomedical areas). In fact, the key to the success of the use of Fuzzy Systems is the availability of experts who can, based on experience and intuition, describe the object of research better than machine learning techniques. The number of such problems is limited and

consists mainly of situations for which it is necessary to get interpreted solutions and it is not allowed to use the "black box" model.

Currently thanks to a significant increase in the accumulated data and the availability of computer capacity more and more researchers are reorienting to use more universal and effective approaches based on ML. Unfortunately flow of Russian publications on FS remains consistently high and reflects a certain inertia of researchers, their inability to quickly adapt to the use of other methods. To explain this conclusion, let's quote the founder of the Fuzzy Systems L. Zadeh, who formulated the so-called “hammer principle”: “If a person holds a hammer in his hand and it is the only tool at his disposal, then he sees a nail everywhere”.

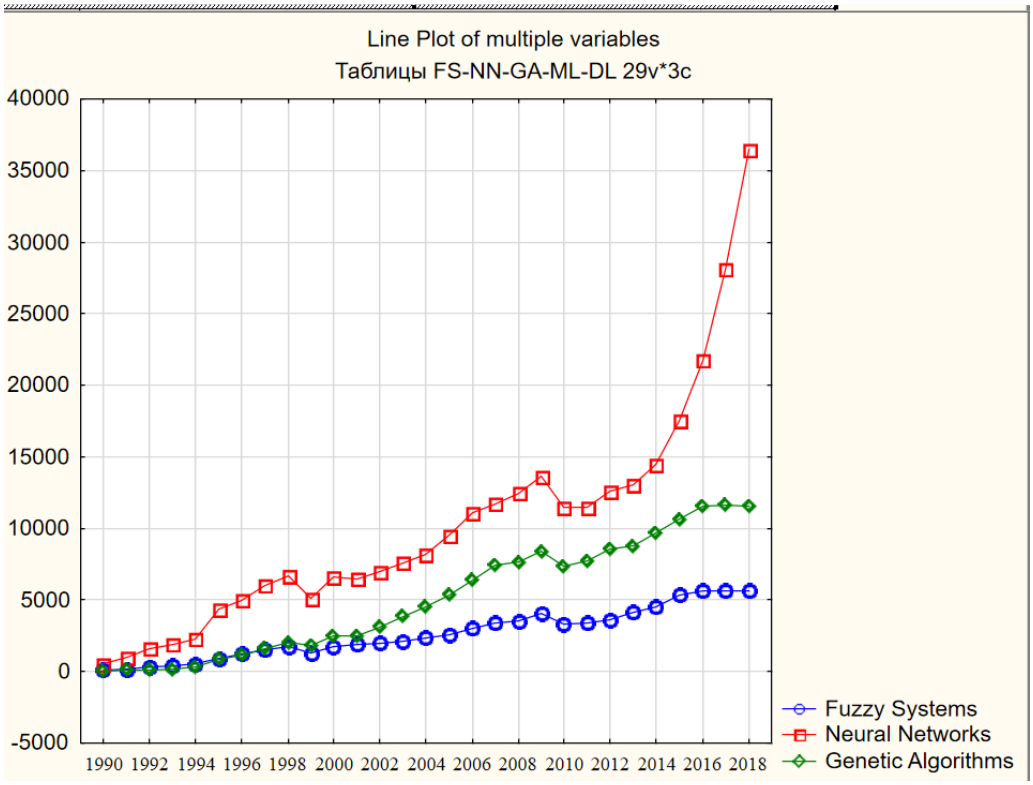

Figure 3: Number of publications on Soft Computing (Neural Networks, Fuzzy Systems, Genetic Algorithms) in 1990-2018

Let's analyze in more detail the publication activity on Soft Computing including Neural Networks, Fuzzy Systems and Genetic Algorithms. The dependencies in

Figure **3** show the development these technologies during the period under review. As can be seen from figure 3 the return of interest in AI in the 90s (after the end of the second “winter”) is primarily associated with the development of Neural Networks, while the number of publications on Fuzzy Systems and Genetic Algorithms is reached a “plateau”.

To identify the relationships between the considered topics, we will conduct an extended analysis, which also includes Expert Systems and Soft Computing. We will calculate an estimate of the occurrence of the topics in the bibliographic descriptions over the entire time period 1990-2018 (i.e., we will determine how often these terms were found together in titles, annotations, and keywords). The results are summarized in

Table 2. The analysis confirms previously made conclusions about the leading place of Neural Networks in AI publications and their regular appearance together with Machine and Deep Learning, as well as Genetic Algorithms.

The intersection of the terms FL and GA in some publications is quite high. However, their co-occurrence with DL and ML is very small and is limited with survey works. A good interpretation is given to the high frequency of co-occurrence of ES and FL, reflecting the direction of “fuzzy expert systems”, but it is difficult to explain the low co-occurrence of FL, GA in combination with (their unifying) term Soft Computing.

Table 2
The number of co-occurrence of topics in publications on AI

	Expert Systems	Fuzzy Systems	Neural Networks	Genetic Algorithms	Soft Computing	Machine Learning	Deep Learning
Expert Systems	1	2653	3242	913	125	921	55
Fuzzy Systems	2653	1	12658	6835	1689	2220	129
Neural Networks	3242	12658	1	20122	2428	29264	29296
Genetic Algorithms	913	6835	20122	1	1317	4308	308
Soft Computing	125	1689	2428	1317	1	379	30
Machine Learning	921	2220	29264	4308	379	1	10240
Deep Learning	55	129	29296	308	30	10240	1

Figure 4 shows the co-occurrence of the term "Neural network", which is most often found in AI publications, with other keywords. The pie chart shows that NN most often appear in the context of Machine (or Deep) Learning.

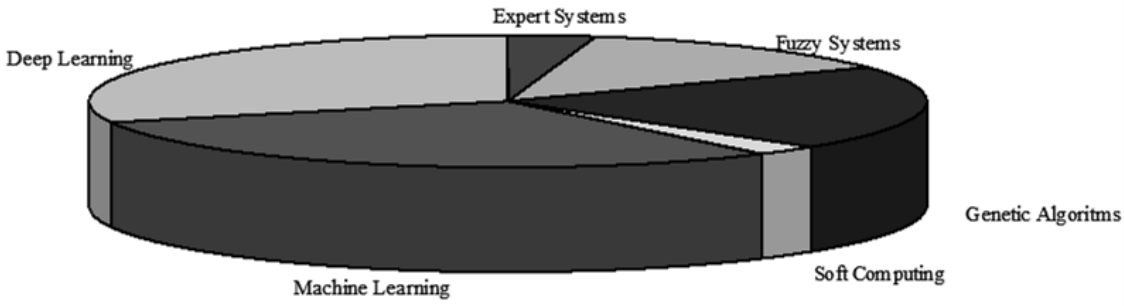

Figure 4: Co-occurrence of NN and other AI topics.

4. Comparison of the results obtained with the conclusions of other scientometric studies of the subject area "artificial intelligence"

An important condition for the correctness of the results of the scientometric method is their consistency with other conclusions made by experts or by processing documentary arrays. When preparing this article, the authors were not able to find similar studies that would study the publication activity in the above areas of AI development. At the same time, there are currently quite a lot of forecasts for the development of artificial intelligence in the medium term, with the most promising areas identified [3,4]. Notice the research conducted by MIT Technology Review, which studied trends in AI development over the past 25 years based on the analysis of 16,625 scientific papers from arXiv.org [5]. This paper highlights three key trends that developed during this period: Machine Learning (primarily learning with a teacher), Neural Networks, and new types of machine learning

(primarily learning with reinforcement). The occurrence of effective innovations in the field of machine learning is also confirmed by our other research, which conducted a scientometric analysis of the subject area "deep learning and deep neural networks". The largest number of papers indexed in the WOS database was related to the following rapidly developing learning methods [6]:

- Transfer Learning – 399 articles;
- Feature Learning – 270;
- (Deep) Reinforcement Learning – 231;
- Multi-Task Learning - 119;
- Active Learning – 79;
- Ensemble Learning – 73.

The MIT Technology Review suggests a gradual decline in the popularity of deep learning and predicts that this technology will be replaced by other technologies, many of which will be based on rethinking already known algorithms. In this case, special attention will be paid to getting more interpreted and understandable results for the user (Explainable Artificial Intelligence). Similar forecasts are also given in [7]. However, in our opinion, it is premature to draw such conclusions and in the medium term, the dominance of deep learning and deep neural networks in the field of AI will continue.

In our opinion, interesting results based on scientometric analysis of the Microsoft Academic database are presented in the publication of the Center for national technological initiative (NTI) of the Moscow Institute of physics and technology [8]. The object of the study is publication activity in the field of AI in 2019. Specialists of the NTI Center found 165.09 thousand publications in the Microsoft Academic database on AI topics. Special attention is paid in [8] to assessing the participation of Russian high-tech firms in AI work. It is noted that Russian leading companies in 2019 almost did not submit research results in this area (ABBYY - 58 articles and reports, Yandex - 20, Mail.Ru Group – 1, and Sberbank - 0). The most actively published employees were IBM (5 022 articles), Microsoft (2 658) and Google (2 466).

We obtained similar results when analyzing the subject area "deep learning and deep neural networks" using data from WOS [6]. The most frequently mentioned places of work for Russian authors were the Skolkovo Institute of Science and Technology (Skoltech), Yandex, and MIPT (Moscow Institute of Physics and Technology). V.Lempitsky (Skoltech) has the highest citation index among specialists from Russia, 12.

5. Conclusions

It is well known that the scientometric approach has a number of disadvantages and there is always a risk of losing some of the information due to incomplete queries, using only English-language documents, not taking into account synonymous terms, etc.

However, it seems that these distortions cannot significantly change the results. The formation of samples from other databases will primarily lead to changes in quantitative indicators and is unlikely to modify the identified patterns.

The conducted scientometric analysis shows that the scientific community is currently in the process of another transformation of the subject area of AI, changing of “epochs”, the period of dominance of neural networks, machine and deep learning. However, the adaptation of specialists to new approaches takes quite a long time and now there is a coexistence of several generations of AI technologies that have their own “niche” of application.

Our research did not allow us to find a significant intersection of the leading topics (NN, ML and DL) and the “classic” directions (FL, GA). These terms rarely appear together in the same publications. However, the experience of development suggests a possible return of interest in previously proposed methods, when they can be upgraded using modern advances in computing technology, accumulated data and knowledge. In this regard, it is necessary to talk about finding the new ways to jointly apply different approaches. In particular, it is promising to use FL in hybrid systems that use ML and DL to obtain solutions that are interpreted and understandable to the user [9.10].

We also note that expert forecasts of AI development pay attention to the formation of new perspective directions in the field of artificial intelligence, related both to neural network technology – neuromorphic computing, and to the "reincarnation" of the probabilistic approach – probabilistic computers, the construction of which may suggest the feasibility of using fuzzy logic [11]. In our opinion, the scientometric analysis conducted in this paper provides a fairly objective and consistent view of current priorities in the field of artificial intelligence and allows us to rank key AI technologies by the degree of intensity of scientific research.

6. References

[1] A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks, Commun. ACM, 60, 6 (2017) 84-90.

[2] Y. Bengio, P. Lamblin, D. Popovici et al., Greedy layer-wise training of deep networks, Adv. Neural Inf. Process. Syst. 2007, 153-160.

[3] M. Craglia, 100 Radical Innovation Breakthroughs for the future, 2019. URL: https://ec.europa.eu/jrc/communities/en/community/digitranscope/document/100-radical-innovation-breakthroughs-future.

[4] J. Ding, Deciphering. China's AI dream, University of Oxford ,2018. URL: https://www.fhi.ox.ac.uk/wp-content/ uploads/Deciphering_Chinas_AI-Dream.pdf.

[5] K. Hao, We analyzed 16,625 papers to figure out where AI is headed next, MIT Technology Review, 2019, URL: https://www.technologyreview.com/2019/01/25/1436/we-analyzed-16625-papers-to-figure-out-where-ai-is-headed-next.

[6] V. O. Tolcheev, Research and Analysis of the Subject Area of Deep Learning, Automatic Documentation and Mathematical Linguistics, 53, 3 (2019) 103-113.

[7] A. Averkin, Hybrid Intelligent Systems Based on Fuzzy Logic and Deep Learning. In: Osipov G., Panov A., Yakovlev K. (eds) Artificial Intelligence. Lecture Notes in Computer Science, 2019, 3-12.

[8] A. Balashova, E. Balenko, Issledovateli po iskusstvennomu intellektu iz Rossii ustupili konkurentam po chislu publikacii, RBC: Technologii i media, 2019. URL: https://www.rbc.ru/technology_and_media/ 13/09/2019/5d7a71079a79474bcb4b6144.

[9] E. Tjoa, C. Guan, A Survey on Explainable Artificial Intelligence (XAI): Towards Medical XAI. arXiv:1907.07374 (2019).

[10] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, J. Zhu, Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. CCF International Conference on Natural Language Processing and Chinese Computing NLPCC, 2019, 563-574.

[11] Intel Scales Neuromorphic Research System to 100 Million Neurons, 2020. URL: https://newsroom.intel.com/news/intel-scales-neuromorphic-research-system-100-million-neurons/