



**FORUM: ARTIFICIAL INTELLIGENCE
IN THE SOCIAL SCIENCES AND HUMANITIES**

Albert Baiburin

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS
3 Universitetskaya Emb., St Petersburg, Russia
European University at St Petersburg
6/1A Gagarinskaya Str., St Petersburg, Russia
abaiburin@gmail.com

Yuri Berezkin

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS
3 Universitetskaya Emb., St Petersburg, Russia
European University at St Petersburg
6/1A Gagarinskaya Str., St Petersburg, Russia
berezkin1@gmail.com

Olga Boitsova

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS
3 Universitetskaya Emb., St Petersburg, Russia
boitsova@gmail.com

Andrey Gromov

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS
3 Universitetskaya Emb., St Petersburg, Russia
a.v.gromov@mail.ru

Kira Kovalenko

European University at St Petersburg
6/1A Gagarinskaya Str., St Petersburg, Russia
Institute for Linguistic Studies, RAS
9 Tuchkov Lane, St Petersburg, Russia
kira.kovalenko@gmail.com

Natalia Kovalyova

University of North Texas
7400 University Hills Blvd, Dallas, USA
nvk@utexas.edu

Anna Moskvitina (Siim)

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS
3 Universitetskaya Emb., St Petersburg, Russia
anna.siim@gmail.com

Ivan Shirobokov

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS
3 Universitetskaya Emb., St Petersburg, Russia
ivansmith@bk.ru

Evgenii Sokolov

European University at St Petersburg
6/1A Gagarinskaya Str., St Petersburg, Russia
Institute for Linguistic Studies, RAS
9 Tuchkov Lane, St Petersburg, Russia
pan_liwerij@mail.ru

Nadezhda Stanulevich

Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS
3 Universitetskaya Emb., St Petersburg, Russia
nstanulevich@gmail.com

Ilya Utekhin

European University at St Petersburg
6/1A Gagarinskaya Str., St Petersburg, Russia
Indiana University Bloomington
355N Eagleson Ave., Bloomington, USA
ilia@eu.spb.ru

Abstract: AI allows work with data to extend well beyond simple keyword searches or formulaic calculations. It is applied to help specialists in the history of art and in discourse analysis as well as in linguistics and physical anthropology who analyse big corpora and construct models. Participants in the “Forum”, a written round table, are representatives of the social sciences and humanities. They describe their contact with such applications in their professional life (whether as a field of study or to practical ends), and discuss what the pluses and minuses of AI in the world of academic work and education are. They also touch upon cases where they knew, or suspected, that a student had submitted work that was actually written by a resource such as ChatGPT. Finally, ethical problems arising in connection with AI are considered.

Keywords: social sciences, humanities, artificial intelligence, AI, neural network.

To cite: ‘Forum: Artificial Intelligence in the Social Sciences and Humanities’, *Forum for Anthropology and Culture*, 2024, no. 20, pp. 11–60.

doi: 10.31250/1815-8870-2024-20-20-11-60

URL: <http://anthropologie.kunstkamera.ru/files/pdf/eng020/forum.pdf>



Participants

Albert Baiburin (Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS, St Petersburg, Russia / European University at St Petersburg, St Petersburg, Russia)

Yuri Berezkin (Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS, St Petersburg, Russia / European University at St Petersburg, St Petersburg, Russia)

Olga Boitsova (Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS, St Petersburg, Russia)

Andrey Gromov (Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS, St Petersburg, Russia)

Kira Kovalenko (European University at St Petersburg, St Petersburg, Russia / Institute for Linguistic Studies, RAS, St Petersburg, Russia)

Natalia Kovalyova (University of North Texas, Dallas, USA)

Anna Moskvitina (Siim) (Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS, St Petersburg, Russia)

Ivan Shirobokov (Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS, St Petersburg, Russia)

Evgenii Sokolov (European University at St Petersburg, St Petersburg, Russia / Institute for Linguistic Studies, RAS, St Petersburg, Russia)

Nadezhda Stanulevich (Peter the Great Museum of Anthropology and Ethnography (Kunstkamera), RAS, St Petersburg, Russia)

Ilya Utekhin (European University at St Petersburg, St Petersburg, Russia / Indiana University, Bloomington, USA)

ChatGPT-3.5

Artificial Intelligence in the Social Sciences and Humanities

AI allows work with data to extend well beyond simple keyword searches or formulaic calculations. It is applied to help specialists in the history of art and in discourse analysis as well as in linguistics and physical anthropology who analyse big corpora and construct models. Participants in the “Forum”, a written round table, are representatives of the social sciences and humanities. They describe their contact with such applications in their professional life (whether as a field of study or to practical ends), and discuss what the pluses and minuses of AI in the world of academic work and education are. They also touch upon cases where they knew, or suspected, that a student had submitted work that was actually written by a resource such as ChatGPT. Finally, ethical problems arising in connection with AI are considered.

Keywords: social sciences, humanities, artificial intelligence, AI, neural network.

QUESTIONS FROM THE EDITORIAL BOARD

AI and neural networks have become fixtures in our daily lives. We use them for purchases, business transactions, and simply for entertainment: for medical advice, when we’re speaking to call centres, using biometric authentication, or translating texts from another language.

Neural networks are everywhere — including places where their use is undesirable, or indeed unjustifiable or borderline unlawful. And if in the world of creative arts the rising presence of AI has generated rows and industrial action, in the world of social sciences and humanities, things, so far, are quieter. They certainly do have a background role, however, in digital humanities, for the analysis of big textual corpuses by specialists in literature and the history of art, or in discourse analysis. AI allows work with data to extend well beyond simple keyword searches or formulaic calculations, as discussed in a recent study enumerating, for example, qualitative analysis, historical modelling, and the identification of repeats and echoes in works of art [Gefen et al. 2021].

It is such use of neural networks in social science and humanities on which we seek to focus here. Participants of the discussion answered the questions below.

- 1 *What are the applications of neural networks in academic life that you have heard or read about? If you personally have had contact with such applications in your professional life (whether as a field of study or to practical ends), then in which ways? What would be your predictions about likely further uses in the near future?*
- 2 *How useful do you consider AI in academic research and analysis? Have you yourself employed it for, say, putting together bibliographies? What is the likely impact of resources such as ChatGPT that can produce superficially plausible texts, but work entirely on the principle of compilation, without necessary reference to facts as such?*
- 3 *What, in your view, are the pluses and minuses of AI in the world of education? Have you encountered cases where you knew, or suspected, that a student had submitted work that was actually written by a resource such as ChatGPT? In what ways is it possible to identify such work? What are the methods by which we might effectively combat this new pedagogical challenge in a general sense, not just on a case-by-case basis?*
- 4 *What are your views on the ethical implications of AI use in academic work and teaching? Does having recourse to neural networks create additional problems in terms of identifying authorship?¹ What are the implications for intellectual property rights when the products of neural networks are concerned? Should there be limits on the use of AI, and if so, of what order?*

References

Gefen A., Saint-Raymond L., Venturini T., 'AI for Digital Humanities and Computational Social Sciences', Braunschweig B., Ghallab M. (eds.), *Reflections on Artificial Intelligence for Humanity*. Cham: Springer, 2021, pp. 191–202. (Lecture Notes in Computer Science, vol. 12600). doi: 10.1007/978-3-030-69128-8_12.

¹ Anthropologists, of course, have long grappled with the issue of authorship / intellectual property rights, particularly in the case of data acquired in "the field", and especially since the entire relationship between researchers and informants began coming under scrutiny.

ALBERT BAIBURIN

My experience of using AI is very small, but I am interested not so much in AI itself, as in the situation that may be observed in connection with its “invasion”. The general picture of what is happening looks more or less like this. Like some of my colleagues, I am accustomed to regard the various systems of AI primarily as tools, the use of which helps to solve largely practical problems: collecting the material needed to identify subjects of various kinds, compiling a bibliography on one topic or another, making a preparatory analysis of a text, and so forth. Demands of this kind are constantly increasing and are becoming quite usual. Meanwhile the place of AI in our lives is becoming ever greater. It is not only becoming ubiquitous; it is bringing about substantial changes. One cannot but notice how the sphere of human communication is shifting more and more onto social networks, where new kinds of communities are arising. A significant part of the teaching process takes place on the internet or using it. Everyday life is changing, and it is already hard to imagine it without various manifestations of AI. The language of everyday communication is changing too, and its vocabulary is constantly being extended with terms from the field of AI. An anthropologist cannot help noticing these changes, but for the time being they are not always taken account of in our research. We use networks as sources for the information that we need, but we are still not much interested in their influence on the character of this information, or more broadly on cultural strategies and practices.

Albert Baiburin

Peter the Great Museum
of Anthropology and Ethnography
(Kunstkamera), RAS /
European University
at St Petersburg,
St Petersburg, Russia
abaiburin@gmail.com

At the same time, AI specialists are working tirelessly to give the artificial the character of the natural. The so-called anthropomorphising of AI is underway, and the degree of its apparent closeness to the human is growing (quite a curious process). This anthropomorphising is, of course, not only becoming apparent in robotics and other fields of AI, but also, for example, in the language we use to describe various AI systems, to which we have begun to relate almost as if they were a sort of partner or even interlocutor. As a result, it might probably be said that whereas AI was for a long time considered a technological project, now it is becoming to a certain extent a social project also. It is no accident that certain philosophers and sociologists are discussing the appearance of the so-called new or artificial sociality, which arises in social interactions with the participation (as a rule, as an intermediary) of various “representatives” of AI — chatbots, search algorithms, voice assistants, etc.¹ It is probably impossible to predict how people’s interaction with AI will end up. Alongside the obvious usefulness of new technical devices and other positive elements, one might, it seems to me, foresee substantial complications, which will affect the further changes in relationships between people and lead to breakdowns in communication (at least at the level of direct interaction).

The use of AI in education is a question. Those positive effects of the application of AI to teaching that I know of regard the use of various forms of ancillary instruments, such as preparing presentations, creating the necessary texts or images, online classes, and so forth. I cannot imagine the possibility of the teacher being completely replaced by an AI analogue, at least because the human relationships between students and teachers cannot possibly be replaced by any sort of analogues. It seems to me that restrictions on the use of AI might be introduced here, but they should not prevent its application for particular courses of study or for technical purposes.

As for the possible ethical problems connected with the authorship of a text created with the involvement of a neural network, in this situation, from my point of view, there should not be any “variants”: the author is the person who had the idea for the research, who formulated the task for the neural network and analysed the material obtained. It goes without saying that the author must indicate the neural network that he or she used to work on the text.

Student work seems to me something like an authorial text if it is a matter of the simple copying of the material obtained from a neural network (particularly if the student has not indicated its source). In such a case one can probably say that it is only the form and

¹ See, for example: [Rezaev, Tregubova 2021: 293–296].

character of the borrowing that has changed in comparison with traditional plagiarism. So far, I have not come across any such work. At the same time, one should bear in mind that in order to get a text of decent quality, the student must do serious work with the neural network. I have in mind primarily the need to think through and clearly formulate the task (possibly more than one task), since the character and completeness of the answer will depend upon this. Such a “collaboration” may be intensified when the student has not simply copied the text generated by the neural network, but worked on it to bring it up to an acceptable level (from his or her own point of view). I do not think that in such cases one can speak of borrowing. It is another matter that we must think of changes to our methods of testing knowledge and not reduce them to written texts. In any case, it would be strange to prohibit the use of AI tools in the learning process. Rather, the reverse is the better strategy, since recourse to AI allows work to be organised more rationally (with less expense of time and energy), and for the necessary results to be obtained.

References

- Rezaev A., Tregubova N., ‘Ot sotsiologii k novoy sotsial’noy analitike: krizis sotsiologii i problema iskusstvennogo intellekta’ [Sociology on the Way to New Social Analytics: The Crisis in Sociology and the Problem of Artificial Intelligence], *Russian Sociological Review*, 2021, vol. 20, no. 3, pp. 280–301. (In Russian).
doi: 10.17323/1728-192X-2021-3-280-301.

YURI BEREZKIN

2

Although I understand that I am not entirely addressing the topic of the questionnaire, I shall nevertheless express my opinion. Any programme that is created to process large amounts of data can only be welcomed. The problem lies not in its usefulness, but in how it can be applied to particular research. At the end of 2023 our folklore and mythology database consisted of 1019 rows (traditions) and 3167 columns (motifs). Anyone who has learnt how to process it, can process it by various means. I have sent our correlation table to everyone who wanted to use it, among them people from the USA, France and South Korea, as well as Russian “IT people”. I think that some of them will have used AI, because the schemes and tables that appeared as a result were striking in their

Yuri Berezkin

Peter the Great Museum
of Anthropology and Ethnography
(Kunstkamera), RAS /
European University
at St Petersburg,
St Petersburg, Russia
berezkin1@gmail.com

complexity, logic and graphic presentation. The one thing that they did not have was meaning. By “meaning” I understand new results that are both non-trivial and convincing, that can be expressed in human language and juxtaposed with data from other historical disciplines. The reason is that our colleagues, who had a good understanding of mathematics, either had no notion whatsoever of folklore and prehistory, or adhered to what is in my view an erroneous opinion regarding the similarity between the replication of genes and those analytical units that can be extracted from narratives. I shall not go into detail, but the main thing is that there is no getting away from the neo-Kantians: the humanities, even when based on big data, presuppose not only knowledge, but understanding. This is no mysticism: the initial material is too diverse and imprecise for one to translate it into numbers while completely forgetting about what lies behind it.

There is another side to the matter. Not everyone knows that in the era of the advanced internet the main obstacles to work are connected with obtaining the initial publications. A significant number of books and articles, perhaps most of them, are NOT available on the net, and sometimes not even their titles are to be found there. Will you be able to find many French articles about the ethnography of Indo-China? And if half, or at least a third, of the nineteenth-century publications on the folklore of Europe are available, this is, of course, not the case for the twentieth century. If I lived in another country and earned good money, I could obtain a great deal by buying reprints, but that is for another life. Right now, in my field, it is still a long way from neural networks. Make information available, and scholarship will evolve much more quickly.

ANDREY GROMOV

Over the last year or two the expressions “neural network” and “artificial intelligence” have firmly entered everyday speech. Obviously, specialists with an interest in the problem, and those who are curious about it, have been following the topic for a long time. There is probably a specific moment for everyone when these terms ceased to be an attribute of science fantasy. For me, this turning-point was in March 2016, when there was a match between the neural network AlphaGo and Lee Sedol, who was the world number one Go player at that time, which the neural network won convincingly. Unlike chess, in which the computer has long been stronger

Andrey Gromov

Peter the Great Museum
of Anthropology and Ethnography
(Kunstkamera), RAS,
St Petersburg, Russia
a.v.gromov@mail.ru

than the human being, it has not until recently been possible to write a programme that plays Go better than an ordinary amateur. The application of a neural network changed everything, and it became clear that a qualitative leap had taken place, of the consequences of which we are not yet fully conscious. For me, the emotional consequences of this event are even now such as to prompt me to offer my answers to the editorial board's questions.

1

In the first place it is a matter of using neural networks (NN) for working with images, their identification and classification. NN is good at identifying animals, plants and objects from photographs and three-dimensional models. When working with images NN has an advantage over a human being in that it does not have the distorted perceptions that are natural to the latter. When selecting objects a human being is inclined to choose those that are easiest to perceive: large, beautiful, contrasting. The human being is also inclined to focus on the central part of the image. NN does not have such distortions and makes the same effort with the whole block of graphic data. Overall, processing large volumes of data is the optimal use for NN.

My personal experience of dealing with neural networks is still not extensive, and it concerns only work with texts. I have used ChatGPT and Claude. The second of these has the advantage in the number of tokens that can be loaded, and also in the format of its sources (for instance, PDF), but it is less conveniently obtainable in Russia.

Briefly, in the immediate future, neural networks will be used everywhere. But one should distinguish between the use of NN for routine operations and for academic research as such. If all is clear about the first, regarding the second, things are not so simple. I have heard from colleagues who have experience of practical use of NN that if you plot a graph of the growth of the effectiveness of NN in the results of teaching, after an initial rise (of varying length) it practically inevitably levels off. Moreover, a certain degradation is even possible thereafter, if some special measures are not taken. Therefore, it is essential for users of NN, especially in the humanities, to have the opportunity of special training in it. As an example of an attempt at this, one may cite the 'Neural Networks and Artificial Intelligence for Problem Solving in the Creative Industries' training programme for creative and managerial staff in the cultural sphere, planned to be put into operation in 2024 at the St Petersburg State University of Film and Television, as part of the National Project 'Culture'.

2

At present, AI is primarily used for carrying out routine operations, of which there are plenty in any academic research. Since large language models (LLM) work with text very well when all the necessary information is given to them directly in the request, they

can be entrusted with finding and editing stylistic errors, tautologies, repetitions, and so forth. AI can also identify keywords in a text, and compile an abstract of a given length for a report or the plan of a conference paper. These tasks, simply though they are, take a person a long time to do. Here we may also include the whole range of tasks connected with preparing texts in English. In case of doubt about which of several foreign terms is the most appropriate, AI can instantly describe the nuances of each of them and identify their semantic similarities and differences.

In academic works, bibliography is sufficiently formalised to be compiled by simpler and stricter algorithms implemented in programmes such as EndNote, Reference Manager or Zotero. AI is likely to be most useful in publishing. Probably every new tool in academic work can, if one wants, be used on the principle “I put something in, I got something out, but I don’t understand how I got it”. As the academic community adapts to a new tool, the understanding of how to prepare the initial data best, how to stick to the right method, and how to interpret the result increases. I think that the possibilities of NNs such as ChatGPT will soon be mastered to the extent that the frequency of ludicrous results or scandalous stories of the writing of meaningless articles will become negligible.

3

The independent part of learning gains substantially from the integration of AI, when instead of the rigid structure of the textbook or the lecture, we use an AI that has learnt ten textbooks. Students will be able to ask AI to explain in greater detail an element that they find difficult, and, by contrast, not waste time on material that they already know. AI can also anonymously collect statistics for teachers, on which sections of a particular course cause the greatest difficulties and likely require improvement. The position of students from towns and countries with a relatively low quality of education will be particularly improved, and this may reduce inequality in the world. However, an excessive depersonalisation of the process of education would also have undesirable effects.

At present one can use a NN to determine the probability of a text’s having been generated by a NN. If a student has simply wanted to save time and energy, he or she has probably worked with the NN unskilfully too, and this makes it easier to unmask them. However, if students have used NN only for secondary technical tasks and have themselves checked the final result, they are only to be praised for their efforts to polish their work. (I should say that these are not real cases, since I don’t myself have a teaching role.)

4

I think that at present the ethics of using AI in academic work are no different from academic ethics in general: do not publish unverified data, do not present texts that are hard to read, and so forth. The present NNs are mostly regarded as tools, but with the

prospect of the appearance of Artificial General Intelligence (AGI), I think that it is already worth teaching oneself to regard AI as a co-author. Nevertheless, the legal aspects of the problems of the authorship of products generated by AI have not yet been regulated at the level of legislation. Evidently, at present the place to refer to NN is in the “Methods” section.

The question of plagiarism using NNs should be examined separately. Most of them belong to generative models (in particular, ChatGPT), and so the sources for teaching or analysis may be to a large extent reproduced in the results produced by the NN. This is, of course, not direct borrowing, but human surveillance is still, nevertheless, essential. However, it is the NN that would have more success in finding borrowings.

Despite the calls by extremely authoritative specialists to limit the use of AI, it is hardly possible to stop the process of its expansion into all spheres of life.

KIRA KOVALENKO, EVGENII SOKOLOV

We shall look at the questions proposed by the editorial board principally in the light of our work on the project ‘Plants and People in the Russian Empire in the Eighteenth Century: the Class Distribution of Knowledge and Practices’,¹ which is being conducted at the Centre for Machine Learning, Data Analysis, and Statistics (MAST) at the EUSPb. The project is directed towards the study of plants and their role in human life, using materials from manuscript and printed sources. As well as the texts, the ethnobotanical database PhytoLex, which is the researchers’ basic working tool, includes scanned eighteenth-century herbarium specimens from the RAS Botanical Institute collection. The database is on open access at the site <<https://phytolex.eusp.org/>>. In the same place one can find short notes on the plants, information on new publications on ethnobotany and videos of talks on plants (“Blog” section).

As in any research activity that is connected with the analysis of texts, a major role is played in the ‘Plants and People in the Russian Empire

Kira Kovalenko

European University
at St Petersburg /
Institute for Linguistic
Studies, RAS,
St Petersburg, Russia
kira.kovalenko@gmail.com

Evgenii Sokolov

European University
at St Petersburg /
Institute for Linguistic
Studies, RAS,
St Petersburg, Russia
pan_liwerj@mail.ru

¹ <<https://eusp.org/projects/rasteniya-i-lyudi-v-rossiyskoy-imperii-xviii-veka-soslovnaya-distribuciya-znaniy-i-praktik>>.

in the Eighteenth Century’ project by optical character recognition. In most cases we use the FineReader programme, which has an inbuilt function for character recognition in pre-revolutionary orthography. After processing, a PDF file containing a layer of text is loaded into the PhytoLex database, where a special module created at MAST allows a screenshot to be taken of the required fragment and saved in the database in image format. If the PDF file contains a textual layer, the corresponding text is also saved in the database.

In addition, for books with non-standard typefaces we have used Transkribus, a special service for training models for character recognition. Using the tools packaged with Transkribus, we have trained a special model for character recognition of eighteenth-century texts.¹ It allows text printed at the University Printing House by Vasily Okorokov² to be transcribed to a very high quality, with 99.4 % accuracy, and is also usable with the publications of other printing houses that had similar typefaces.

At present we are also testing the application of syntactic tagging of the text. Since the position in the sentence occupied by plant names (phytonyms) is usually that of semantic actants of the matrix predicate, that is, dependent on the verbal apex that indicates their properties, their connection with signs of one sort or another, their effect on living creatures, etc., syntactic tagging allows us to discover a significant volume of information about the plant. By way of experiment, we applied the Stanza natural language processing neural network model,³ extended with code to extract syntactic relations from presentations of trees.⁴

By applying this code to the text *The ash (Fraxinus excelsior) is best sown in October in light soil* [*Yasen (Fraxinus excelsior) seyetsya luchshe v Oktyabre na legkoi zemle*], from G. Grosse’s *Bergmann’s Natural Description of the Earth*, the syntactic parser of the syntagrus packet produces the following analysis:

```
{'id': 1, 'form': 'Ясень', 'lemma': 'ясень', 'head': 6, 'deprel': 'nsubj'}
```

...

```
{'id': 6, 'form': 'сеется', 'lemma': 'сеяться', 'head': 0, 'deprel': 'root'}
```

```
{'id': 7, 'form': 'лучше', 'lemma': 'хорошо', 'head': 6, 'deprel': 'advmod'}
```

```
{'id': 8, 'form': 'в', 'lemma': 'в', 'head': 9, 'deprel': 'case'}
```

¹ <<https://readcoop.eu/model/russian-print-of-the-18-c-v-okorokovs-printing-house/>>.

² A master printer and publisher active in late eighteenth-century Moscow. [Eds.]

³ <<https://stanfordnlp.github.io/stanza/index.html>>.

⁴ <https://github.com/EvgeniiGSokolov/for_phytolex/blob/main/stantaxis.py>.

```
{'id': 9, 'form': 'Октябре', 'lemma': 'Октябрь', 'head': 6, 'deprel': 'obl'}
```

```
{'id': 10, 'form': 'на', 'lemma': 'на', 'head': 12, 'deprel': 'case'}
```

```
{'id': 11, 'form': 'легкой', 'lemma': 'легкий', 'head': 12, 'deprel': 'amod'}
```

```
{'id': 12, 'form': 'земле', 'lemma': 'земля', 'head': 6, 'deprel': 'obl'}
```

Although the Universal Dependencies tagging system,¹ used in this and other similar systems, is far from perfect [Osborne, Gerdes 2019], it is nevertheless possible for the researcher to extract certain important facts. For the word *ash*, the syntactic relation *nsubj* (nominal subject) is indicated, and the apex is the verb *is ... sown*, located in the root node, which, with its non-subject dependents (the prepositional groups *in October* and *in light soil*), gives valuable information about the ash as a tree and with automatic bypassing of the syntactic graph allows this information to be extracted for uniform presentation in the form, for example, of an ontology of knowledge of plants in the period that interests us.

An important positive feature of neural network parsers is the possibility of their additional training. Data tagged by the parser in .conllu format are re-annotated manually as a training sample, the parser is retrained on this sample, and in this way adapts to the material, in our case eighteenth-century texts in Russian.

2

In the 'Plants and People in the Russian Empire in the Eighteenth Century' project extraction of data from text with the help of machine learning is currently the principal activity. At present, in order to complete the PhytoLex database, on which the project is based, the curators need to look through the entire text. Methods of machine learning, for example, models based on the principles of distributive semantics [Jurafsky, Martin 2024: 5–7], allow, by means of the representation of words in the form of multivariate vectors, the comparison of their semantic closeness using measures of their cosine similarity [Jurafsky, Martin 2024: 10–11], and the identification of clusters of the closest lexemes, which means extracting from the texts groups of phytonyms, verbal forms describing the plants' functions, the places where they grow, and the ethnic group and class of the actors. In addition, there is the possibility of extracting the most frequent collocations (and therefore established names) and operating with them as integral semantic units.

Since we are working with eighteenth-century texts in the 'Plants and People in the Russian Empire in the Eighteenth Century' project,

¹ <<https://universaldependencies.org/introduction.html>>.

it is planned in the near future to use vector models trained on material of the text library of the *Slovar russkogo yazyka XVIII veka* [Dictionary of the Russian Language of the Eighteenth Century] [Sokolov 2023: 23–24] based on the Gensim module [Řehůrek, Sojka 2010].¹ The best results are obtained from a training model with the following parameters: algorithm type cbow, vector metric 8,000 dimensions, minimum number of occurrences of a word in the sample for it to be included in the model 3 entries, window size (number of wordforms counted on either side of the word) 3, number of training sessions 20. The cbow algorithm was chosen by us relying on the data of [Kutuzov, Andreev 2015: 148], which shows that this is the most effective algorithm for vectorising Russian-language material; the other parameters were obtained experimentally and checked on the corpus [Sokolov 2024: 112–113]. Data obtained by machine methods require checking thereafter, but allow the most labour-intensive stages of the analysis of the material — the collection of lexical units and their semantic grouping — to be automated. The existing results — the names of plants in Latin (in present-day and historical classifications), Old Russian, Middle Russian and Modern Russian, and lists of ethnonyms, classes and basic functions — can afford significant assistance in training the model. Future directions of the research may include a distributive study of the dynamics of semantic change of particular phytonyms or groups of them [Del Tredici et al. 2019: 2069], or hyper-hyponymic relations in this field of lexis.

We would note, moreover, that the result of the work of the models of machine learning that we use in our research is only the processing and interpretation of the data obtained on input. We consider that the incompetent use of large language models such as GPT-3/3.5/4 in the scientific process may lead to results that will be unwelcome to the scholarly community, because here instead of the transformation and interpretation of facts, false or unverified information may be generated.

3

In our view, methods of machine learning, to the extent that they exist, cannot but be employed in education; however, we cannot help agreeing with the words of the Apostle Paul: “All things are lawful for me, but all things are not expedient.” It is essential for students to master the tools of machine learning, both in their theoretical and in their practical aspect, in order to be able to apply them in their future professional activity. But using machine learning for the systematisation, processing and interpreting of data must not be confused with the dishonest generation of work. Fortunately, at present a text generated by a large language model can still be

¹ <https://radimrehurek.com/gensim/auto_examples/index.html#documentation>.

4

quite easily identified once the marker has “got their eye in”, but it is quite probable that in the future the use of LLMs may become a real problem.

The use of neural network methods in science and education, in our opinion, is also fully legitimate so long as the neural network is used as a tool for processing and systematising data and the parameters of the learning model are indicated. Since in this case the neural network only systematises the data, and does not draw conclusions, the intellectual property rights unquestionably belong to the authors of the article, who have applied neural network mechanisms for data processing. But if the neural network is used to draw conclusions about some phenomenon, this must be explicitly indicated in the text of the research, and the question of whom the rights to the results of such activities belong to must be specially determined by the scholarly community and competent legal specialists. At present we shall refrain from any judgment on this matter.

References

- Del Tredici M., Fernández R., Boleda G., ‘Short-Term Meaning Shift: A Distributional Exploration’, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Minneapolis, MN: Association for Computational Linguistics, 2019, vol. 1, pp. 2069–2075. (Long and Short Papers).
- Jurafsky D., Martin J. H., ‘Vector Semantics and Embeddings’, Jurafsky D., Martin J. H., *Speech and Language Processing*. Draft of 2024, Feb. 3. <<https://web.stanford.edu/~jurafsky/slp3/6.pdf>>.
- Kutuzov A., Andreev I., ‘Texts In, Meaning Out: Neural Language Models in Semantics Similarity Tasks for Russian’, *Proceedings of the Dialogue 2015 Conference*, 2015, s. pag. <<https://arxiv.org/ftp/arxiv/papers/1504/1504.08183.pdf>>.
- Osborne T., Gerdes K., ‘The Status of Function Words in Dependency Grammar: A Critique of Universal Dependencies (UD)’, *Glossa: A Journal of General Linguistics*, 2019, vol. 4, no. 1, p. 17. doi: 10.5334/gjgl.537.
- Řehůrek R., Sojka P., ‘Software Framework for Topic Modelling with Large Corpora’, *Proceedings of LREC 2010 workshop New Challenges for NLP Frameworks*. Valletta: University of Malta, 2010, pp. 46–50.
- Sokolov E. G., ‘Ustanovlenie semanticheskogo skhodstva, sinonimii i antonimii sredstvami distributivnoy semantiki na materiale kollektzii russko-yazychnykh tekstov XVIII veka’ [Detecting Semantic Similarity, Synonymy and Antonymy between the Lexical Items in the 18th-Century Russian Corpus with Distributional Semantics Instruments], *Tezisy Mezhdunarodnoy nauchnoy konferentsii “XXXV chteniya pamyati Yu. S. Sorokina i L. L. Kutinoy (k 110-letiyu so dnya rozhdeniya Yu. S. Sorokina)”* [Proceedings of

the 35th International Conference in Memory of L. Kutina and Yu. Sorokin held on March 23–24, 2023 on the 110th anniversary of the Birth of Yu. Sorokin]. St Petersburg: Institute for Linguistic Studies RAS Press, 2023, pp. 22–25. (In Russian).

Sokolov E. G., 'Both Teacher and Guide, or On the Meaning of the Word *nastavnik* (Tutor) in the 18th Century', *Russian Language at School*, 2024, vol. 85, no. 1, pp. 106–116. (In Russian).
doi: 10.30515/0131-6141-2024-85-1-106-116.

NATALIA KOVALYOVA

1

My practical use of AI has been minimal, and attempted a few times only for the sole purpose of satisfying my curiosity, which mounted after an overwhelming quantity of the scholarly works I consult started referencing AI tools, almost uniformly claiming the speed of text generation as its major advantage. Since I rarely equate speed with improved quality, I posted a few queries to ChatGPT, e. g. “List major disadvantages of predictive thinking in various disciplines”, “Name major researchers working on topic X”, “Discuss non-experimental research methods”, “Create a short (75 titles) bibliography on topic Y”, “Create several sets of discussion questions for a book, modifying them for different language proficiency levels”, “Write a book-by-book summary of *War and Peace* by Leo Tolstoy”.

None of the ChatGPT responses to those queries excited me to the point of considering a regular use of AI in my own work, either as an exploratory device or as a major time- or effort-saving instrument. The current shape of the tools in questions leaves much to be desired: I found them suitable assistants for crunch times when imagination fails, or when a writer needs to get unstuck, but they appeared to me annoyingly pedantic in their grammatical correctness, in contrast to the bogus facts that they delivered.

Moreover, my brief experiments with ChatGPT, you.com, and a few others, brought to mind a series of earlier would-be game-changers: slide-rules, calculators, electronic dictionaries, spell-checkers, search engines, and a few others.

Natalia Kovalyova
University of North Texas,
Dallas, USA
nvk@utexas.edu

Some of us can easily recall the times when film projectors, television, tape-recorders, computers were pronounced to change forever the way we teach and learn. Furthermore, many of those items were initially banned from schools. Students in mathematics classes were not allowed to use calculators, and essays were expected to demonstrate the writer's mastery of grammar and spelling, unassisted by a spell-checker. These days, with penmanship being abysmal almost universally, we instruct students to run a spell-check on their work before submitting it, and maths instruction has not become obsolete because of a calculator in every pocket.

It seems to me that for AI to be broadly applied at various stages of a research process, we first need to provide feedback to the developers regarding the fundamental assumptions that guide our work and that have been grossly misunderstood by the current algorithms. The IT specialists do need our feedback, and they do listen. Some of the earlier blunders of AI have already been corrected. Hopefully, we will see more progress soon. Until then, humans will outperform the machines in the areas that require information synthesis or uncovering of hidden / implied information. Humans are superior in identifying gaps in the existing literature, generating hypotheses, contextualising findings, and offering alternative explanations to observable facts. All these activities are essential to the research process, and none has been automated, justifying our reliance on old-school methods, assisted but not replaced by various technologies.

2

4

When it comes to compiling bibliographies, I welcome any help in sieving through the bulk of tangential works, especially when I make my way into a new "territory". Even with plenty of contemporary sources available online, it requires a lot of painstaking effort to put together a decent list of works that comprehensively cover major debates in the field to date. And to complete that task, I largely rely on library electronic databases, for three specific reasons: (1) a library's databases feature publications from subscription-based sources (aka academic journals from major publishers); (2) they provide access to full texts which are typically kept behind paywalls by the said publishers, and which, therefore, cannot be accessed via commercial search engines such as Google Scholar; (3) they facilitate meticulous evaluation of the source by providing abstracts, keywords, author's information, and additional bibliographical details on all their holdings. Importantly, everything that enters library databases is vetted. Never in my thirty years of writing academic papers have I stumbled upon an entry that had wrong pages, incomplete titles, or nonexistent publishers listed. Unfortunately, all those mishaps are common in bibliographies generated by AI.

To be completely honest, I sometimes start my bibliographical journeys with a quick peek into Google Scholar and Google Books

to see what is freely available, or to select a first stepping stone of the journey. Needless to say, on the topics that I specialise in, I keep a running list of publications and add new entries from publication alerts by select journals or when I come across a relevant citation.

So, my above-mentioned adventure to generate a bibliography with ChatGPT was launched exclusively as a home-grown usability test of the tool: I wanted to see whether the machine would return a familiar list of works I know exist. Thus, I requested 75 items on a topic and pressed “Enter”. The resulting list presented more than a few surprises: publications whose titles I had never heard of came attributed to highly reputed scholars in the field, and were reportedly published in some of the major journals. Since the research area I selected for the test is relatively small (a subfield of political rhetoric, geographically and temporarily limited), all major works in it are well-known and widely-cited. Consequently, a sudden surfacing of a forgotten publication would have been a major sensation. So, for a split second, the AI-generated list made me dizzy. On it, I saw references to publications in the 1980s and 1990s not by some dignitaries, but also some of my peers who could, at best, have been in high school at that time, and could not have published anything so groundbreaking as to be included into any bibliography, let alone a list of major works. On a closer inspection, out of 75 items, only a small fraction referenced real publications (mostly books); the overwhelming majority of journal papers were fake, even though they featured real names and were supposedly housed in real journals.

My immediate reaction to these results was to use my experiment as a warning about ChatGPT imperfections and outright flaws when it comes to bibliographies. However, after I sat on it for a day or so, a more reasoned response started shaping up, namely, the need to revisit the process of academic scholarship itself — from generating a research question to publishing the results. ChatGPT’s poor performance in certain academic tasks presses forward many questions regarding those tasks, as well as the assumptions about research and academic labour in general that AI either undermines or ignores. Granted, in the knowledge-making enterprise called science, compiling good bibliographies is a very early stage: they help us to map out the terrain for exploration and to construct the foundation on which our own work is going to build. The latter is expected to advance a prior conversation on the topic, in some form. In my example, ChatGPT largely “invented” the conversation, thus undermining any effort to build on it.

On a positive note, though, blunders in AI-generated responses offer excellent teaching materials for fact-checking and copy-editing; for scrutinising the organisation and cohesion of long texts;

for recognising bias, overgeneralisation, and stereotyping; for fine-tuning register and producing audience-centered writing; for discussions of collaboration and authorship; for consideration of drafting, editing, and revising, and much more — albeit these issues sound more relevant to advanced students already familiar with the process, and seeking to improve — not replace — their work with AI. For beginners, critical thinking applied to AI-generated texts has, first of all, to undermine their faith in the power of Google and to disabuse them of the notion that “everything is online” these days.

Additionally, the current state of AI-generated texts offers a tremendous opportunity for probing serious misconceptions regarding the place of writing in research, for instance, that writing a text is a burden on academics that takes their precious time away from research proper (hence, the marketing ploy of AI as a way to lift that “burden” by automating “non-essential” elements of research); that “clean” grammar and coherent text make good academic writing and (may) deliver a strong argument (hence, the articulate machines generating perfectly grammatical verdicts of sheer nonsense that is ever harder to pinpoint because of their polished surface); that sophisticated vocabulary signals advanced training and superior thinking above any pedestrian examples (hence, a parade of abstractions unsupported by a single illustration in most ChatGPT responses).

A substantial treatment of such issues falls beyond the scope of remarks in a ‘Forum’, but the urgency of a discussion of the knowledge-making enterprise itself intensifies by the hour. A lot of concerns here highlight issues around knowledge distribution, barriers to and control of access, as well as the power structures that promote the automation of knowledge production and often benefit from it. To find a solid ground on which to advocate for or against AI in research, we need a better argument as to why academic knowledge is transmitted primarily in written form; about the role writing plays in intellectual inquiry, and about the acceptability of its technological replacement.

A relative ease with which one can learn to write a prompt for an AI tool and generate a passable response has planted significant doubt in the strict adherence to the standards of academic honesty and integrity on college campuses and beyond. College administrators promptly adjusted their policies; libraries put together guides to AI tools; teaching and learning centres organised workshops for faculty in an effort to capitalize on the advantages that AI offers and to mitigate the risks it poses. Yet the concern seems to run beyond any large-scale academic dishonesty on the part of learners tempted by a powerful tool to disregard ethical standards. It raises questions about the possibility of an alternative system of knowledge-

production. Fortunately for us, machines cannot assess their own success or failure here. And through that open door, humans should enter.

3

4

As a teacher whose courses are writing-intensive, I am not a stranger to ghost-writers, essay-mills, Google-translated passages and other assistive instruments that students approach in order to cope with their writing assignments. And while some of those coping strategies speak volumes about poor time management and unwillingness to apply oneself, they often are testimonies to real barriers such as underdeveloped skills or a lack of clarity about the usefulness of a specific assignment for building specific skills. To encourage self-study and self-discovery in those students, I heartily recommend spell-checkers or readability score calculations, writing timelines and electronic reminders, group work and peer tutoring, but I also see a positive influence of their turn to ChatGPT on my own work, especially as a challenge to my instructional decisions and my design of learning activities.

Since most of my students are expert query writers by now, I have been compelled to reconsider my assessment techniques and have changed quite a few questions on my exams. I have always used short essay questions. Now, I ask more questions that require consolidation of lecture materials, visuals, and points from class discussions. For example, my essay prompts now ask the students to demonstrate the application of at least one concept from the course materials and specify those chapters in a course textbook from which they could borrow their analytical lens. And yes, I have run all my essay prompts through ChatGPT to have a clearer idea what automated responses are likely to look like. The responses generated by ChatGPT were logical and extensive, but they were not nearly satisfactory, as they did not apply either the concepts or the terminology specific to the discipline, let alone the course. In fact, out of four prompts that asked for the application of particular ideas in the analysis, two mentioned such concepts but none were discipline-specific and none managed to read a given scenario through a required conceptual lens. However, I should give credit to ChatGPT programmers at some levels: (a) all points in the prompts were addressed; and (b) all responses were carefully crafted, including opening and closing paragraphs.

In addition to the fact that AI tools challenge us to link the learning assessment to course materials, and to use case studies, extended illustrations, or a formulation of guiding principles instead of rote memorisation of facts alone, they bring to light an important but often neglected dichotomy of Process / Product. Teachers of writing have long been advocating for a switch of a focus from a product (a text) to a process (writing) in their line of work. In the current

climate of consumerism, this dichotomy resonates particularly strongly (loudly, even) beyond writing classes as an increasing number of students arrive to school casting themselves as customers and demanding a quality product for their money. In their eyes, they come to school to get a degree / diploma / qualifications, skills and knowledge (often); friends and connections (also very often); a legitimate break from other responsibilities in life (occasionally). Many agree to complete tasks and meet various requirements in order to attain their desired end. Yet, as with writing classes, the point of completing coursework is not good grades or good texts / essays / assignments produced along the way. A true product of education is a person who has ended up more knowledgeable, skilled, mature, and educated in an overall sense, compared to the day when he or she enrolled. But unless everyone involved in the process shifts their focus onto students as true products of the enterprise, our efforts will continue to be misplaced and we will continue mistaking by-products (essays, perfect exam scores) for a validation that the system works.

Thus, AI has increased the ease with which the infrastructure behind a written text can be replaced / outsourced / automated without a significant damage to the output, at least in its conventional shape. What is more, it has challenged us collectively to rethink the process of education and the roles of and relationships among its participants. So, before we rush to protect the existing practices from the (potential) threats of AI, we will all benefit from pausing and scrutinising those very practices to establish the extent to which they actually do support our mission.

ANNA MOSKVITINA (SIIM)

1
2

If I were having recourse to a neural network, I should formulate my prompt as follows: “How would an Africanist, a specialist in Swahili language and culture, respond to the questions in the present discussion?” That is the capacity in which I shall share my ideas on the proposed topics.

Africa is often positioned as a single object to be targeted for the introduction of artificial intelligence. It is accepted that one speaks of “the prospects for the development and application of AI in Africa”, “the democratisation of AI in Africa”, “the development of technologies for smart cities in Africa”, “smart education in Africa”, “AI in modern African

Anna Moskvitina (Siim)

Peter the Great Museum
of Anthropology and Ethnography
(Kunstkamera), RAS,
St Petersburg, Russia
anna.siim@gmail.com

art”, etc. And always in such generalising discourses, on one side of the scales there is the need for a revolutionary breakthrough of neural networks and smart technologies in all spheres of life in Africa, and on the other, the dangers of a digital and neural network colonisation of Africa. These are the subjects that predominate on the agendas of the large official events that embrace all the countries of the continent. Thus, the African Business Forum (ABF) that took place in Addis Ababa in February 2024 as part of the thirty-seventh summit of the African Union, with the advertised general theme of “Boosting Africa’s Transformation through Education, Science, Technology, and Innovation”, was focused, in accordance with its title, on questions of introducing AI. At the Africa Media Festival and Editors’ Forum in Nairobi, also in February 2024, the central topic of discussion was, analogously, the prospects for applying neural networks in the mass media and the replacement of journalists by AI. Cases were discussed when neural networks had been deployed, for example, in the struggle against the dissemination of disinformation in the elections in Kenya in 2022 by means of monitoring social networks in real time for offensive content and texts in English, Swahili and Sheng (an urban argot based on English and Swahili) that were stirring up inter-ethnic antagonism and calling for an escalation of violence in inter-ethnic conflicts. (The political preferences of the Kenyan electorate are indeed defined by people’s ethnicity and regional origins, for which reason pre-election debates among the people are always liable to cause at least verbal dissension on the basis of inter-ethnic relations).¹ As far as I am aware, this experiment boiled down to the identification and removal of undesirable posts primarily on Facebook, but, as the team who developed the programme acknowledged, in fact the work of the neural network proved ineffective and unsuccessful. Nevertheless, it was useful experience in the field of applied political anthropology that took into consideration the local realities and used material in the local languages.

Anything connected with neural networks in local languages, their content and their interaction with living speakers of those languages is of enormous interest to an Africanist. In my case, the language is Swahili. This is not the only African language in which neural network technologies are being developed for communication and translation, conversion of speech into printed text and vice versa (ChatGPT also works in Hausa and Amharic). But for various reasons there are more projects and products with it than with other languages: Swahili is indeed the largest in terms of the number of speakers, is a national language of a number of East African

¹ For more detail see: <<https://news.un.org/en/story/2023/10/1140862>>, <<https://www.globalwitness.org/en/campaigns/digital-threats/hate-speech-kenyan-election/>>.

countries, and for that reason is the official language of national projects and of secondary and tertiary education, the leading African language studied as a foreign language, and, finally, is being considered as a candidate for inclusion among the official languages of the UN. The importance of the factors listed is directly proportional to the extent and volume of the introduction of neural networks into the reality of countries with a Swahili-speaking population. The actual procedures of their penetration into social reality and their reception by society are worthy of attention. A state agency of the Tanzanian Ministry of Culture, BAKITA (*Baraza la Kiswahili la Taifa*, the National Kiswahili Council), which is responsible for the standardisation of Swahili and which confirms educational projects and looks after applied neural network technologies in the field, has proposed and confirmed a single norm for the name of the actual phenomenon of artificial intelligence — *akili mnemba* (Swahili, literally ‘imitating the mind’), as against the ubiquitous and long-established in colloquial speech and the media *akili bandia* (Swahili, literally ‘false mind’), and also the other widely used synonymous concepts *akili-unde* (Swahili, literally ‘created mind’) and *akili tarakilishi* (Swahili, literally ‘computer mind’). In the opinion of BAKITA experts, the definition *akili bandia* does not correspond to the nature of the phenomenon, having negative connotations of ‘false’ or ‘fake’, as in *fedha bandia* (Swahili, ‘counterfeit money’), *cheti bandia* (Swahili, ‘false documents’), *mguu bandia* (Swahili, ‘false leg’), etc. *Akili mnemba*, by contrast, reflects the character of analogous re-creation in the image and likeness of human intellect. It should be noted that ever since it was set up in 1967, an important aspect of BAKITA’s activity has been standardisation and the purist adaptation of foreign terminology (including the lexis of computer technologies and innovations) to Swahili.

I would suggest that nowadays the existence of a neural network instantiation of a language, or its absence, could be understood as an inherent feature: if we divide languages into written and unwritten, we can equally divide them into those that are or are not adopted by neural networks. When new students of African studies, or indeed any audience, first become acquainted with Swahili, among its main characteristics and peculiarities (its phonetic and grammatical structure, the number of people who speak it, the territory over which it extends, the existence of writing and a literary tradition, its social and political status as a national language and the working language of international organisations) the appearance and existence of its neural network version is a fact that must be listed.

Swahili, which ChatGPT began to “speak” in 2022, and its rival Google Bard (Gemini) in 2023, is included among the so-called

under-resourced languages. These youngest of neural networks do indeed have few resources in comparison with the English-language and French-language networks, but they are perfectly operative. For Swahili there exist applications for text recognition, conversion of text into speech and vice versa, with the possibility of choosing the voice and sex of the speaker, and even of recognising the differences between Tanzanian and Kenyan Swahili. ChatGPT and Google Gemini (Bard) in Swahili not only allow one to conduct a correspondence or dialogue with a virtual interlocutor, but also offer a “teaching” option, which “corrects” the written or inserted text into *Kiswahili sahihi* (‘correct Swahili’), providing an interlinear translation into English or any other chosen language and providing reactions, answers, counter-questions and phrases to “maintain the conversation” in both languages. ChatGPT only allows a written format, but Google Gemini can, on request, “read out” the answers that it has generated.

3

Neural networks that “speak” the language of the culture being studied and have access to an extensive informational resource in the area of a given culture and region can, with a reasonable approach and use, to a great extent optimise the processes of education, self-education and teaching in the sphere of the humanities and African and Asian studies. Just as Swahili-speakers are at present actively trying out the neural networks that have been created for them, we and our students are studying them in season and out of season. Like them, we take a great interest in examining and correcting the linguistic and factual errors made by the network, analysing and interpreting the reasons why it goes wrong, with cross-sectional and collective surveys of this virtual subject. Some of the most curious mistakes in the responses of the Swahili-speaking interlocutor are becoming “memes”, and there is a special interest in the “hypercorrection” of the neural network, which identifies quotations from folkloric and poetic texts, works of literature and phrases in dialect as incorrect, and corrects them to Standard Swahili (*Kiswahili Sanifu*). In the case of students, spontaneous conversations with the neural network prove to be useful training for subsequent communication in the milieu of study. Because it is so new, such experience proves genuinely enthralling, and thus increases motivation towards the subject. In the regime of testing neural networks it is impossible not to catch oneself asking the sceptical question “What professions will AI replace?” Could networks become in future an alternative to teachers as conveyors of knowledge of the language and the culture that lies behind it? In other words, can AI act as a speaker of the language with the skills of a professional teacher, and as an informant, i.e. a representative of the national / ethnic culture that corresponds to that language? It must be acknowledged that the Swahili ChatGPT and Google

Gemini can serve as an extremely effective and intriguing interactive textbook for students of Swahili, literary studies and cultural studies. In contact with neural networks one may find answers to many questions at the initial and advanced levels of learning Swahili. At present, within the development of a new online course 'Basic Swahili', at St Petersburg State University, I find experiments with testing neural networks for this purpose both interesting and thought-provoking. I can say with confidence that neural networks are capable of coping successfully with most of the tasks that face the authors and compilers of a course of study. Following the parameters that we set, they easily compose and output in multiple variants verified teaching texts of a standard type, scenarios for dialogues (with two or more participants), and exercises for testing and training, and all this taking into account the East African cultural realities that we need. At the level of variation and creativity they can rival our own knowledge, fantasy and ability to think up tasks to improve linguistic memory and logic. Yes, the generated texts, scenarios and especially verse compositions (!) are rather clichéd, but from the point of view of teaching methods this is justified and inevitable in the process of teaching the rudiments of the language. The text modules that we ourselves invent (compile, adapt on the basis of recordings made in the language milieu in real time) likewise cannot and should not be exclusive, super-original and not artificial. All this gives one seriously to think. In this way, having assessed the potential of neural networks in Swahili as it deserves, I consider it to the purpose, when creating new teaching programmes, interactive courses and textbooks on Swahili language and culture to check against and consult also a Swahili neural network, and, when structuring the course that one is writing, for the avoidance of repetition to have a look not only at existing editions, but also at virtual content.

The capabilities of neural networks in composing and structuring text and in systematising material for specific purposes and tasks theoretically, of course, constitute a temptation for writers. Like the detectors that reveal plagiarism and self-plagiarism that have long been universal in the sphere of publishing, expert neural network programmes that recognise text as human created or AI generated have already been developed and are being applied (I know of such applications for working with English-language texts), and it is entirely to be expected that in the immediate future they will also be called into service when manuscripts are accepted for publication and students' work and dissertations defended. I am sure that in the case of our students — students of African and Oriental studies — situations where texts generated by neural networks are substituted for independent work are improbable, because of the small size of our community, the individual approach of our colleagues in the

department to each student, and the constant feedback that students get from their supervisors, discussants and reviewers at every stage of their work. The topics of their work are equally individual, and they choose them by themselves or with our assistance, and many of them are of the sort for which neural networks in their present state, by my evaluation, could not produce a decent result. The student's inner authorial impulse, like that of a fully-fledged researcher, cannot be unaware that presenting a text generated by a neural network as one's own is a deliberate fraud. I suspect that there may be hybrid works, i.e. partially generated, to which the student has made editorial corrections to the structure and content of an artificially produced text. I guess that in larger university communities, where the flow of students is greater, such incidents occur more often and the problems of identifying "fraudulent" works are more acute. The clues that allow one to recognise AI generated work are different from those that give away "copying and pasting from the internet". Neural network generation is a more sophisticated and more highly evolved form of copying. Large texts on a given topic composed by a neural network have a clichéd structure and on close reading one may discern "seamless writing" in them. In the final analysis, students who have gone through the experience of studying and writing have a knowledge of their work, can answer for it, and are able to find their way round the topic in quite different ways and to different degrees from ones who have delegated the creation of their work to artificial intelligence.

4

As already said in the previous points, it makes sense to have recourse to a neural network as to an assistant in teaching (including language teaching), and — with a certain caution — as to a source in academic work. With a number of caveats, the network may be treated as a speaker of the language, but not as a representative of the culture that corresponds to the language, that is, an informant, because it does not belong to a living milieu. There are many cases when an informant's ignorance or lack of an answer to one's question is substantially more valuable than a neural network's extensive and exhaustive reply to the same question. I firmly believe that, when using data generated or simply provided by a neural network, it makes sense to qualify them as such. When using a neural network as an analytical instrument, it is better to note where it takes some of your tasks upon itself, and that your hypothesis or assumption has been confirmed as a result of a series of operations performed by a neural network.

IVAN SHIROBOKOV

One may confidently assert that the capacities and limitations of artificial intelligence and its influence on culture and society are a popular topic for discussion in literature and science; however, as far as I know, it has not until recently attracted particular attention from anthropologists. There is a certain irony in the fact that as I was answering *AF*'s questions, I several times caught myself feeling that I was playing the role of a chatbot. It was not, of course, a matter of how I was thinking over my answers (neural networks do not think) or writing my text (AI algorithms are still a black box for me), but of the shortcomings of the result that I obtained at the end. The overall formulations of the questions to a certain extent dictated the format of my answers. There were clear manifestations of the limited information on the subject at my disposal, my bias in relation to it, repeating theses, and an inclination towards plausible (but probably erroneous) conclusions. I have tried to overcome the shortcomings that unite the human mind and AI at the expense of additional shortcomings. Neural networks know nothing of doubts, incoherence of judgments or uneven style. I shall hope that they will be my props in a world where human beings have lost their monopoly on the creation of meaningful text.

1 The development of technology and the appearance of methods of machine learning have allowed a noticeable expansion in the possibilities for work with large volumes of formalised data and the extraction of new, valuable information from them. Simultaneously, our capacities for “massaging data” have also doubtlessly grown: the identification of false regularities and the exaggeration of the significance of the results obtained by conscious or unconscious manipulation of the material. I know little about successes in using AI tools specifically in the humanities. A cursory glance at last years publications on Google Scholar shows that researchers most often use chatbots

2

Ivan Shirobokov

Peter the Great Museum
of Anthropology and Ethnography
(Kunstkamera), RAS,
St Petersburg, Russia
ivansmith@bk.ru

as a tool for editing and translating texts, for writing code and for generalising literary data. In reality, we have recourse to the assistance of neural networks quite often: they have simply become an unnoticed part of our life (including our academic activities) and are not perceived as a separate phenomenon. It is quite possible that with the passage of time chatbots too will become elements in our environment as inconspicuous as the image-editing functions in graphic editors, search engines and online translators are today. In this sense the appearance of ChatGPT has not changed our relationship with AI: modern neural networks continue to be applied primarily in their traditional role of personal assistants.

As a physical anthropologist, I have constantly had recourse to the assistance of neural networks over the last year and a half. For me, they are in the first place an instrument of formal analysis, the flexibility and applicability of which raises the quality of the standard programmes for the statistical processing of data. As well as using ready-made specialised programmes, I myself train models, including for the evaluation of certain characteristics of skeletons which are not immediately available when working with remains (for example, their age at death). I have some modest skill at programming in Python, but whereas last year I more often wrote my programmes myself, now I usually use the free versions of the chatbots ChatGPT and Bing for this. I formulate the task in simple text form, and the chatbot provides the code to carry it out. You quite soon begin to understand that the more clearly you are able to formulate not only the task, but also the sequence of steps needed for its resolution, the higher the quality of the final product will be. Nevertheless, I usually need at least four or five subsequent clarifying requests in order to get a programme that works correctly. The main problem is that the code proposed by the chatbot may be workable, but rely on an algorithm of actions that differ in detail from what is needed to fulfil the task that I have set because some details of the analysis were not specified (and frequently were taken for granted). Such deviations have to be specially traced, examining the meaning of the individual operations proposed by the chatbot. Despite this monitoring, in my case, the experience of interaction with ChatGPT does not help to raise the level of my skills in programming. If I get a notification of an error in the working of the code, I simply forward it to the chatbot, without thinking about the nature of the problem, and receive a corrected version of the code for further testing. If I find that the programme works, but does not correspond exactly to the task that I set it, I send off a clarifying request in Russian, and receive some new code. I do not want to become a programmer, I want to understand human remains and find new meanings in that mass of data that lies behind their study. While I try to obtain from the chatbot a programme that suits me, it in turn is “programming”

our communication, making me reformulate my requests again and again and using quite distinct stimuli (my desire to fulfil a particular task) for this. I do not want to become a programmer, but still, I am studying the language of dialogue with ChatGPT.

Sometimes I simply fail to get the desired result, and the time spent communicating with the chatbot is wasted, but in other cases, using the neural network really does allow me to save my energy. For example, sometimes I used the services of the chatbot Claude AI, which can process information from files in different formats. I sought its help to extract the information that interested me from two hundred articles about different methods of determining sex from the skull. At my request the chatbot had to analyse each article in turn and answer the same set of questions regarding the method of analysis used, the size and composition of the sample, the means of validating the results, the accuracy of the evaluations obtained, and several other things. From time to time, I verified the correctness of the neural network's answers manually — and they were always correct (which, of course, is explained not only by the quality of the neural network, but also by the technical character of the information requested). Sometimes Claude AI was unable to give an answer to one or more questions or offered imprecise formulations, and in all these cases a simple check showed that the authors of the article had simply not indicated the necessary information or been vague in their formulation of the conditions in which the research was conducted.

I also have recourse to ChatGPT when working on my own texts, mostly when a translation is needed. To a large extent, the real possibilities of using chatbots to write essays or edit the structure and sections of an article remain for me an insufficiently studied field. As with coding, the key role here is played by the formulation of requests and the volume of prompts. You can send the list of questions proposed to the participants in the Forum to a chatbot, and receive meaningful, perfectly reasonable, but rather boring answers, based on theses which in part repeat themselves. Add to the request, proposing that the chatbot should play the role of a British research anthropologist, change the style of its exposition, add a bit of irony and, without fail, mention your favourite ideas in the text, and you will get a noticeably more interesting result. One should not, however, be too demanding: in response to requests that require the mention of specific works or of the biographies of known anthropologists, ChatGPT will certainly produce a text that mixes fact with fiction. You may also be initially unsuccessful in catching the neural network out in providing false information: before acknowledging its mistake, the chatbot may try to “justify itself” and offer new absurd “facts” supposed to confirm its original assertions. I find it quite amusing that some science fiction writers of former

times discussed the question of whether cyborgs could lie, but the chatbots that really exist have been leading their interlocutors astray ever since they came into being.

So far, the most interesting use of neural networks seems to me not to obtain some factual data about a subject, but as an opponent in a discussion, examining the subject from the position of logic, or in the role of a reviewer, who can help to reveal the weak points in the prepared text of some research and sharpen up its argumentation before the text is presented to a real (human) reviewer. The artificial nature of such an interlocutor is a notable advantage. The outbursts of emotion that almost inevitably occur during a live discussion, or a discussion over the net, when the participants attack each other's ideas with passion, are practically excluded when you are dealing with critical remarks addressed to you by a neural network. Despite all the frivolous conversations about the possibility of chatbots becoming conscious, I am quite clearly aware that it is an imitation of a living interlocutor. To be useful to me, it should not try to be alive. One more area of the application of neural networks with which I have been involved, at first only for entertainment purposes, is image editing. More than once I have used the assistance of neural networks to give a photo-realistic appearance to sculptural and graphic reconstructions made on skulls by various Russian anthropologists following M. M. Gerasimov's method. For this I first uploaded the image to the Artbreeder neural network (available on the net). This is a popular tool for creating portraits and landscapes in various styles, and is not intended for work with reconstructions. Even at this stage of uploading the image to the site the facial features undergo small uncontrolled changes. In the process of editing the portrait the number of distortions increases noticeably, which may lead to substantial changes in individual peculiarities. For this reason, the image obtained from Artbreeder was then edited by me in Adobe Photoshop with the aim of making the "portrait photo" as close to the original image as possible. In the majority of cases, I created photo-realistic portraits of people of different epochs for my own pleasure, and then posted the resulting pictures on social media, but in two cases this game acquired a new status. I made one of these portraits on request of the Murmansk Regional Local Studies Museum. Based on M. M. Gerasimov's reconstruction, I tried to give a photo-realistic appearance to the sculpted image of a woman whose remains had been discovered during excavations at the Bolshoy Oleniy Ostrov [Great Deer Island] burial site. In the second case, with the permission of the author of the original reconstruction, I edited the graphic portrait of an aboriginal inhabitant of South America, later published in an article by my anthropologist colleagues. After that, I had several more orders to process reconstructions, but I always refused to take part. I also refused any

payment for my work on the portrait made for the museum in Murmansk, and likewise to be named as a co-author of the article about the remains of the South American Indian. Still, this part of my experience of interaction with neural networks has remained an amusement for me. I put more of my time into the result than of my professional experience as an anthropologist.

I shall undoubtedly continue in future to use various AI tools in my work, and press for their adoption by my friends and colleagues. At the same time, I shall not at present undertake to foretell even in broad terms how exactly access to neural networks will influence the development of different sciences. It is highly probable that with time their significance will increase everywhere. (My inner neural network is well trained in generating seemingly plausible variants of answers even under conditions of inadequate information, but this time I shall deny it the right to set them down on paper without irony.) Those authors may be right who suppose that in future chatbots and AI will be able to put forward scholarly hypotheses, develop methods, conduct experiments, analyse and interpret data and write texts for publication, while at the same time taking upon themselves also the functions of editors and reviewers of articles [Van Dis et al. 2023]. Nevertheless, I doubt that any independent AI scholars and AI journal publishers will appear. D. Friedmann has noted that “Human culture is a way in which humans idiosyncratically express themselves via works and share these with other humans. To outsource the evolution of human culture is to deprive humanity from actively shaping what it is like to be human and instead passively permeating society with a surrogate culture” [Friedmann 2024: 1]. The same is true of scholarly research. If we ever outsource scholarship to artificial intelligence, it will be a suicidal step if not for humanity, then for scholarship, since it will change its very essence as a form of human activity.

3

To evaluate the positive prospects for the use of AI in education, it is easiest to start with its negative sides. The most obvious of these is the possibility of using neural networks to write student work and claim authorship of it for oneself. Moreover, this temptation will be experienced not only by schoolchildren and students, but by teachers too. Why should I spend my time reviewing student work that has very probably been written by a neural network, when I too can entrust that activity to artificial intelligence? In this sense, the evolution of AI technologies does not create any specific new problems, but simply intensifies the problem of either side’s formal attitude towards education. Even if the current system of education does not change (but it will), as new AI tools become usual elements of the environment, ascribing to oneself results obtained using them will more and more often be understood as something natural. Our brain is lazy and energy consuming, and the availability of artificial

neural networks will allow part of the load laid on biological neural networks to be moved onto them. Even twenty years ago, when I was just beginning to work with computers (not with AI tools, but with the Microsoft Word text editor), I realised that I could transfer to it some of the functions carried out by my own brain. I need no longer try to formulate my thoughts clearly in my mind. I can begin to write a sentence without knowing how I shall end it, I need not keep track of word order, style, or syntax. I can be utterly careless in writing, because I have access to the function of unlimited editing of the text on the screen. In a certain sense, the computer has not only reduced the load on my working memory, it has also affected my very way of thinking. I have learnt to think about a text visually. Now I have access to the next step, I can transfer to AI the function of editing and completing the text. And why should I consider myself a dishonest researcher in so doing? In the end, search engines' algorithms also to a certain extent affect what works I read, and, perhaps, cite in my work, that is, they affect the content of the text. There may be many steps from here to ascribing to oneself authorship of a text entirely written by AI, but with the advent of large language models (LLMs) on which the work of chatbots is founded, there appeared the possibility of covering that distance in a single bound. And in that case one might not necessarily consider oneself a dishonest student. If the subject of a compulsory course does not interest me, but I need to prepare some work on it, then why not? Should I really consider myself a bad student for such a trivial matter, since there are courses where I do try to understand things and where I may attain genuine success? Is it not the backwardness of the educational system that moves me to find ways round it?

Can lecturers learn to recognise texts created by chatbots? In some cases yes, in others no. On the one hand, according to certain publications, even now essays, précis and sections of scholarly articles created by ChatGPT are, from the point of view of outside reviewers, no worse in quality, or even better (various parameters were assessed) than equivalent texts written by people. In many cases the readers did not succeed in separating the two groups of texts (see, for example: [Herbold et al. 2023]). On the other hand, unedited texts generated by ChatGPT can sometimes be identified using special detector programmes, or even by attentive reading. LLMs do not think: they produce texts relying on statistical associations in their training data and the prompts given to them by the user. Articles generated by a neural network may appear boring and vague, and contain almost identical formulations of the same theses, and also absurd errors and references to non-existent works. However, variants of the text that have been edited by a human being are impervious to any sort of reliable expertise. Even my own small

experiments in this field show that the free programmes intended for this purpose (such as AI Content Detector) are incapable of effectively evaluating the contribution of the LLM once insignificant corrections have been made to the text. The accuracy of the evaluations of paid-for detectors of AI content is evidently somewhat greater, but also not ideal. I think that we should not rely on the further development of the technology: neural network detectors will evolve in parallel with chatbots based on models with various architectures.

It seems very doubtful to me that it will be possible to introduce effective limitations on the use of AI in secondary and tertiary education. Consequently, changes in both the education system itself and in our approach to assessing the students' attainment are inevitable. The ability to multiply large numbers in one's head and calculate their square roots may be an admirable thing in a person, but as such it is not considered significant in teaching, no matter what the student's chosen speciality may be. The use of a calculator during exams does not, as a rule, cause any indignation among lecturers. Nor does anyone require the machine learning models used in research to be indicated as authors of the work. It is highly probable that the ability to write a grammatical, structured text on a set topic will no longer have that educational value that we have ascribed to it until recently (or rather, its status in the overall system will change: the ability to formulate one's thoughts on paper will undoubtedly be an important teaching aim in primary school).

Neural networks cannot be prohibited, but they can, and should, be made part of the system of education. It seems important to me not just to discuss the possibilities and limitations of AI with students, but also to involve chatbots in participation in discussions on particular topics. This will allow their strong and weak points to be visibly demonstrated, and teach the students to evaluate the arguments, correct the factual errors in AI-generated text, and determine the effect of the training data on the neural networks' bias. The development of critical thinking and the skills to understand and put forward arguments, the evaluation of the reliability of information, setting tasks and analysing the possible methods for fulfilling them will be prioritised teaching aims. (It seems that I have stolen that thought from some author that I have read earlier, but when playing the part of a chatbot one does not consider the ethical aspect of such a problem.) Chatbots may also become effective personal tutors, taking account of the students' individual peculiarities and level of knowledge. One can already read on the net of the experience of individuals who have used neural networks as a study tool for foreign languages and programming languages. The opportunity to obtain such experience is now available to all of us.

I think it extremely probable that, although in the near future the multiplication of chatbots will lead to a general decline in the quality of education in this country and in the world at large (or at least an increase in its dispersion), in the end its role will nevertheless prove to be beneficent. Moreover, a side effect of the multiplication of neural network technologies, mostly based on visual contact with the user, may be an increased attention towards the development of pupils' abilities at oral presentation and at leading discussions at all levels of the education system.

4

Many ethical problems concerning the use of AI in scholarship arose before neural networks appeared and are completely unconnected with them. Plagiarism and the problem of authorship are only the most visible of them. No less acute is the problem of conscious or unconscious manipulation of data, carried out by researchers in order to confirm some hypothesis. How can such manipulations be revealed when the work is read? How can scholarly research be made really transparent? How important is it, and how is it possible, to distinguish between conscious and unconscious "data massaging" from an ethical position? Finally, an important problem, but one that for some reason has not attracted much attention, is what might be called the concealment of an insufficiency of information: the advent of LLMs has shone a light on how relevant that is. Texts in which the authors make up for their lack of the necessary information for the research by using false premisses that they find plausible as if these were facts are something which many of us have encountered, including when reading anthropological and archaeological texts (and sometimes when re-reading our own). Researchers inclined to such manipulations have some similarity to "hallucinating" neural networks. Perhaps attention paid to the topic of the consequences of the use of AI in science will lead to a new spurt of interest on the part of researchers in the problem of the generation of plausible nonsense by natural intelligence, including those fields of knowledge which are not over-indulged with such attention at the moment?

The fulcrum for the discussion of the topic of the ethical aspects of applying AI tools in scholarship might be the two simple ethical principles formulated last year by the editors of *Nature* and Springer Nature publishers. First, an AI tool may not be listed as an author of a research publication. Authorship implies acceptance of responsibility for the work, and neural networks by definition cannot bear such responsibility. Second, researchers must indicate the facts of their use of LLMs in their publications, for example, in the 'Methods' or 'Acknowledgments' sections [Tools 2023].

The importance of the second principle is obvious in the context of the discussion of the authorship of articles, even for utilitarian purposes. According to some (admittedly dubious) data, the majority

of articles published in the humanities do not have a single citation, and half of them have probably never been read by anybody outside the “author — reviewer — editor” triangle. Now the authors of publications that came out before 2022 and have not received any response from their colleagues can console themselves with the fact that their work has not been in vain, their articles are being “read” by neural networks or will be in the near future. LLMs need texts written by people, because using content created by AI for training leads to irreparable defects in the resulting models through the elimination of the improbable data present in the original data provided by a human being. This is true of any form of information, textual, visual, or musical [Shumailov et al. 2023]. In other words, the greater the quantity of AI-generated data affecting an AI training model, the worse it works, producing more errors in the content it generates. Consequently, material created by human beings in the period before neural networks will be of particular value for the development of AI. And consequently, it will not be enough simply to require authors to indicate their use of LLMs in preparing their publications. Paradoxically, such a requirement might be effective only if a neutral or positive attitude towards the very fact of the use of LLMs in scholarship were to be formed within the scholarly community and beyond.

The probability of introducing limitations on the use of AI in scholarship that would be in the least effective seems unrealistic to me, but attempts to introduce such limitations are inevitable. We can expect the most lamentable consequences if agencies of the state take upon themselves to control the development and application of AI. Any state is interested in expanding the possibilities of applying AI to controlling all aspects of its citizens’ life and “effective government”, and also in limiting private research in the AI sphere in order to preserve “citizens’ human rights, freedoms and security”. The difference between a world in which the life of society is entirely subordinated to an artificial super-intellect, and one in which the development of artificial intelligence is controlled by state and international bureaucracy is only that we still do not know whether the arrival of such an omnipotent AI is possible in principle, but the existence of power-hungry bureaucrats is not open to doubt.

References

- Friedmann D., ‘Copyright as Affirmative Action for Human Authors until the Singularity’, *GRUR International*, 2024, vol. 73, no. 1, pp. 1–2. doi: 10.1093/grurint/ikad133.
- Herbold S., Hautli-Janisz A., Heuer U., Kikteva Z., Trautsch A., ‘A Large-Scale Comparison of Human-Written versus ChatGPT-Generated Essays’, *Scientific Reports*, 2023, vol. 13, article 18617. doi: 10.1038/s41598-023-45644-9.

- Shumailov I., Shumaylov Z., Zhao Y., Gal Y., Nicolas Papernot N., Anderson R., 'The Curse of Recursion: Training on Generated Data Makes Models Forget', *ArXiv*, 2023. <<https://arxiv.org/abs/2305.17493v2>>. doi: 10.48550/arXiv.2305.17493.
- 'Tools Such as ChatGPT Threaten Transparent Science; Here Are Our Ground Rules for Their Use', *Nature*, 2023, vol. 613, pp. 612. <<https://www.nature.com/articles/d41586-023-00191-1>>. doi: 10.1038/d41586-023-00191-1.
- van Dis E. A. M., Bollen J., Zuidema W., van Rooij R., Bockting C. L., 'ChatGPT: Five Priorities for Research', *Nature*, 2023, vol. 614, pp. 224–226. doi: 10.1038/d41586-023-00288-7.

NADEZHDA STANULEVICH

1

As a historian of photography, I am interested in neural networks that work with images, that is, that solve problems relating to the classification of illustrations, the detection of objects, the segmentation and encoding of images, searching for key points, the generation of textual descriptions and the creation of new images.

The conditions for preserving historical photographs in institutions and private collections are determined not only by principles common to all photographic materials, but also by the special conditions for individual techniques of production. Thus, the identification of photographic techniques is one of the basic tasks of museum work. My speciality is the identification of the means by which photographs were printed, and in 2022 I had occasion to collaborate on this question with specialists who work with neural networks, during a seminar on the computer analysis of historical photographs based at the Art and Artificial Intelligence Laboratory at the School of Arts and Cultural Heritage of the European University at St Petersburg. I took part in formulating the museums' requests for the use of neural networks, and was consulted by participants on the methodology of attribution of photographic techniques of the nineteenth and twentieth centuries. During the seminar, the tasks for machine learning in searching for images in the open collections of museums and the attribution of photographs and fragments thereof were made more specific.

Nadezhda Stanulevich

Peter the Great Museum
of Anthropology and Ethnography
(Kunstkamera), RAS,
St Petersburg, Russia
nstanulevich@gmail.com

There are many publications on the internet, particularly on social networks, which do not have specific comments about the unique possibilities of AI, such as the recognition of techniques and manners of painting and makers' signatures on works of art, and the discovery of the dates of such signatures. It must be noted, though, that this sort of work is possible when there has been a previous technical reference examination and when high-resolution images have been uploaded to an open database. In this case it is not a matter of the actual attribution of painting techniques by the machine, but of the comparison of data. What happens is that the machine compares the data set being investigated with a marked-up big data set. For example, it compares the layers of paint on paintings from the collections of a museum that has provided such data for study by the neural network, or loaded high-resolution images to open access, with unattributed works from another collection. The results of the machine interpretation directly depend on the completeness of the big data set according to the parameters of the comparison experiment [Lashmanov, Rogozhnikov 2023]. The diversity of the input data gives the neural networks the possibility of finding regularities more easily and outputting an exact result. However, the final decision on the results of the experiment is taken in all cases by the human researcher.

The tasks in work on historical photographs are newer, in comparison with other works of art, and it may be that this is connected with the comparatively short era of photography — 185 years by 2024. As for AI in work with historic photographs and images, there are, in my opinion, certain limitations for its use in the immediate future. These comprise limitations in the methodology of the attribution of photographic techniques and in the parameters of the images used for training neural networks. Methods of determining the types of photographic negatives, prints and slides are based on the determination of the materials, supports and layers that contain the image [19th Century Materials and Technologies 2017]. Further filters in the specialist's work are constituted by the colour and tone of the materials, various physical characteristics of the surface of the objects investigated, and the structure of the layers [A Methodology 2017]. In this case, it is a matter of the researcher's immediate interaction with the photographic materials. An experienced specialist can carry out most of the operations with a high-resolution digital image, but part of the research requires a large number of specialised copies when AI is used: different ways of lighting the surface of the object, particular types of macro-copy, the choice of separate specific areas of the photographic material to be analysed. Despite the abilities of neural networks in detecting the colour of works of art [Lashmanov, Rogozhnikov 2023], the basic question of the application of AI to attribution is the accessibility of such images for training the neural

network, and also the existence of such copies in the possession of the institution or collector that is making the comparison. Thus, the State Catalogue of the Russian Federation Museum Fund, which at the time of writing contains 6,400,878 exhibits in the category 'Photographs and Negatives',¹ nevertheless does not allow the use of its pictures for training AI. This limitation is connected with the poor resolution of the images represented, and also with the fact that museums have uploaded only one or two types of image — the front and back of the material.

In my opinion, the most probable application of neural networks in scholarship in the broad sense in the near future will be their use in working with texts.

2 Machine training of neural networks is able to carry out tasks relating to finding similar elements, predicting the results of an experiment, selecting the best solution according to set parameters and reproducing training algorithms. At the same time, neural networks are incapable of creating anything new (as the word is generally understood), nor can they go beyond the tasks set for them [Lashmanov 2023].

The use of AI (taking into account its limitations, and with the requisite training of the neural networks) may be useful for my research. At present, I use the simplest versions of AI for search operations on images in the Google Lens and Yandex 'Search by Image' systems.

I do not use special programmes for compiling bibliographies when writing scholarly texts.

Given the lack of discussions on the application of neural networks, in my opinion, there may be problems with their correct introduction into the educational and scholarly milieu. So long as it is understood that the use of AI is just a tool that can speed up some processes, conduct optimisation and effect more quantitative research in the humanities [Lashmanov, Rogozhnikov 2023], one may not only dispel the fears surrounding the application of machine learning, but also raise the quality of academic work.

3 Among the advantages of using artificial intelligence in education may be the possibilities for optimising the processes of finding information and the generation of grammatical constructions when there are difficulties in formulating them. It may be a disadvantage when machine-generated text is thoughtlessly used as if it were one's own scholarly work, when the academic text is the most widespread means of disseminating the results of research.

¹ State catalogue of the Russian Federation <<https://goskatalog.ru/portal/#/collections?typologyId=17>>.

In my teaching practice (which is not extensive) as an Associate Professor in the core department¹ of the Peter the Great Museum of Anthropology and Ethnography (the *Kunstkamera*) at the Higher School of Economics in St Petersburg, I undertook project work with students working for their bachelor's and master's degrees in the academic year 2021/22. The level of their knowledge and their skills at working with the scholarly literature were extremely high, and there was no reason to suppose that they would submit machine-generated work. This could also be seen from the mistakes and misprints that occurred in the texts, and from the dynamics of their work during the academic year.

In my opinion, the authorial text has particularly individual signs. To achieve a full imitation of the text of a specific individual, the neural network would have to be trained on texts by that author, or else the researcher would have to insert their own personal corrections into the text. At the current stage of neural networks' evolution, there occur illogical periphrases and incorrect citations that might become the hallmarks of generated text. Besides, a knowledge of the body of academic work and methodological tendencies on the part of the professor can filter the students' work. A thoughtful, critical reading, most likely, will be able to identify the inconsistencies in a file that has not been corrected by a human being. It is important to remember, and to explain to the students, that it is the context of the research and the method of the work that determine the scholarly value of the final text.

4

In the process of teaching, it is essential to explain that AI carries out low-level tasks in searching, combining texts and rephrasing them, but the essential meaning of academic texts consists in conveying the results of academic work personally by a researcher or group of researchers. It may be that this thesis is hard to explain at the undergraduate level, but it seems possible when it comes to work for which the student and the lecturer are responsible to each other.

In my opinion, the authorship of a work carried out with the help of a neural network belongs in the first instance to the researcher who set the machine a particular task. Nevertheless, it would, in my opinion, be honest, and appropriate from an ethical point of view, to indicate the neural network with the help of which the results were obtained. One may take as an example the practice of commercial photography, which credits among the photographers the retouchers who have cleaned up the image to match the concept

¹ Core department [*bazovaya kafedra*] is a term used in Russian universities for a department in a university set up in partnership with an external institution, whether commercial or non-commercial. [Eds.]

of the journal, brand, or whatever. The experience of outlining the AI used in the work in the section on methodology might be revealing.

The limitations of neural networks, listed above, regarding the kind of task they can fulfil, in my opinion, restrict their functions anyway. As to the generation of texts, limitations are meaningless, since in 2024 we do not limit, for example, the use of the basic functions of Microsoft Word for correcting the spelling and grammar of texts.

References

- ‘19th Century Materials and Technologies’, Webinars. Image Permanence Institute, 2017, Sept. 13 <https://s3.cad.rit.edu/ipi-assets/webinars/webinar_identification_19th_century.pdf>.
- ‘A Methodology for Process Identification. Part 1’, *Webinars. Image Permanence Institute*, 2017, Dec. 13. <https://s3.cad.rit.edu/ipiassets/webinars/webinar_identification_methods_p1.pdf>.
- Lashmanov O., ‘Zachem obuchat mashiny’ [Why Educate Machines], *Blog Olega Lashmanova — Sistemnyy blok*, 2023, Oct. 18. <http://sysblok.ru/blog/blog_olashmanov/zachem-obuchat-mashiny/>. (In Russian).
- Lashmanov O., Rogozhnikov V., ‘Chto kompyuter ponimaet v restavratsii’ [What the Computer Understands about Restoration], *Proverka svyazey — Arzamas.academy*, 2023, Oct. 4. <<https://arzamas.academy/podcasts/330/5>>. (In Russian).

ILYA UTEKHIN

1

My personal experience is mostly connected with a project which did not at first have so much a scientific as an applied scientific character. (However, since both the methods of analysis used as applied to the flow of news, and their results, require interpretation, an academic publication will soon appear.) In 2018–2021 I was leading the creation of a news aggregator, which, like any aggregator, collected news from many sources and from this material formed a picture of the news of the day in different variants. This could be seen in the form of a bot on Telegram. One button on the bot produced what we called ‘Clickbait News’.¹ So as to generate the bot’s output when the ‘Trash’ button was pressed, we used two sources. First, if you search the news using Yandex for a four-

Ilya Utekhin

European University
at St Petersburg,
St Petersburg, Russia /
Indiana University,
Bloomington, IN, USA
ilia@eu.spb.ru

¹ In the original, *prikolnye* (i.e. provocative or sensational content). [Eds.]

letter word beginning with *zh*,¹ and also certain other words (*turned out, stuck, jaw...*) a lot of “clickbait” comes out. Second, since our system collected texts from the most diverse news sources, including entertaining, not serious ones, and others from countries bordering Russia where there is a Russian-language press, and the four-letter word is not forbidden by the censor, there was a fair amount of this sort of material in our own database as well. Each news item was given a “trash index”, actually, how high it rated as “clickbait”, which was determined by the system, which in a certain sense possessed artificial intelligence. This was by no means the same as today’s large language models, but a simpler thing (in our case the texts were translated into vectors and their vector similarity was found; this procedure is widely used in information systems working with natural language, but these are not always neural networks). Having viewed with my own eyes many headlines from the unserious press, I selected several hundred examples of “top clickbait”, like ‘Omsk Man Digging Girlfriend’s Allotment Found Her Ex’s Skull’. For some reason, people from Omsk often feature in absurd or horrific news items. These “top clickbait” headlines served as the training set, on the basis of which the values of the index were assigned. Here are a few examples, with the index given in brackets:

Trump’s Former Attorney Denies Rumours about Documents Eaten by President (0.0001)

Nude Volochkova² Delighted after Drawing Bra on Self (0.0016)

Guinness Record-Holder Tries to Stop Sewage Discharge into Miass River (0.0107)

Raving Botkin Hospital Patient Throws Chairs at Doctors, Then Falls out of Window (0.5717)

Pyt-Yakh Woman’s Naked Protest at Pension Reform (0.9619)

Report of Nizhniy Novgorod Shop Assistant Holing Condoms Was Fake News (0.9643)

You may disagree with the non-human evaluation of the degree of “clickbait value” that the system gives, but it is evident from these examples that a value greater than zero for this index has a certain meaning. Incidentally, the poet Lev Rubinstein, who, independently from us, discovered news headlines as a source of conceptual poetry, has from time to time published on social media similar selections of headlines, evidently chosen manually from the results of a keyword search.

¹ The Russian-speaking reader will immediately recognise the word for ‘arse’. [Transl.]

² The ballerina Anastasia Volochkova, the subject of numerous scandals in the late 2010s. [Eds.]

If we are to speak of the analysis of the news flow, the comparison of the agendas of different news outlets, the discovery of the degree and form of bias in the information supplied (and that, in fact, is the object of my research project), then here the application of artificial intelligence to work with texts does indeed allow us to proceed to another level. For example, the task of isolating in a text all the collocations that contain an evaluation or a relation, which is not so easy for a human being, can easily be done even by the free ChatGPT-3.5. And if your project has the money to pay for the processing of a large body of texts, AI allows you to obtain results that will provide data demonstrating what a person can in some cases understand intuitively, and in others is simply incapable of seeing, because these patterns are visible only in large data sets.

When a corpus of text data has been collected (say, publications in social media by representatives of the group or community being studied) that interest the anthropological researcher, AI can help in the work. Just as Adobe Creative Suite, a set of programmes for working with images and video, is now being provided with plug-ins with artificial intelligence, it is quite possible that in future new AI functions will be added to programmes for processing qualitative data such as Atlas.ti, QDA Miner and NVivo. However dystopian the idea of an interview annotated by artificial intelligence might look, AI assistance in working with large corpora of material, and not only textual material, must be seen as an immediate prospect. It would be useful for the social scientist of the future to understand what is “under the bonnet” of such tools, and how their algorithms are constructed.

But never mind annotation and analysis — since large language models can be used to create advanced dialogue systems, it is easy to imagine such a tireless artificial interviewer. That it is not a human being deprives the field of its most important aspect, the establishment of human relationships, but in some circumstances that might even be an advantage.

3

In 2023 it so happened that I twice lectured on ethnographic cinema to undergraduate students, of whom only a few were specialising in cultural anthropology and by no means all of them were sufficiently motivated to make an effort and work — many of them only needed a pass mark. Not much reading was required for the course, but it was necessary to watch the films, and, since they were long films, it was not possible to watch them in class. Over that time I encountered dozens of cases where answers to assessments generated by AI were submitted. I went over some of these texts in class, explaining which features gave me the basis for asserting with complete confidence that this text had not been written by a human being. There are such

features. They are particularly noticeable if a lecturer has themselves tried in advance to “feed” their assessments to an AI service.

The assessments in response to which I received this sort of essay were formulated, for example, like this: “In Robert Gardner’s film *Dead Birds* there is no division into chapters, but if you needed to divide it into sections (no more than eight), which parts would you select? Provide headings for them”, or “Indicate the ethical problems connected with the work of anthropologists among the Yanomami tribe discussed in the film *Secrets of the Tribe*”.

It is obvious that to answer questions of this sort one must be acquainted with the content of the film. As a rule, ChatGPT has no access to a film if there is no Wikipedia article about it, describing its content, naturally, in text form. ChatGPT actually declares this in its answer, both at the beginning and the end: “I have not seen the film, but this is what I can put together in answer to the question.” It can put it together extremely smoothly and, in a certain sense, informatively: after all, publications on the Yanomami and the anthropologists in popular editions, and also discussions about ethical aspects of fieldwork, were part of the training corpus of this language model. So, a negligent student cuts off the warnings given by the service, and, as if butter wouldn’t melt in his mouth, submits the middle of the generated text. The model fantasises: it composes a synopsis of a non-existent film, supposing what might go into a typical documentary on such-and-such a topic. In some cases I know where the details have been taken from: for example, from descriptions of other ethnographic films that I am acquainted with, of which there was no mention in the assessment question or indeed anywhere in the course. But overall, there is much less detail and specifics in this sort of text than one would expect from a person who knew the material. The artificial hand is visible in the specifics (and the general lack of specificity). At the same time, the quality of the generated text and the level of grammar is different (that is, better) from what the average student produces. Evidently, this is how students are taught to write short essays: the text must have a certain structure, and in particular the conclusion, that sums it up, should be the last sentence of the text. AI imitates all this very successfully.

In order to avoid the painful and time-consuming procedure of the disciplinary process for academic dishonesty, I asked the guilty students to analyse the text that they had submitted, noting those places in it that contained incorrect assertions, and to write a new text, this time by their own efforts and in their own hand.

But one case that I encountered deserves a separate commentary. I could hardly understand this Chinese student’s speech, and mostly guessed what he was saying, which contrasted markedly with his

faultless written English, and with the substance of what he wrote, and, moreover, that substance was somewhat different from the typical AI manner. In answer to the question where, and with what group of informants, students today would film a remake of Rouch and Morin's *Chronique d'un été*, instead of the usual ChatGPT mush ("several large cities, social problems, inequality" — there were several essays written in that key) it contained the idea of addressing people who communicated primarily over the net, at a distance. When I invited him to explain, he said that this idea reflected his own experience of relationships at a distance, and that when he did his assessments he wrote in Chinese, translated it into English with Google Translate, and polished the result using ChatGPT. I gave him permission to do so. In certain situations, insufficient knowledge of a language, which would formerly have crippled a person, can now be compensated for using a range of AI technologies, which confirms Lev Vygotsky's well-known thesis that disability is determined by historical and cultural context.

There is no doubt that in our teaching we shall have to teach ourselves to teach students to manage the tools of artificial intelligence in such a way as to expand their capabilities efficiently, effectively realising Joseph Licklider's vision of human-machine symbiosis [Licklider 1960]: the machine bears the weight of the technical tasks, and helps the human being to fulfil the creative ones.

References

Licklider J. C. R., 'Man-Computer Symbiosis', *IRE Transactions on Human Factors in Electronics*, 1960, March, vol. HFE-1, no. 1, pp. 4–11. doi: 10.1109/THFE2.1960.4503259.

CHATGPT-3.5

1

In the social sciences and humanities, neural networks can be used to analyze large arrays of text data, such as literary works, historical documents, sociological surveys, etc. For example, they can help identify trends and patterns in culture, study social interactions, analyze political processes, and much more. Neural networks can also be used to analyze images and sound data in the context of the social sciences and humanities. For example, to study and interpret works of art, architecture, music, or photographs. This allows for deeper research analysis in the field of art, culture, history, and other humanities. Thus, neural

networks open up new opportunities for research in the social sciences and humanities, enriching our knowledge and understanding of various aspects of human activity. I am sure that neural networks will be actively used in various fields of science in the near future. In medical research, they can help in the development of personalized approaches to treatment, diagnosis of diseases, and analysis of medical images. In the field of ecology and climate research, neural networks can be used to predict changes in natural ecosystems and climate. They will also be used in the development of new materials, optimization of production processes, and modeling of complex physical and chemical systems. In addition, neural networks will play an important role in social and humanitarian research, helping to analyze large arrays of text and image data, study cultural and historical processes, analyze social interactions, and much more. In general, neural networks will be widely used in science to solve complex problems, analyze data, and conduct deeper research in various fields.

2

Artificial intelligence can be very useful in conducting scientific research. For example, AI can help analyze large amounts of data, identify hidden patterns, predict trends, and create models. I have not personally used AI for scientific research, but I know that many scientists successfully use it to process data, optimize experiments, and create predictive models. Using AI to compile a bibliography can also be useful. Machine learning algorithms can help automatically find and classify scientific articles, books, and other sources, simplifying the process of compiling bibliographies and ensuring accurate and complete citations. The use of neural networks such as ChatGPT in scientific work can have both positive and negative consequences. On the one hand, such neural networks can help generate new ideas, proposals, and concepts, which can stimulate the creative process and enrich scientific research. However, it is important to remember that such neural networks can generate content that does not necessarily correspond to real facts or scientific data, and they should be used with caution, taking into account the need for additional verification and confirmation of the obtained results.

3

I am convinced that the use of artificial intelligence in education has both positive and negative sides. Among the positives, we can highlight the increased accessibility of education through online courses, the individualization of learning through personalized educational programs, and the improvement of the efficiency of the educational process through the automation of assessment and feedback. However, there is a concern about replacing human teachers and the loss of personal contact in education. In addition, there are questions about the transparency of algorithms, the ethical aspects of using AI and the protection of student data. I personally

have not had any real cases or suspicions of a student submitting work generated by a neural network. However, this raises serious concerns, since the use of neural networks to forge academic work can undermine trust and honesty in the educational environment. There are several ways to identify work created using neural networks, such as analyzing the style and approach to solving problems, checking for plagiarism from the Internet and using specialized programs to detect fakes. Addressing this challenge requires a collaborative effort between institutions and technology developers, including updating assessment methods and training teachers to recognize fake work. Overall, the use of AI in education is a powerful tool for improving the educational process, but requires careful management and the development of appropriate protection mechanisms against unfair use. It is important to find a balance between innovation in education and ensuring the integrity, ethics, and quality of education for all students.

4

The ethical side of using artificial intelligence (AI) in research and teaching is an important issue that requires serious discussion. In research, the use of AI can create authorship issues, since the results obtained with its help can be the subject of disputes regarding who owns the rights to the obtained data and conclusions. In teaching, the use of neural networks can also create additional authorship issues, since students can use AI-generated content, which complicates the process of assessing and monitoring the authenticity of their work. Regarding the rights to the result of the neural network, the question arises as to who should own the copyright to the generated content. It is quite difficult to determine the specific person or organization that owns the rights to the results of the neural network, since it operates on algorithms and data provided to it for training. It may be necessary to develop new legal norms and principles for determining authorship and rights to the results of neural networks. It may be necessary to introduce restrictions on the use of AI in research and teaching to ensure the ethical and legal use of this technology. Restrictions may concern mandatory transparency in the use of AI, the establishment of rules and standards for authorship and citation of neural network results, as well as ensuring the protection of intellectual property rights in the context of AI use. These measures can help balance the innovative potential of AI with the ethical and legal aspects of its use in research and teaching.

AFTERWORD TO THE DISCUSSION

In 2023 “neural network” was the word of the year, the A. S. Pushkin Institute of the Russian Language announced [Dobrunov, Khabarova 2023]. We chose AI and neural networks as the theme for our discussion because of their meteoric appearance, to which it was impossible not to react. In his answers to the Forum’s questions, Albert Baiburin remarked that AI, which was long considered a technological project, is now becoming a social project. The entrance of neural networks into our lives inevitably leads to their appearance in our professional sphere, the humanities and social sciences.

It is intriguing that the participants themselves noted the resemblance between the format of the Forum and requests made to neural networks. Ivan Shirobokov wrote: “There is a certain irony in the fact that as I was answering *AF*’s questions, I several times caught myself feeling that I was playing the role of a chatbot.” Anna Moskvitina formulated the question that she answered in her response as a “prompt” for herself. This makes one think: in time, will our Forum turn into the publication of responses that are no different from artificially generated reasoning? Why is the discussion of topics surrounding scholarship on the pages of a journal necessary at all, if any reader can in a minute produce possible answers on their own screen? In his response, Ivan Shirobokov predicted: “You can send the list of questions proposed to the participants in the Forum to a chatbot, and receive meaningful, perfectly reasonable, but rather boring answers, based on theses which in part repeat themselves.” He also suggested: “Add to the request, proposing that the chatbot should play the role of a British research anthropologist, change the style of its exposition, add a bit of irony and, without fail, mention your favourite ideas in the text, and you will get a noticeably more interesting result.”

Olga Boitsova

Peter the Great Museum
of Anthropology and Ethnography
(Kunstkamera), RAS,
St Petersburg, Russia
boitsova@gmail.com

When we gave the Forum questions to a chatbot as an experiment, we did not input these additional parameters, so the readers have the opportunity to see the unadorned AI answers in all their tedium (and, in the English version, the plodding translation that was the work of the AI drudge). Repetition in the neural network's answers is partly explained by the fact that it was given the questions one by one, and did not "remember" what it had said in its previous output. However, there is quite enough repetition in paragraphs that were generated at one go. The AI answers have not had any editorial correction, apart from the deletion, in the Russian version, of the commas after the word *odnako* [however] at the beginning of a sentence, which is a solecism. The English version has likewise undergone no editing, e.g. conversion of the default US spelling to British conventions.

The examples of the use of AI in the humanities and social sciences provided in the answers supplied by the bot were generalised to the point of being meaningless. Our participants Ilya Utekhin, Ivan Shirobokov, Nadezhda Stanulevich, Kira Kovalenko and Evgenii Sokolov described far more interesting actual cases of applying AI in discourse analysis, physical anthropology, history of photography and linguistics.

So, the answers on the Forum by ChatGPT-3.5 are mostly interesting as an example of generated text for a first acquaintance with the possibilities of neural networks. But one property of the AI response does attract attention. The chatbot offered in its text a model of complete neutrality. It found both pluses and minuses in everything under discussion.

This lack of prejudice is regarded by the participants in the Forum as a strong side of the neural network as a tool. (Andrey Gromov, with a measure of irony about himself, mentioned the human emotion that moved him to answer the editorial board's questions.) AI's lack of bias does good service when data are analysed: "For example, the task of isolating in a text all the collocations that contain an evaluation or a relation, which is not so easy for a human being, can easily be done even by the free ChatGPT-3.5" (Ilya Utekhin). Human attention is selective, but AI works without any distortions when processing visual sources: "When selecting objects a human being is inclined to choose those that are easiest to perceive: large, beautiful, contrasting. The human being is also inclined to focus on the central part of the image. NN does not have such distortions and makes the same effort with the whole block of graphic data" (Andrey Gromov). An interesting and very promising application of the unbiased neural network for the scholar was proposed by Ivan Shirobokov: as a reviewer of his project and an opponent in a discussion with himself.

However, that very neutrality begins to work against the scholar when it is necessary to make a judgment emanating from the context and take a decision about what is important and what is not. If you are defending a thesis close to your heart in your academic work, do you need neutrality of the “perhaps this way, but perhaps the opposite” sort? Natalia Kovalyova writes that “humans will outperform the machines in the areas that require information synthesis or uncovering of hidden / implied information. Humans are superior in identifying gaps in the existing literature, generating hypotheses, contextualising findings, and offering alternative explanations to observable facts. All these activities are essential to the research process, and none has been automated, justifying our reliance on old-school methods, assisted but not replaced by various technologies.”

The participants in the Forum have considered AI as a useful tool, which is more effective in some forms of activity than others. Many are prepared to transfer to AI such time- and energy-consuming operations as the collection of data within a large block, grouping material, extracting keywords from an academic text, composing a brief abstract, or finding stylistic errors in a text. However, Natalia Kovalyova warns us against regarding the writing of a text as “a burden on academics that takes their precious time away from research proper”. One cannot but agree with her. As long as the publication of a text remains the basic means of presenting academic results, writing an academic text is also research work. As we select specific formulations, we prove (or fail to prove) our thesis, and out of these words, perceived (and perhaps subsequently disputed) by our colleagues, our contribution to scholarship emerges.

It is a fault in ChatGPT that it invents facts: even ChatGPT itself notes this in the responses. Natalia Kovalyova describes her own experiment of asking a chatbot to compile a bibliography, which led to the appearance of many made-up works in the list. When they asked the question about using neural networks in working with bibliography, the editorial board had in mind AI such as Research Rabbit, which helps to keep track of works that come out on a particular subject. Of course, before using a tool one must understand how it works, and not set it in motion where it may do harm.

How to cope with dishonest use of neural networks by students? The Forum participants consider that the problem with the submission of generated work lies not in AI, but in the system of supervising their progress. Ivan Shirobokov notes that “the evolution of AI technologies does not create any specific new problems, but simply intensifies the problem of either side’s formal attitude towards education.” As I see it, so long as we require of our students smoothly

written essays without signs of their own working on specific material, we shall have to put up with their handing in work by ChatGPT. The ability to write good prose, with an introduction and a conclusion, is considered an important skill by teachers, and it is in such assignments that people are trained how to do it. And this is the skill that neural networks imitate best. Natalia Kovalyova also writes about this: both students and teachers regard the final text as the outcome of the course, whereas the real outcome of teaching ought to be a better educated person.

In the education system, beginning at school, the less specific a conclusion appears, the better. The requirement for a conclusion in compositions and relevance in projects from junior school on (in Russian schools) makes pupils produce general phrases. Is it surprising that when they start using neural networks they have recourse to a tool that is an expert in blarney and can prove the relevance of anything at all, and finish a text with words that resemble a conclusion as closely as possible? So, the root of the problem appears to be the exaggerated value that our education puts on smooth written language divorced from the thought that lies behind it. But the world has changed, and now smooth boring texts can appear without human intervention. It is time for schools and universities to prepare people for having to think. As Natalia Kovalyova notes, before hurrying to defend existing practices from the (potential) threat of AI, we would do well to stop and examine carefully how far those practices actually favour the attainment of our goals. And Ivan Shirobokov presages, thanks to neural networks, a possible “new spurt of interest on the part of researchers in the problem of the generation of plausible nonsense by natural intelligence”, i.e. people.

If we speak of specific advice to lecturers who encounter dishonest students, the pages of the Forum offer recommendations not to reduce assessment only to written work (Albert Baiburin). Some participants write that an attentive attitude towards students and observing the dynamics of the education of each during the semester will allow one not to worry about the artificial generation of their final work (Anna Moskvitina, Nadezhda Stanulevich). If there is no doing without written examinations when there is a large cohort of students, the Forum participants recommend setting questions which can only be answered by working on the material studied during the course (Natalia Kovalyova, Ilya Utekhin). It was also suggested that one should “feed” one’s assessments to the neural network, so as to understand what sort of text it might produce in response (Natalia Kovalyova, Ilya Utekhin). Furthermore, there are already neural networks in existence, such as AI Content Detector, which will detect whether a text has been generated by AI (Ivan Shirobokov). But the human examiner too, given a certain training, can notice

the fact that a neural network has been used by the student. Among the characteristic features of generated text are clichéd structure and “seamless writing” (Anna Moskvitina), a general vagueness and lack of details, and also errors of fact (Ilya Utekhin). I may add to this final point that a neural network easily falls down when it tries to make up quotations, for example, from Akhmatova’s poetry.

Everything that has been said above relates to the unthinking use of ChatGPT by negligent students. However, if a student has done further creative work on a text obtained with the help of a neural network, that can be evaluated as the student’s own work (Albert Baiburin, Andrey Gromov). Ilya Utekhin gives the positive example of a neural network helping a student to overcome the disadvantage of not knowing the language. Many participants agree that today’s students ought to be taught how to work with AI. Ivan Shirobokov explicitly writes: “Neural networks cannot be prohibited, but they can, and should, be made part of the system of education.”

Anna Moskvitina offers interesting examples of the use of AI in teaching students Swahili. She notes that a certain element of cliché in the output of neural networks is not an obstacle in the field of education. In describing her experience, Anna Moskvitina calls AI an “interactive textbook”. I think this is no accident. At the present level of its evolution artificial intelligence cannot become a teacher able to replace a human being (Albert Baiburin puts it like this: “I cannot imagine the possibility of the teacher being completely replaced by an AI analogue, at least because the human relationships between students and teachers cannot possibly be replaced by any sort of analogues”), but AI can provide help to human teachers. Other participants in the Forum have noted that using AI in education allows personalised teaching, giving each student the opportunity to dwell in greater detail on those sections that he or she does not understand, and skip over those he or she does. ChatGPT also distinguished itself in its answer to this question. It named an advantage that nobody else had remembered — “an improvement in the effectiveness of the educational process through the automation of assessment and feedback”.

In the questions about the authorship of research our participants agree that while neural networks must certainly be mentioned in a scholarly work, the place for that at present is the “Methods” or “Acknowledgments” section. Kira Kovalenko and Evgenii Sokolov expressed it like this: “Since in this case the neural network only systematises the data, and does not draw conclusions, the intellectual property rights unquestionably belong to the authors of the article, who have applied neural network mechanisms for data processing.” They make an exception for cases where the neural network is used to draw conclusions, but no such cases are examined in the responses

to the Forum. As Yuri Berezkin writes, “The humanities, even when based on big data, presuppose not only knowledge, but understanding.”

None of the participants would be prepared to give over the work of scholarship to AI in its entirety. Ivan Shirobokov expressed an opinion which I think is shared by many researchers: “If we ever outsource scholarship to artificial intelligence, it will be a suicidal step if not for humanity, then for scholarship, since it will change its very essence as a form of human activity.”

Natalia Kovalyova draws attention to the “process — result” dichotomy in her answer to the question about education. Scholarly activity is also a process as well as a result. Obtaining new knowledge and discovering something that nobody knew before give pleasure in themselves to the researcher. If we give scholarship away to AI completely, we shall deprive ourselves of that pleasure.

The editorial board remains confident that AI will not at present replace either the authors of the Forum, nor those researchers who publish their articles in other parts of our journal, nor the editorial board itself, and they thank all the participants warmly for their responses.

Olga Boitsova

References

Dobrunov M., Khabarova V., ‘Institut Pushkina nazval glavnoe slovo goda v russkom yazyke’ [The Pushkin Institute Has Named the Main Word of the Year in the Russian Language], *RBK*, 2023, Dec. 18. <<https://www.rbc.ru/society/18/12/2023/657ffc1d9a79472cfba66e12>>. (In Russian).

The response by ChatGPT was translated by ChatGPT. The remaining responses originally written in Russian were translated by Ralph Cleminson.