M

УДК 004.8; 004.94; 001

A. V. Mishchenko
Institute for Research in Computer Science and Automation, Rocquencourt, France.
Domaine de Voluceau, Rocquencourt - BP 105, 78153 Le Chesnay Cedex, France.

"CIVILIZATION-LICHEN" AND "CIVILIZATION-MAGNET": MODELING SYMBIOSIS OF HUMANITY AND ARTIFICIAL INTELLIGENCE, WITHIN THE MIND-MATTER THEORY

А. В. Мищенко
НИИ информатики и автоматики, г. Роканкур, Франция.
Домэн Волюсо, Роканкур - BP 105, 7815, Франция.

«ЦИВИЛИЗАЦИЯ-ЛИШАЙНИК» И «ЦИВИЛИЗАЦИЯ-МАГНИТ»: МОДЕЛИРОВАНИЕ СИМБИОЗА ЧЕЛОВЕЧЕСТВА И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, В РАМКАХ ТЕОРИИ МЫСЛЯЩЕЙ МАТЕРИИ

О. В. Міщенко
НДІ інформатики і автоматики, м. Роканкур, Франція
Домен Волюсо, Роканкур - BP 105, 7815, Франція.

«ЦИВІЛІЗАЦІЯ-ЛИШАЙНИК» ТА «ЦИВІЛІЗАЦІЯ-МАГНІТ»: МОДЕЛЮВАННЯ СИМБІОЗУ ЛЮДСТВА ТА ШТУЧНОГО ІНТЕЛЕКТУ, У РАМКАХ ТЕОРІЇ МИСЛЯЧОЇ МАТЕРІЇ

The types of future coexistence between human society and future general AI depend both on the types of AI being developed today and types of existing societies. The range of possible partnerships, suggested by science fiction, range from elimination of humans to happiest life under the care of AI. This paper considers evolution of two types of AI (open- and closed-access, controlled by society and authorities, correspondingly), as well as corresponding societies.
Key words: artificial intelligence, modeling, mind-matter theory, sociology, evolution, civilization, human-AI interaction, open-source.

Будущее сосуществование человечества и общего ИИ зависит от разрабатываемого сейчас ИИ и типов современных обществ. Фантастика предлагает широкий спектр будущего - от уничтожения людей до безбедного существования под опекой ИИ. Эта статья рассматривает эволюцию открытого и закрытого общества и соответствующие им открытый и закрытый ИИ (контролируемый обществом и уполномоченными организациями соответственно).
Ключевые слова: искусственный интеллект, моделирование, мыслящая материя. социология, эволюция, цивилизация, человек и ИИ, открытый доступ.

Майбутнє співіснування людства і спільного ШІ залежить від того, який ШІ розвивається зараз, і від типів сучасних суспільств. Наукова фантастика пропонує широкий спектр майбутнього, від знищення людей до комфортного існування під опікою ШІ. У даній статті розглядається еволюція відкритих і закритих суспільств і відповідних їм відкритих і закритих ШІ (які контролюються суспільством і уповноваженими організаціями відповідно).
Ключові слова: штучний інтелект, моделювання, мисляча матерія, соціологія, еволюція, цивілізація, людина та штучний інтелект, відкритий доступ.

1 Actuality of the problem

The question whether the development of strong general AI should be strictly controlled by government or, on the contrary, open-source and, therefore, society-controlled, is becoming increasingly important. For example, at the mobile technology congress in Barcelona in Mars 2024, Brad Smith, president of Microsoft (the main owner of GPT chat and other advanced developments in the field of general AI) announced new «AI Access Principles», according to which, all developments of Microsoft in the field of strong AI will become open source and available for study and use by anyone in any projects [1].

The previous (first) event of this kind occurred in 2023, when another giant of AI development in the field of social networks opened its large language model «Llama 2» [2].

Taking into account the AI's ability to create weapons and control society, it seems that putting the components of general AI into the public domain is analogous to putting all the necessary ingredients of nuclear weapons on some «open warehouse», so that anyone can construct his own bomb. It may seem that, may be, more wisely is to impose a strict governmental control over development of strong general AI.

But, on the other hand, this analogy from previous, industrial age, may be incorrect in our information age. To understand which type of AI is more dangerous, the modeling within the theory of Mind-Matter [3] may be useful. Such modeling implies construction of mathematical and computer models, structured according to informational meta-model of Mind-Matter [3]. The results of application of this meta-model and simulations of symbiosis of different types of AI are discussed in this paper.

2 Methodology: Mind-Matter meta-model

In this section we remind the structure of interaction between information and matter in the special type of organization of matter, defined in [3] as "Mind-Matter".

As defined in [3], the Mind-Matter (MM) is a such type of organization of matter, where the methods of its organization are modeled (created, improved or tested) separately from the organized matter. In other words, any compound (structured) matter is a Mind-Matter if it is composite, consisting of two parts, named in [3] IC and OC.

IC is "information component" (artificial or natural intellect) creating models, which holds and processes the information about organization of OC. OC is "organized component (environment, reality), being organized according to these models.

Since these IC and OC components of MM interact, the parts providing this interaction can also be discerned. On the one hand, it is the "creating component" (IC-to-OC), providing the process of organization of OC. On the other hand, it is the "observing component" (OC-to-IC, representation of environment), providing information from the reality(OC) to the models(IC) component.

If some mathematical or computer model for simulation some part of reality has such "informational" component and "organized" component, behaving as described above, this model is said to be structured according to the "Mind-Matter meta-model". Corresponding part of reality is said to be the "Mind-Matter".

Figure 1 illustrates information-exchange between environment (OC), information component (IC) and representation of environment (OC-to-IC). As shown in [3], the MM meta-model can be used to construct the concrete models and computer programs, simulating this information-exchange in case of concrete examples of evolution of intellect (whether artificial [3], natural, hybrid or group intellect).

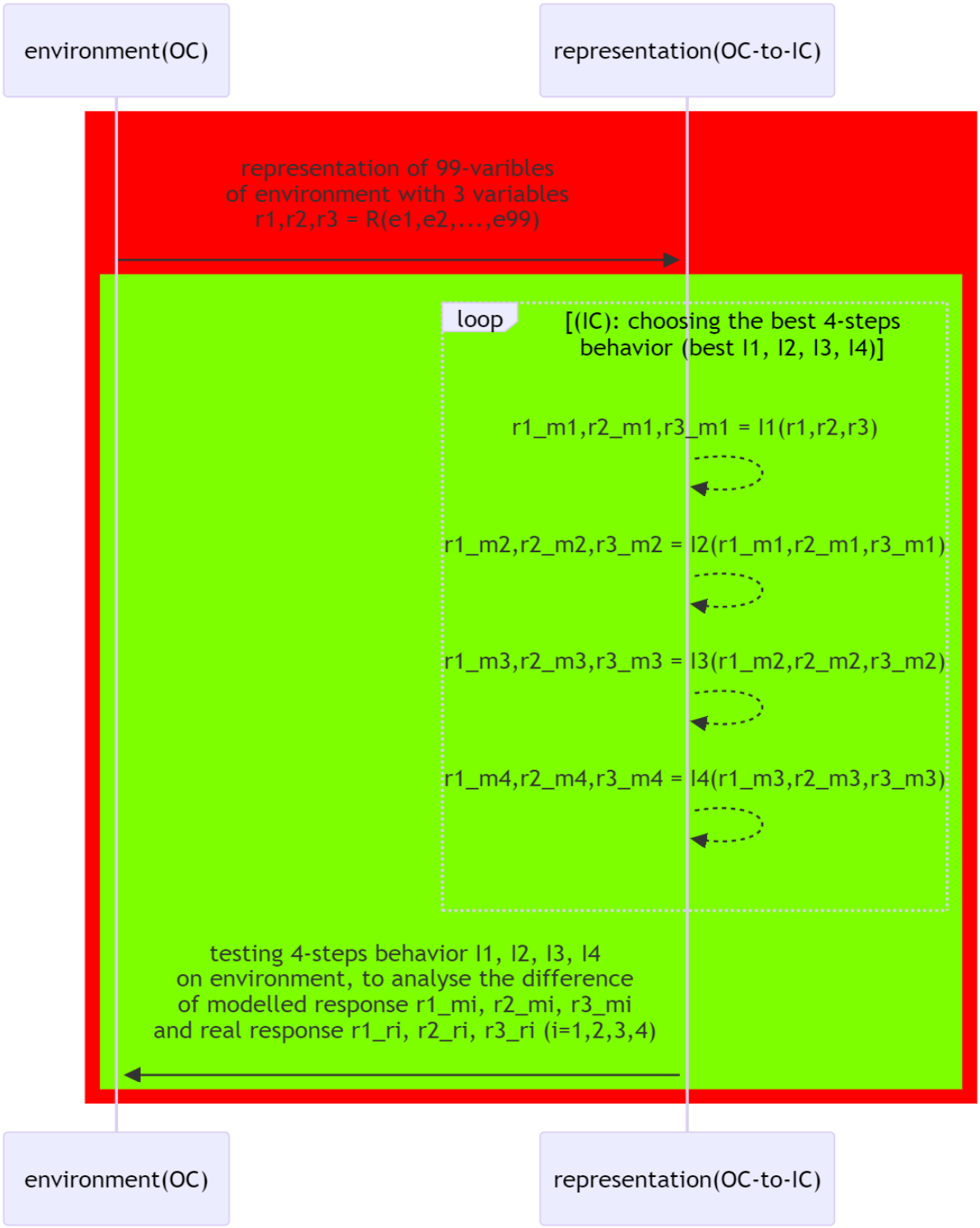

Figure 1 – Example of Mind-Matter model with 99 variables of environment (OC) and 3 variable of representation of environment (in IC). Information-exchange between environment (OC) and information component (IC), as well as processing models in IC (discretee 4-steps planning used in this example correspond to planning-ability equal to 4). The representation of environment (OC-to-IC) is also shown

The biggest known example of Mind-Matter is the entire mankind (its knowledge being IC) and environment (reality), organized by it (OC). This MM is learning how to change and organize reality around. In this case, the environment (OC) is all reality interacting with humankind, behaving according laws of physics and other sciences. One example of models of reality in IC are physical models, processed in IC within corresponding physical theories and tested against reality to correct corresponding theories.

Both information(knowledge) processing in IC and organization of environment in OC are structured according Mind-Matter meta-model and, therefore, represent the Mind-Matter.

In [3], on the contrary, one of the simplest environments are considered: OC consists from chess-board and chess-figures, behaving according chess laws or horse-move-game laws (playing the role of physical laws of reality). This OC is organized by self-learning AI (IC). The aim of this AI is to observe, finally, a winning-state on a chess-board. The ability of such AI is measured in number of game-steps which it can plan ahead (in chess-simulators this is called «depth of a decision-tree»). The relation of self-learning AI and its environment (computer program, simulating chess-laws or horse-move-game laws) is structured, also, according to the Mind-Matter meta-model.

3 Methodology: Mind-Matter meta-model for modeling of abstract behavior

As noted in the previous section, in cases of a concrete examples of environment (which can be any reality which is learned and organized by intellect), it is possible to make a concrete specific model and simulate how this intellect learns its behavior on environment organization. Some concrete examples are also given above.

Another such concrete example, for instance, is any actor (human being) learning how to organize his proper life and reality around. In this case environment(OC) is society, behaving according social laws. The IC is human's consciousness(created in social interactions), developing models of himself and his place in society, as well as planning his behaviour.

The aim of a human-actor is to use his ability to plan ahead in order to observe only positive-for-him states of the neighboring-environment. The amount of possible positive states of the environment, discovered by human-actor is described later as the region of freedom within which this actor can move. In following sections we model behavior of multiple independent Mind-Matter actors (many humans in society). These actors can mimic successful behaviors of their neighbors to reach the same positive state.

In the above mentioned case of concrete society and human, as well as in other concrete specific cases (such as entire mankind or chess-games, mentioned above) the environment was not an abstract, but a concrete one (such as chess, society or universe), therefore it was possible to make a computer program, simulating how intellect organizes this environment, intending to observe a subset of environmental states which are positive for him ($[r_1,r_2,r_3]^{POSITIVE}$ in 3D case) and avoid a negative ones ($[r_1,r_2,r_3]^{NEGATIVE}$).

For a simple concrete cases (such as chess), the corresponding program was implemented [3]. For complex cases, such as mankind or behavior of human in society, such complex programs can be the directions of future work. This work will include modeling of environment, behaving according laws of physics, sociology, etc.

In this paper we consider not a concrete, but an abstract environment(OC), therefore it is impossible to model how intellect (artificial, natural or hybrid) learns its concrete behavior. However, as computational procedures can be represented as functions [4], [5], it

is possible to analyze the abstract functions of transformation of information from environment (OC) to its representation in the intellectual models (IC). Such functions (in approximation of functions of 99 and 3 variables) are illustrated in Figure 1.

The example of approximation shown in a Figure 1 is as follows. Transformation from 99 observed environment variables e1,e2,...,e99 to only 3 environment-representation variables is made by the abstract function R: r1,r2,r3 = R(e1,e2,...,e99).

Another important abstract functions of this example is the ones used for testing these models in IC. In this example, 4-steps behavior is tested inside IC: $r1_1,r2_1,r3_1$ = I1(r1,r2,r3), $r1_2,r2_2,r3_2$ = I2($1_1,r2_1,r3_1$), $r1_3,r2_3,r3_3$ = I3($r1_2,r2_2,r3_2$), $r1_4,r2_4,r3_4$ = I4($r1_3,r2_3,r3_3$). Remind that the planning-ability of Mind-Matter actor is measured in amount of behavioral steps that it is able to plan, so, the Figure 1 illustrates the case where this planning-ability is 4. Note that continuous-systems can be also reduced to discrete-steps planning, with a help of procedures, such as in [6],[7].

In two sections below we consider hybrid civilization with human-actors and AI-actors. We analyze how restriction on representation of reality (whether simple restrictions on function R(e1,e2,...,e99) or more complicated policies of model functioning) affect on regions of freedom of human-actors.

These simulations with 2 types of AI-actors (enabling and restrictive AI) reveal 2 types of humans-AI symbiosis, called below «Lichen-civilization» and «Magnet-civilization» correspondingly.

4 Details of the simulation of human-AI symbiosis

Let us consider two types of Mind-Matter (humans and AI). Human-actors are characterized by constant planning-ability (limited by brain capacity). The AI-actors are characterized by growing planning-ability (as it is growing with the progress). Note, that, despite there are always some theoretical limitations on calculations speed [8], nethertheless, as compared with human-actor, growth of capacity of AI-actor can be considered unlimited.

In particular, for human-actors of Mind-Matter, we consider the following approximation: 99 variables for environment(e1,e2,...,e99), 3 variables for environment-representation (r1,r2,r3) and cubic polynomial function R for transformation from 99 observed environment variables to only 3 environment-representation variables: r1,r2,r3 = R(e1,e2,...,e99).

Along with this, other functions, illustrated in Figure 1, correspond to testing these models in IC. In Figure 1, 4-steps behavior is tested inside IC (I1, I2, I3, I4).

The modeling of abstract behavior consists of dividing the environment space into random connected positive and negative regions. Transformation from 99 observed environment variables e1,e2,...,e99 to 3 environment-representation variables results in positive and negative regions in 3D space ($[r_1,r_2,r_3]^{POSITIVE}$ and $[r_1,r_2,r_3]^{NEGATIVE}$, correspondingly).

Figure 2 (left) illustrates modeling the behavior of 8 human-actors inside $[r_1,r_2,r_3]^{POSITIVE}$, avoiding $[r_1,r_2,r_3]^{NEGATIVE}$ regions. The trajectories of these human-actors are shown in blue, orange, magenta and cyan color. The $[r_1,r_2,r_3]^{NEGATIVE}$ regions are shown in gray. The $[r_1,r_2,r_3]^{POSITIVE}$ regions occupy the rest, non-grey 3D space. Mimicking behavior of neighboring human-actors leads to some similarities in their trajectories.

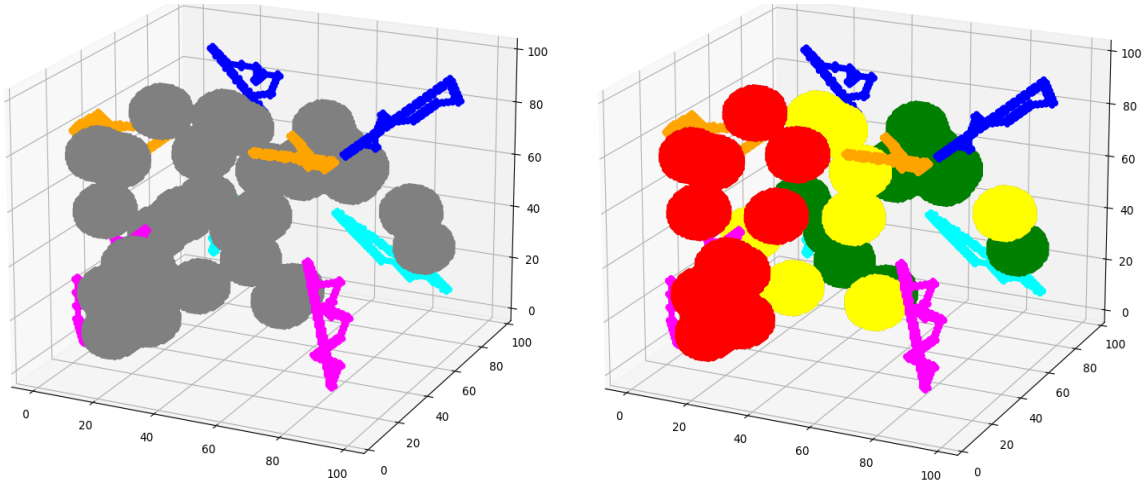

Figure 2 – Left: Simulation of behavior of 8 human-actors inside $[r_1,r_2,r_3]^{POSITIVE}$, avoiding $[r_1,r_2,r_3]^{NEGATIVE}$ regions. The trajectories of these human-actors are shown in blue, orange, magenta and cyan color. The $[r_1,r_2,r_3]^{NEGATIVE}$ regions are shown in gray. The $[r_1,r_2,r_3]^{POSITIVE}$ regions occupy the rest, non-grey 3D space. Environment-representation variables are measured from 0 to 100% from their maximum value.
Right: The example of restriction «r3>66%». The area satisfying r3>66% is highlighted in green. The «a little bit» unsatisfactory area (33%<r3<66%) is highlighted in yellow. The «very» unsatisfactory area (r3<33%) is highlighted in red.

5 Results of simulation of enabling-AI: "Civilization-lichen"

Details of this simulation are illustrated in a Figure 2 (left): differently colored trajectories correspond to different 4-steps behavior from Figure 1: $r1_1,r2_1,r3_1$ = I1(r1,r2,r3), $r1_2,r2_2,r3_2$ = I2(1_1,$r2_1$,$r3_1$), $r1_3,r2_3,r3_3$ = I3($r1_2$,$r2_2$,$r3_2$), $r1_4,r2_4,r3_4$ = I4($r1_3$,$r2_3$,$r3_3$). The simple linear I1, I2, I3, I4 functions are used, providing a shift in one direction (the longer straight line of each trajectory). When applying these steps to real 99-variables environment(e1,e2,...,e99) and back-projecting the result to 3-variable observation (r1,r2,r3), the straight-line trajectories are replaced by more complicated trajectories, shown in a Figure 2 (left), each along with its straight-line trajectory. If this more complicated trajectory do not intersect with $[r_1,r_2,r_3]^{NEGATIVE}$ regions shown in a Figure 2 (left) in gray, then the test of modelled 4-steps behavior on a real environment was successful. The amount of all successful (positive) trajectories, reached using a given functions and given number of steps (4-steps in our case), correspond to a region of freedom of each MM-actor.

Apart from a very simple simulation with 4 linear functions I1, I2, I3, I4 and 3D observation space (useful for illustrations such as in a Figure 2), we used also a more complex simulation with 7 cubic functions I1,...,I7 (corresponding to 7-steps planning) and 20D observation space (r1,...,r20). The results of simulation of growing regions of freedom for 100 human-type MM-actors are illustrated in a Figure 3 (left), with 20D space of representation of environment projected into 2D image-space. The area of each circle correspond to the region of freedom of this human-actor (to the volume in 20D space r1,...,r20, covrered by reached positive, successful trajectories). This simulation results in building small non-intersecting regions of freedom for all MM-actors.

Another simulation includes 100 human-type and 4 AI-type MM-actors located, then projected to image space, at middles of 4 sides of image (in a Figure 3 only human-type MM-actors are shown). Remind that AI-actors differ by growing number of planning steps (that corresponds to growing AI capabilities), whereas number of planning steps of human-type MM-actors remain constant (7 in this simulation).

In this case, mimicking steps of neighboring AI-actors allow human-actor to explore more successful trajectories and building bigger regions of freedom, as illustrated in a Figure 3 (center). Finally, the number of explored successful trajectories stops growing, being limited by negative regions themselves, as well as by the accuracy of cubic approximation for functions I1,...,I7 and accuracy of representation of 99D environment with smaller, 20D observation space (Figure 3, right).

The main conclusion of this simulation is that growing of regions of freedom of MM-actors with constant-abilities (human-actors) is not restricted but, on the contrary, enabled by mimicking behavior of growing-ability MM-actors (AI-actors). This corresponds to growing possibilities of humans when using solutions, developed by AI.

This type of human-AI symbiosis is named in [9] “The Civilization-lichen”. We can imagine that such dependency of humans on AI leads to a symbiotic society, where each person delegates all survival questions to AI – similarly as it happens in algae-fungi symbiosis, which we see as lichen.

Human in the “Civilization-lichen”-symbiosis lives in his own comfortable world, designed and maintained for him by global AI – in the same way as each algae in a lichen live in its own capsule having a sea-like microclimat, despite the sea around it does not exist anymore. Instead, a network of mycelium of fungi creates this sea-like microclimat for it and develops all food-providing solutions (similar as AI do for humans).

Note that the second result of simulation (limitation of regions of freedom by the internal limitations on accuracy of I1,...,I7 and accuracy of representation of environment) may be seen as the fact that in such “Civilization-lichen” every human will take as much responsibility for his own life as his own «laziness» allows him. More precisely, not laziness, but his determination to change his biological basis(accuracy of I1,...,I7), thus turning from an “algae” into some kind of “new branch of lichen”[9].

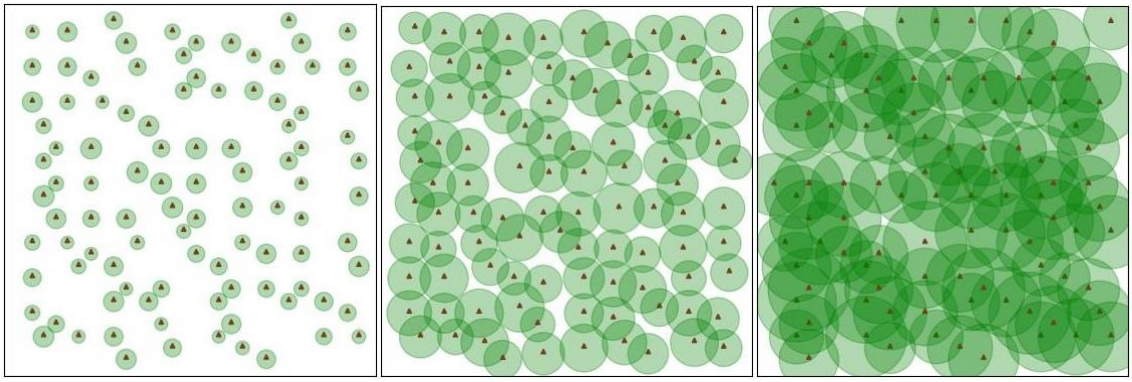

Figure 3. Visualization of modeling the development of human-civilization (left) into hybrid AI-human civilization of type «Civilization-lichen»(right).
The region of freedom for each human-actor, is represented by a green circle, with radius, corresponding to the amount of reached positive human-actor trajectories.

Note that, however, not every type of AI has the characteristics of enabling AI or can grow into such enabling AI, with which humans can behave freely, delegating all their survival questions to AI.

As shown in [10], only the open-source and commercial AI, visible and controlled by society can have such a potential. Such AI is called in [10] a «public AI». Such public AI includes both open source AI and commercial AI, which has closed-source but open specifications and goals. All public AI is open in this sense and available to the public.

All «not-hidden» parts of paid and free software (that we well-know and use) form an environment for the growth of public AI (the development of such an environment is the commercial benefit of making AI open access, which started to happen with public AI [1],[2]).

This open, competitive and simultaneous developments ensure the most sustainable, evolutionary development of AI. The beginnings of public AI are products like chat GPT, applications like Smart Home, Open Commercial Maps (such as GoogleMaps or Mapquest), as well as open-source maps (such as OpenstreetMap) and so on.

As noted above and in [10], this public-AI has a potential to grow into enabling general AI, building, finally, the «Civilization-lichen».

In the section below we consider another, non-public AI, which may lead to alternative type of civilization.

6 Results of simulation of restricting-AI: "Civilization-magnet"

The summary of the simulation from the previous section can be formulated as follows. In the case of absence of restrictions imposed by AI on humans, the regions of freedom of Mind-Matter are growing and the mimicking neighbors behavior does not affect this growing.

In this section we simulate the opposite case: presence of restriction on the representation of reality, imposed on humans by another type of AI (restricting AI).

In a case of modeling the symbiosis of humans and restricting AI, we used the same set-up of the simulation of human-AI symbiosis, described above, with the difference that the AI-actors impose on neighboring human-actors some restrictions on their behavior, such as on their representation of environment. These restrictions may be simple (such as restriction on outputs of R(e1,e2,...,e99) functions, used below, or more complicated policies on functioning the OC-to-IC module).

In this research, we consider the simplest example of such restrictions: equality or inequality restriction for the value of some representation-variable or combination of representation-variables.

For example, in a simulation illustrated in a Figure 1, transformation from 99 observed environment variables e1,e2,...,e99 to 3 environment-representation variables r1,r2,r3 = R(e1,e2,...,e99) is used. The example of restriction r3>66% is illustrated in a Figure 2(right): the area satisfying r3>66% is highlighted in green. The «a little bit» unsatisfactory area (33%<r3<66% instead of r3>66%) is highlighted in yellow. The «very» unsatisfactory area (r3<33% instead of r3>66%) is highlighted in red.

This policy impose that if some neighboring human-actor arrive at representation of reality with 33%<r3<66%, then the representation function of this human-actor is corrected by averaging with the representation functions of his «most correct» (having biggest r3) neighbor. If some human-actor arrive at representation of environment with r3<33%, then the representation function of this human-actor is replaced by average of representation functions of his «correct» neighbors.

Here, the average of some polynomial representation functions R1 and R2 mean using function with vector of coefficients (at different powers of e1,e2,...,e99) equal to average of vectors of coefficients of functions R1 and R2.

Similarly as with modeling «Lichen-civilization», we used not only a very simple simulation, useful for illustrations such as in a Figure 2, but also more complex simulation with 7 cubic functions I1,...,I7 (corresponding to 7-steps planning) and 20D observation space (r1,...,r20). In this 20D observation space, several variants of restrictive-AI-actors in different locations were considered.

The results of simulation for 100 human-actors and 3 or 4 restrictive AI-actors is illustrated in a Figure 4. Upper row represent 4 AI-actors, lower row represent 3 AI-actors. The area of each circle correspond to the volume in 20D space (r1,...,r20), covrered by reached successful trajectories (region of freedom of human-actor).

Before adding to simulation these restrictive AI-actors (Figure 4, left column), the human-actors build small non-intersecting regions of freedom, corresponding to purely human civilization (same as in a Figure 3, left).

Then, adding restrictive AI-actors (Figure 4, center column) results in similarity of functions R(e1,e2,...,e99) of representation of reality (alignment of vectors of coefficients of these functions). The degree of such vectors alignments is illustrated as the length of vectors.

The interesting observation is the alignment of representations of reality, which, with time, become dependent solely on the policy of the nearest restrictive-AI (center of the Magnet-civilization). As a result, everything become divided into regions of influence of different centers (different Magnet-civilizations). This process looks similar to mutual magnetization of human-actors in the strong magnetic field, generated by restrictive AI. Therefore, this type of symbiosis of humans and AI is called «Civilization-magnet».

Another interesting observation is the shrinkage of regions of freedom. As it is visible in a Figure 4 (central column), the volume, covered by reached successful trajectories (corresponding to the area of each circle) become smaller as compared to purely human-civilization in a Figure 4 (left column).

Another interesting observation is that, with increase of restrictions imposed on humans by AI, the total volume of regions of freedom is decreasing monotonously, tending to zero after some threshold of maximal restrictions(Figure 4 right column). The most interesting here is that even growing planning-ability of AI-actors does not provide the growing the regions of freedom, as it was in simulation of enabling AI (AI without restrictions imposed on behavior of human-actors). On the contrary, with the presence of policies imposed by restrictive AI, the regions of freedom shrink.

In [10] these results are discussed from today’s viewpoint. In particular, it is discussed that this restrictive AI can evolve from a hidden AI, developed for surveillance and incorrect-behavior detection, and operating today secretly from society on a dedicated hardware (such as network of smart surveillance cameras) or hiddenly on people’s hardware or sites.

Small temporary programs of such hidden AI can also be developed (in the same way as computer viruses) by fraudsters. Large, long-term projects of secret-AI can be developed by the state and intelligence agencies. The goals of such AI could be not only the fight against the above-mentioned fraudsters or other criminals, but also the identification and neutralization of opposition to preserve the existing government for a long time.

This type of AI is developed secretly and run undetectably, therefore, it is more difficult to give examples of specific programs, than in a case of public-AI. However, sometimes some of its results become visible – in particular, due to its errors.

For example, even now, those arrested by the error of mass facial recognition systems cannot be easily released by the police, who have neither the authority nor the desire to take responsibility for such release. Another example is that some restrictive-AI

alternatives to the «Smart Home» may include the ability to forcibly close the entrances of houses and offices. Similar closures, most likely, occurred during COVID in China [11], [12]. Thus, several forcibly locked people were burned in a fire accident from which they were not able to escape [13], [14].

Another example is North Korean computers and mobile devices equipped with a hiddenly functioning program «Red Flag» that sends screenshots to enable (if unwanted activity is detected) something like an automatic denunciation to state security agencies [15].

The evolution of such secret AI can proceed naturally by increasing its percentage in government systems. In the future, this type of AI could almost completely replace the entire work of not only state security agencies, but also courts and correctional institutions.

Of course, even at maximal replacement of humans by restrictive-AI, at the highest level, this AI will not choose by itself, but offer possible alternative solutions to government (similar to how current specialists do). The government itself (or a single ruler) will be able not to govern, but just to set, according to their own taste, the general paradigm for the development of society (which domains of economics to develop, against whom to start wars, etc). The task of AI, in this case, will be to prove to ordinary people the correctness and necessity of any chosen direction (impose the «correct» representation of reality) and to implement this direction, using the future analogs of models of society [16-19]. As mentioned above, this type of civilization is called in [10] the “Civilization-magnet”, by analogy with how all objects around a magnetic pole become magnetized in a certain direction. In a Figure 4 upper and lower rows represent division of space between 3 and 4 such poles, correspondingly, each preforming magnetization in its own direction.

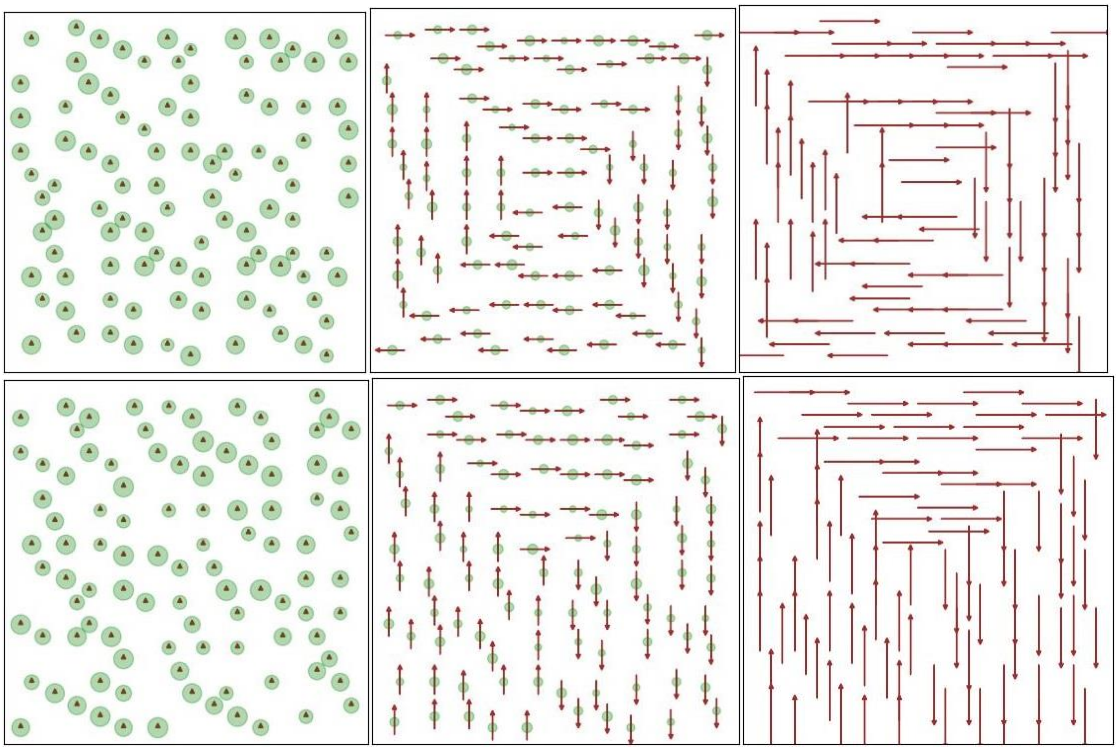

Figure 4 – Visualization of modeling the development of human-civilization(left column) into hybrid AI-human civilization of type «Civilization-magnet»(right-column). Upper-row: 4 restrictive-AI at middles of 4 sides of image. Lower-row: 3 restrictive-AIs at right, upper and lower sides of image. The alignments of representations of reality («magnetization» of human-actors) are illustrated by length of vectors. 4(or 3) directions of vectors correspond to 4(or 3) representations of reality, imposed by 4(or 3) restrictive-AIs

As noted in [10], theoretically, with a sufficient level of development of robotics, the presence of people themselves in such a civilization is not necessary (except, of course, the political elite or even one, for example, immortal ruler). In this case, one of the asymptotic stable solutions is when everything is divided into zones of influence located in different “magnetic fields” of certain rulers [10]. Without humans, these zones can be maintained entirely by advanced AI and robotics, implementing the ruler’s will.

However, in the presence of neighboring civilization of Lichen-type (resulted by not a restrictive, but a public, enabling AI), such completely-magnetized situation is not reached, as described in the next section/

7 Results of simulation of neighborhood of 2 types of AI and corresponding 2 types of civilization

In a case of modeling the neighborhood of “Civilization-lichen” and “Civilization-magnet”, we used the same set-up with 3 and 4 restrictive-AIs, as described above for «Civilization-magnet» (and illustrated in a Figure 4) with addition of enabling-AI at the center (Figure 5, upper row) or at middle of left side (Figure 5, lower row).

As it is visible in a Figure 5, the presence of “Civilization-lichen” maintained by enabling-AI, results in the situation where regions of freedom of some human-actors remain (not all regions of freedom dye-out to zero as it was with “Civilization-magnet”). The human-actors in the zone of influence of “Civilization-lichen” (near enabling-AI in the center for upper row of a Figure 5 or in the middle of left side for lower row of a Figure 5) save both their regions of freedom and remain non-magnetized.

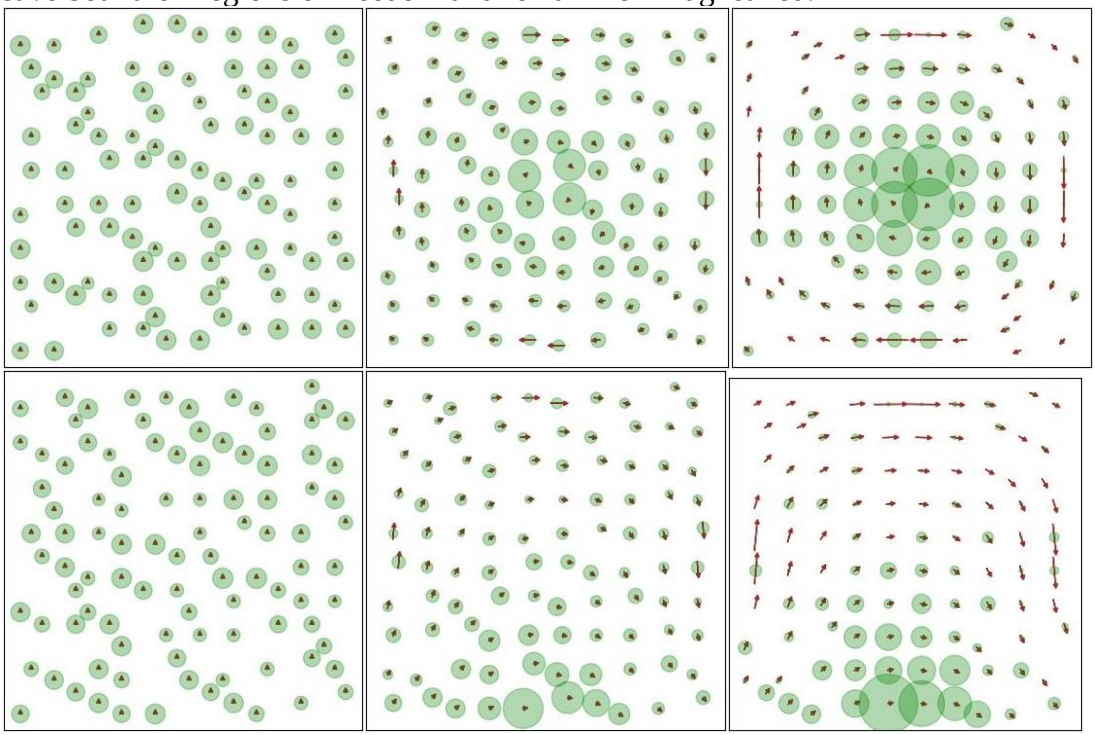

Figure 5 – Visualization of modeling the neighborship of 2 hybrid AI-human civilizations: «Civilization-lichen» and «Civilization-magnet». Left-column: human-civilization before AI involvement (similar to left of Figures 3 and 4). Central-column: early-stage of neighborship of Civilization-lichen and Civilization-magnet. Right-column: final-stage of neighborship of Civilization-lichen and Civilization-magnet. Upper-row: restrictive-AIs of Civilization-magnet are located at 4 middles of 4 sides of image, enabling-AI of Civilization-lichen is located at the center of an image. Lower-row: restrictive-AIs of Civilization-magnet are located at 3 middles of 3 sides of image (upper, right and lower side), enabling-AI of Civilization-lichen is located at middle of left side of an image

The human-actors in the zone of influence of “Civilization-magnet” (near restrictive-AIs) fail to save their regions of freedom and become magnetized according to policies of the nearest “Civilization-magnet” (nearest restrictive-AI).

From today’s viewpoint (as it is done in [10]), this result corresponds to the necessity of public AI, controlled by society, that Microsoft's president spoke at the mobile technology congress in Barcelona in Mars 2024. After announcing decision to make Microsoft’s general AI into open-source and responding to fears that private attackers or totalitarian states could use the open source code of strong AI, he mentioned that this will help «using AI as a defense shield even if someone uses it as sword» [1].

In the section below we discuss the stability of configuration of neighborship between these 2 types of civilization. In particular, the next section first describes the instability of the neighborhood of 2 civilizations in case of temporary disappearance of «Civilization-lichen» (with rapid transition into purely Magnet-civilization). Secondly, next section describes inability of Civilization-magnet to impose its restrictions and reduce regions of freedom in case of changing its «magnetization direction» (direction of changing the representation function).

8 Results of simulation of transition between 2 types of AI and corresponding 2 types of civilization

In a case of modeling the transition of neighborship of two civilizations into one “Civilization-magnet”, we used the same set-up with 3 restrictive-AI-actors and 1 enabling AI-actor, as described in the section above and illustrated in the lower row of a Figure 5.

As it is visible in a Figure 6, as soon as the enabling-AI is removed, the influence of the “Civilization-lichen” is decreased rapidly: first the magnetization of human-actors is rapidly increased (Figure 6, center) and then the regions of freedom of all human-actors dye-out to zero, leading to “Civilization-magnet” (Figure 6, right).

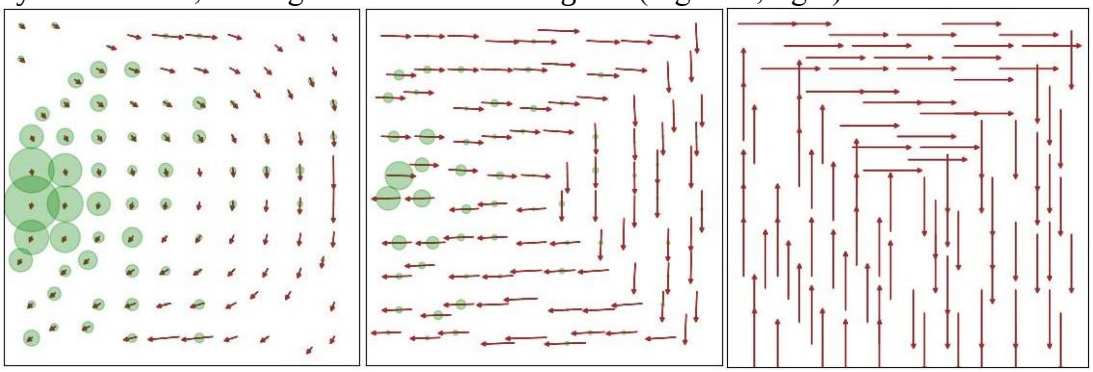

Figure 6 – Development of neighborship of «Civilization-lichen» and «Civilization-magnet» (showed here in left image and in a Figure 5 in lower-right image) into pure «Civilization-magnet» in a case of disappearance of enabling-AI assuring existence of «Civilization-lichen».

Another interesting instability concerns early stages of “Civilization-magnet”. This happens if the regions of freedom are not yet completely died out and if the centers of “Civilization-magnet” or directions of magnetization (the direction of changing policies) are not constant but changing over time.

To simulate this situation, we used the same set-up with 4 restrictive-AI-actors, as described above and illustrated in the upper row of a Figure 4.

As it is visible in a Figure 7, if the restrictive-AI (center of “Civilization-magnet”) frequently changes its policy, or, the center is changed itself, then, such human-AI symbiosis lead (during just a little bit slower time) to the state similar to «Civilization-lichen», with regions of freedom just a little bit smaller than those in a Figure 3. The Figure 7 shows that changing center or changing directions of magnetization is able to maintain the size of regions of freedom and keep the magnetization small for an unlimited amount of time.

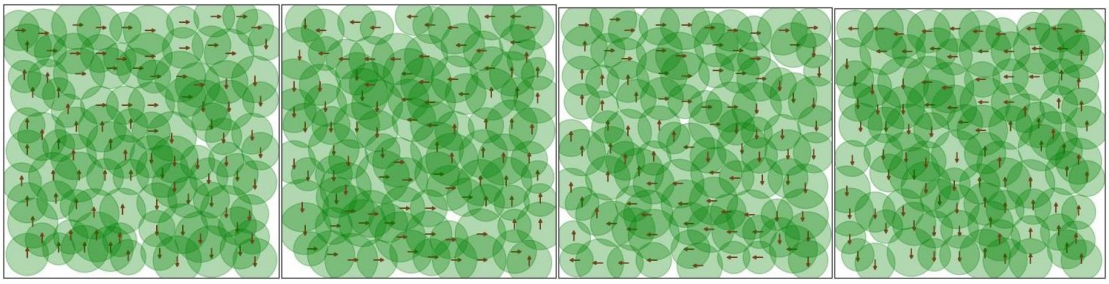

Figure 7 – Inability of «Civilization-magnet» to align representations of environment of all human-actors (to provide the maximum magnetization as in left of the Figures 3 and 6) and inability of «Civilization-magnet» to reduce the regions of freedom – in case if the direction of change in environment-representation (direction of «magnetization») is changing over time.

This result mean that the ability of the «Civilization-magnet» to reduce the regions of freedom of human-actors may not be realized if the policy of its center is changing over time or if the center of «Civilization-magnet» itself is changing.

9 Discussion: legacy of human civilization to hybrid human-AI civilization

As discussed in [20], the properties of AI are largely inherited from human civilization. This is true both for its organization (like human intelligence, it is built on the basis of neural networks [21], [22]) and for the direction of its development. Especially it is true for interaction of large, civilization-scale, general AI and human intelligence.

In particular, future joint human-AI civilizations may inherit two types of human society design: an open society (in which order is maintained mainly through a balance of openly competing social movements, frequently replacing each other in power – this is, in essence, how democracy works) and a closed society (in which order is maintained by closed communities, such as intelligence services, shaping opinions and directions of development of society).

Two types of societies – open and closed – tend to develop two types of AI: open-access and even open-source (controlled by society) AI and closed-access (secret or hidden, out of society-control) AI.

The modeling of evolution of our civilization into hybrid AI-human civilization is becoming important nowadays, with growing AI abilities. The type of this future hybrid civilization is determined by the current organization of society.

The Mind-Matter theory and its Mind-Matter meta-model (able to simulate both human and AI civilizations) can be used for this important direction of research.

References

1. David Walsh. "MWC 2024: Microsoft to open up access to its AI models to allow countries to build own AI economies" Published 26/02/2024/ Updated 27/02/2024. URL: https://www.euronews.com/next/2024/02/26/mwc-2024-microsoft-to-open-up-access-to-its-ai-models-to-allow-countries-to-build-own-ai-e
2. Pascale Davies. "Open source vs closed source AI: What's the difference and why does it matter?" Published on 20/02/2024. Updated 26/02/2024. URL: https://www.euronews.com/next/2024/02/20/open-source-vs-closed-source-ai-whats-the-difference-and-why-does-it-matter
3. Mishchenko, A. "Computer Modeling of the Evolution of Civilization within the Futurological Theory of Mind-matter". *Problems of Artificial Intelligence*. 2021. No 3(22). URL: http://paijournal.guiaidn.ru/download_pai/2021_3/1_Мищенко.pdf
4. Зубенко, В.В. К понятию функции как вычислительной процедуре. *Искусственный интеллект*. 2010. URL: https://jai.in.ua/index.php/архів?paper_num=1012
5. Матвийчук, А.В. К вопросу о принципиальной возможности создания искусственного интеллекта. *Искусственный интеллект*. 2010. № 3. URL: https://jai.in.ua/index.php/архів?paper_num=922
6. Павлыш, В. Н. Дискретное моделирование анизотропной сплошной среды [Текст] / В. Н. Павлыш, А. В. Гром. *Международный рецензируемый научно-теоретический журнал «Проблемы искусственного интеллекта»*. 2023. № 1(28). С. 43–49. http://paijournal.guiaidn.ru/download_pai/2023_1/5_Павлыш_Гром%2043_49.pdf
7. Румянцев, В. В. физических объектов / В. В. Румянцев, С. А. Федоров, А. Е. Рыбалка. *Моделирование физических объектов*. 2023. № 1(28). С. 50–59. URL: http://paijournal.guiaidn.ru/download_pai/2023_1/6_Румянцев_50-59.pdf
8. Настасенко, В.А. Основы концепции определения предельного быстродействия компьютерных систем искусственного интеллекта / Настасенко В.А., Настасенко Е.В. *Искусственный интеллект.* 2008. № 4. URL: https://jai.in.ua/index.php/архів?paper_num=571
9. Алесь Мищенко Цивилизация-лишайник как альтернатива технологической сингулярности. *Инвест-Форсайт*. Январь 2020. URL: https://www.if24.ru/tsivilizatsiya-lishajnik/
10. Алесь Мищенко "Цивилизация-лишайник" и "Цивилизация-магнит": два типа симбиоза человечества и искусственного интеллекта будущего. *Инвест-Форсайт*. Март, 2024. URL: https://www.if24.ru/tsivilizatsiya-lishajnik-i-tsivilizatsiya-magnit/
11. Jessie Yeung. Beijing locks down office building with workers still inside after single Omicron case detected. *CNN*, January 17, 2022. URL: https://edition.cnn.com/2022/01/17/china/beijing-omicron-olympics-mic-intl-hnk/index.html
12. "China locks down entire buildings and even cities as "stealth Omicron" variant fuels record COVID cases" *CBSnews*, March 14, 2022. URL: https://www.cbsnews.com/news/covid-china-record-cases-lockdown-changchun-shenzhen-stealth-omicron-variant
13. "Covid lockdown protests break out in western China after deadly fire". the Guardian. URL: https://www.theguardian.com/world/2022/nov/26/covid-lockdown-protests-break-out-in-western-china-after-deadly-fire
14. Jessie Yeung. "China's lockdown protests: What you need to know". *CNN*. URL: https://edition.cnn.com/2022/11/28/china/china-lockdown-protests-covid-explainer-intl-hnk/index.html
15. "*Ullim"* URL: https://en.wikipedia.org/wiki/Ullim
16. Гарбарчук, В. Методологічні аспекти управління державою як інтелектуальною кібернетичною системою / Гарбарчук, В. *Stuc. Intelekt*. 2010. № 3. URL: https://jai.in.ua/index.php/архів?paper_num=921
17. Шевченко, А.И. Задачи и вопросы экспериментального поиска алгоритмов интеллектуального творческого процесса человека как прототипа машинного интеллекта / Шевченко А.И., Сальников И.С., Сальников Р.И. *Искусственный интеллект*. 2008.№ 3. URL: https://jai.in.ua/index.php/архів?paper_num=470
18. Institutional Dynamics and Organizational Complexity : How Social Rules Have Shaped the Evolution of Human Societies Throughout Human History / Richerson, Peter J., Jenna Bednar, Thomas E. Currie, Sergey Gavrilets, and John Joseph Wallis, eds. *Open Access Book, Cultural Evolution Society*, 2023. – URL:https://institutionaldynamicsbook.culturalevolutionsociety.org/chapters/03_Mathematical_Models_Institutions.pdf

19. Yermolenko, T.V. Application of machine learning in stock market forecasting / T.V. Yermolenko, D.V. Popadin, V.N. Kotenko. *Международный рецензируемый научно-теоретический журнал «Проблемы искусственного интеллекта»*. 2023. № 2(29). URL: http://paijournal.guiaidn.ru/download_pai/2023_2/2.pdf
20. Мищенко, А.В. *Индульгенция людей* [Электронный ресурс] / Мищенко А.В. URL: https://AlesMishchenko.github.io/indulgencia
21. Human-Centered Concept Explanations for Neural Networks / Yeh, C., Kim, B., & Ravikumar, P. 2021. ArXiv, abs/2202.12451
22. Seeliger, K. What comparing deep neural networks can teach us about human vision / Seeliger, K., Hebart, M.N. Nat Mach Intell 6, 122–123 (2024). https://doi.org/10.1038/s42256-024-00789-8

RESUME

A. V. Mishchenko

“Civilization-lichen” and “Civilization-magnet”: modeling symbiosis of humanity and artificial intelligence, within the mind-matter theory

The types of future hybrid human-AI civilization, as function of nowadays AIs and societies.

The types of future coexistence between human society and future general AI depend both on the types of AI being developed today and types of existing societies. The range of possible partnerships, suggested by science fiction, range from elimination or total obedience of humans to happiest life under the care of AI. However, the scientific approaches to simulation of evolution of our civilization into future hybrid human-AI civilization was missing. Similarly, modeling any aspects of possible future hybrid human-AI civilizations was missing as well.

This paper uses Mind-Matter theory for the first simulation(in a multi-player, or, multi-actor system) of evolution into hybrid human-AI civilization, as well as for the first modeling of some aspects of possible human-AI coexistence. This simulation is done using models of 3 or 4 independent AI-actors and 100 independent human-actors, trying to broaden a set of their possible trajectories (successful behaviors) in unknown environment.

The closed and open types of nowadays AI and actuality of the problem.

This paper considers two types of AI (closed-access, controlled by authorities and open-access, controlled by society). In simulation this corresponds to imposing policies on functioning of human-actors and to mimicking many alternatives of behavior, correspondingly.

The question whether the development of strong general AI should be strictly controlled by government or, on the contrary, open-source and, therefore, society-controlled, is becoming increasingly important. For example, at the mobile technology congress in Barcelona in Mars 2024, Brad Smith, president of Microsoft (the main owner of GPT chat and other advanced developments in the field of general AI) announced new «AI Access Principles», according to which, all developments of Microsoft in the field of strong AI will become open source and available for study and use by anyone in any projects[1].

Results of simulation of a multi-player system.

The simulations with two abovementioned types of AI (open and closed) lead to enabling and restrictive AI correspondingly and, finally, to two patterns of humans-AI symbiosis, called «Lichen-civilization» and «Magnet-civilization».

РЕЗЮМЕ

А. В. Мищенко

«Цивилизация-лишайник» и «Цивилизация-магнит»: моделирование симбиоза человечества и искусственного интеллекта, в рамках теории мыслящей материи

Типы будущей гибридной цивилизации человека и ИИ, в зависимости от типов современных ИИ и типов современных обществ.

Возможные типы будущего сосуществования человечества и ИИ зависят от типов разрабатываемого сейчас ИИ и типов современных обществ. Фантастика предлагает широкий спектр возможного будущего - от уничтожения/порабощения человечества до безбедного существования под опекой ИИ. Однако научные подходы к моделированию эволюции нашей цивилизации в будущую гибридную цивилизацию человека и искусственного интеллекта пока отсутствовали. Точно так же отсутствовало и моделирование любых аспектов возможных будущих гибридных цивилизаций человека и искусственного интеллекта.

В данной работе используется теория Мыслящей материи для первого моделирования эволюции нашей цивилизации в гибридную цивилизацию человека и ИИ, а также моделирования некоторых аспектов возможного сосуществования человека и ИИ в системе многих игроков. Эта симуляция выполняется с использованием моделей 3 или 4 независимых ИИ-игроков и 100 моделей независимых людей-игроков, пытающихся расширить набор своих возможных траекторий (успешного поведения) в неизвестной среде.

Закрытый и открытый типы современного ИИ и актуальность проблемы.

В данной статье рассматриваются два типа ИИ (закрытого доступа, контролируемого уполномоченными органами и открытого доступа, контролируемого обществом). В моделировании это соответствует навязыванию правил функционирования людей-игроков и имитации людьми-игроками многочисленных альтернатив поведения.

Вопрос о том, должна ли разработка сильного общего ИИ строго контролироваться государством или, наоборот, иметь открытый исходный код и, следовательно, контролироваться обществом, становится все более важным. Например, на конгрессе мобильных технологий в Барселоне в марте 2024 года, Брэд Смит, президент майкрософта (основного владельца чата GPT и других передовых разработок в области общего ИИ) объявил о новых принципах доступа к ИИ, согласно которым передовые разработки майкрософта в области сильного ИИ, будут выложены в открытый доступ, то есть, станут доступными для изучения и использования кем угодно в каких угодно проектах[1].

Результаты моделирования.

Моделирование эволюции системы многих игроков с открытым и закрытым ИИ приводит к ИИ, расширяющему и ограничивающему людей-игроков, соответственно. Два типа симбиоза человека и ИИ, соответствующие этим типам ИИ, названны «цивилизация-лишайник» и «цивилизация-магнит».

Мищенко Алесь Викторович – кандидат физико-математических наук, доцент, НИИ информатики и автоматики, г. Роканкур, Франция. *Область научных интересов*: искусственный интеллект, моделирование ИИ и ЕИ, компьютерное зрение, машинное обучение, нейронные сети, моделирование и обработка сигналов нейрофизиологии, сознание.

Статья поступила в редакцию 30.03.2024.