Advances in Informatics, Management and Technology in Healthcare
J. Mantas et al. (Eds.)

doi:10.3233/SHTI220783

Artificial Intelligence: On the Way to Doctor's Trust

Michael SHIFRIN[a,1], Anton KHAVTORIN[b] and Gleb DANILOV[a]
[a] *The N.N.Burdenko National Medical Research Centre of Neurosurgery, Moscow, RF*
[b] *Photino LLC, Moscow, RF*

Abstract. The submission is devoted to reflections on the role of trust to modern IT systems, especially based on the AI technologies. Its purpose is to draw the attention of the medical informatics community to the need to achieve trust at all stages of the life cycle of MDSS and other information systems.

Keywords. Artificial intelligence, trust, medical decision support systems, blockchain

1. Introduction

Trust is a value on which all human relationships are based, personal and professional, all relationships between a person and technology, and even the interaction of technical systems. When we board an airplane, our agreement to fly is based on a whole set of "trusts": trust in the aircraft manufacturer and model, airport services, pilot, etc. From all these trusts, our willingness to fly, here and now, follows. The same is true when communicating with doctors. Each act of interaction with a doctor is built both on formal trust (qualification, technique, certificates, documented experience, etc.) and on a deeper, intimate trust of the patient that this particular doctor can help him.

The development of information and communication technologies in the 21st century has led to the formation of a completely new technological environment for communication. In addition, the events of the last two years have radically changed the mass technique of communication and interaction between people and technical systems in almost all areas of human activity. In two years, technologies began to be massively used, which in a normal mode would be introduced within 5-7 years. In order not to lose this unique technological impulse, it is necessary to comprehend, reflect, and, perhaps, in many respects rebuild the very approach to such a fundamental basis of human relationships as trust.

In the current submission, the authors consider the issues of building professional trust in systems built based on artificial intelligence (mainly machine learning methods). There are many works devoted to the problem of the trust in the context of AI systems and closely related topic of their explainability (see [1, 2] for example). Most of them suppose that AI system is developed already and analyze its trustability. Rather, in this presentation, we will 'drill down' and try to identify the trust issue along the path of system development, from problem formulation to implementation and operation.

[1] Corresponding Author Michael Shifrin, e-mail: misha.shifrin@gmail.com

2. On the technological understanding of the term "trust"

It must be said right away that speaking of trust, we in no way affect moral, ethical economical or legal aspects of this concept and will not try to analyze its different definitions. Two participants in the interaction are in a trusting (or confident) relationship if they are completely sure that the interaction process in all its aspects is going the way and only the way it should go in accordance with their ideas and accepted procedures. In other words, each participant may not think about how the other part of the process is performed.

2.1. Technical Example of Ensuring Trust

One of the striking analogs of trust in the traditional IT field is the 7-layer Open Systems Interconnection model (OSI model) [3]. The core feature of this model is that each level of the system is guided by its own internal protocols and does not interfere with the work of neighboring levels, exchanging with them only attribute information related to the transmitted information packet.

Thus, we can talk about the emergence of a "trust stack" or even a "trust graph" that describes the trust relationship between the actors of the business process under consideration.

2.2. A medical example of building trust

Returning to our subject area, let us consider an example in which the issue of trust is resolved systemically and at the global level. This is clinical trials of new drugs. The doctor, prescribing the drug to the patient, must be sure of the following:

- the prescribed drug has passed all the necessary tests
- a specific copy (packaging) of the drug is identical to the one that was tested

In practice, this means that the drug trials were conducted in accordance with the requirements of Good clinical practice (GCP, [4]), and production was organized in strict accordance with the requirements of Good manufacturing practice (GMP, [5]).

3. AI and Trust in healthcare

The special role of trust in the practice of medical informatics is expressively described in the comments to the IMIA Code of Ethics [6].

"in fact the patients also enter into a trust relationship with health informatics professionals (HIPs). To understand why and how this case, it may be useful to recall that the ability of healthcare providers to provide appropriate health care is dependent on data about the patient and on the ability to communicate with other health care professionals as and when this becomes necessary - which, in turn, is dependent on the existence and proper functioning of a technical informatics framework because without it this could not occur. HIPs underwrite all of this through their professional actions… Nevertheless, because HIPs play a pivotal role in the modern health care setting, patients in fact place their trust in the HIPs' ethical and professional actions. The trust relationship is mediated and not direct and may not even be conscious, but it is there."

In the era of introducing AI systems into clinical practice, in which technical systems are entrusted with the task of supporting medical decision-making, achieving trust in such systems becomes one of the central tasks for developers and specialists for implementation. Let us consider that specific chain of "trusts" (trust chain), which arises during the development of medical decision support systems (MDSS).

Its initial level is the formulation of the problem: the search for the solution that the planned system is intended to support. At one time, an integrated approach to this issue was developed at the school of medical informatics by I.M.Gelfand [7]. The idea consists of a sequence of four steps:

- medical goal formulation
- highlighting doctor's decision
- doctor's question formulation
- formal statement of the problem

Let's look at these four steps from the point of view of a doctor who is reading description of a new MDSS, developed and presented accordingly with these principles.

Table 1. A first glance at a new MDSS

Step of development	What do I realise?	I can trust that…
Medical goal formulation	The context of the MDSS development and its KPIs.	…MDSS is intended for real medical problem and it will be possible to evaluate its efficiency.
Doctor's decision	The decision supported by the system.	…MDSS is intended to help doctor to make decision, significant for my concrete patients
Doctor's question	An exact subject of classification or forecasting.	…developers found a question crucial for the decision-making.
Formal statement of the problem	Conditions necessary for the correct operation of the system.	…verification of the decisions proposed by the system can be carried out according to the principles similar to the well-known principles of the GCP, and it was done.

Each of the four is important for building a system that can be implemented in real clinical practice [7]. We especially note the role of the first stage, as setting the context for the operation of the system. This is because the performance of any system can only be adequately assessed in the context of the larger system.

If conclusions of the first three rows of the most right column may be carried out on the basis of medical experience and exploratory statistics, the last one requires special comments. These comments may be found in the Table 2, demonstrating an analogy of developing and validating MDSS and clinical trials.

Table 2. Trust at the implementation and operating stages

Clinical trials phases	Stages of the MDSS development	Found trust
Phase I – first-in-humans	Learning of the system on the training set.	The system works on a special fixed set of data and can be tested on a larger set.
Phase II - to evaluate whether the drug has any biological activity or effect	Test of the system on the control set.	The system is operational on pre-prepared data and can be tested on data obtained during a real clinical process.

Phase III – to assess the effectiveness of the new intervention and, thereby, its value in clinical practice	System operates in real clinical conditions, but without informing medical personnel about results of its work.	The system is operational in real clinical conditions and can be implemented in clinical practice.
Phase IV – drug monitoring trial to assure long-term safety and effectiveness of the drug.	Test of the long-term system efficiency.	The system does not change its basic properties during long-term operation.

As in the case of the Table 1, the last row of this table leads to a new aspect of trust in information systems. It is connected with the self-learning – a hallmark of AI systems.

None of the technical devices and none of the drugs have ever had 'free will' and could not change their properties during use. At the same time, self-learning, is a fundamental property of AI systems. Therefore, developers of AI systems must consider providing another layer of trust on the part of clinicians: confidence that, over time, there will not be serious changes in the properties of the system, that its actions will not become unpredictable. Achieving this degree of confidence in the system should be ensured by a set of rules that are, per se, similar to the rules of GMP, what should be discussed in special paper.

4. Conclusion

Our work is an attempt to demonstrate the omnipresence of trust as the basic condition for the productive interactions of a person and information systems in the context of healthcare. Nevertheless, we cannot help mentioning one more topic – the blockchain technology. This technology allows not only maintain the secure circulation of digital currency, but also allows creating an environment of trusted interaction between participants in distributed business processes. This idea was voiced in [8] in the context of healthcare, and can be applied in the situation discussed above as a tool to ensure trust between the participants in the process of implementation and use of MDSS.

References

[1] Stanton B, Jensen T. Trust and Artificial Intelligence. Available at: https://nvlpubs.nist.gov/nistpubs/ir/2021/nist.ir.8332-draft.pdf

[2] Amann J, Blasimme A, Vayena E, et al. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. BMC Med Inform Decis Mak. 2020;20:310. doi.org/10.1186/s12911-020-01332-6.

[3] Available at: https://en.wikipedia.org/wiki/OSI_model.

[4] Available at: https://www.ema.europa.eu/en/human-regulatory/research-development/compliance/good-clinical-practice

[5] Available at: https://www.ema.europa.eu/en/human-regulatory/research-development/compliance/good-manufacturing-practice

[6] Ethics for Health Informatics Professionals. Available at: https://imia-medinfo.org/wp/wp-content/uploads/2015/07/Handbook-for-revised-Code-of-Ethics.pdf

[7] Aleckseevski A, Gelfand I, Izvekova M, Shifrin M, On the role of formal methods in clinical medicine: from purpose to problem setting (Rus). In: Medicine and Cybernetics, Moscow, "Science" 1997. p. 5-34.

[8] Shifrin M., Khavtorin A., Stepurin V. and Zingerman B., Blockchain as a Process Control Tool for Healthcare, Health Informatics Vision: From Data via Information to Knowledge, J. Mantas et al. (Eds.) IOS Press, 2019 p. 172-175.