Abstract. In presented study we discuss the operation of International Center for Neutron Research based on the PIK high-flux research reactor, and possible adaptation of successful foreign practices for this facility. We consider the phenomena of “digital twins” vs “digital trace”. The first one is a model, possessing the main features of a real scientific facility, while “digital trace” deals with a lake of raw data and scientific outcomes are produced by multiple data processing with various methods. Both approaches lead to dynamic digital infrastructure development; it becomes a prerequisite for the functioning of any megascience facility. We argue this to be the next stage in the scientific facilities development: in nearest future digital twins and digital trace together would eliminate the uniqueness of the megascience, and would lead to the formation of a new supra-infrastructure. This object should be evidently based on taking into account the interaction of man, society, science and technology.

Keywords: Megascience · Digital twins · Digital trace · International scientific cooperation · Infrastructure

1 Introduction

Modern world presentation (and current ideas about the Universe) is largely based on the results of work obtained on unique scientific facilities. These installations are the only possible nowadays scientific research tool capable of solving urgent problems and responding to global challenges of our time, providing an opportunity for interdisciplinary research and projects in various areas [1, 2]. Projects of such a type are defined as a “supranational” organization with “independent representations” [3] or as organizational and managerial innovation [4].

Due to their complexity and the scale of the tasks to be solved, their successfully functioning is possible only in the format of international consortia. In practice, the point is that a characteristic feature of megascience is the fact that the cost of their creation and maintenance is beyond the capabilities of individual countries and/or entities [5]. Though to be constructed by one country solely (for example, ISSI-4 facility in Russia), their practical use upon completing the project is only possible by a number of scientific teams, and this approach has been already included at planning stage (access for international research teams, sale of time and/or data for external users, etc.).
Thus, a megascience class project should be treated as a business-like project. It aims at obtaining new (unique) knowledge and technologies, the last to be later commercialized [4, 6] and adapted in various sectors of the real economy [7, 8]. For instance, super-bright X-ray beams are used for studying living complex systems (proteins and human brain cells). Megascience class facilities form differs from a set of infrastructural objects localized in one territory (single-sited) to a kind of distributed facilities (united under one “umbrella” project).

This gives rise to a whole range of features of their activities: from the philosophical questions and legal issues to maintenance of an appropriate infrastructure. One of the problems that gained additional relevance in 2020 due to the coronavirus pandemic was the organization of work on unique scientific facilities in remote access mode, and the development of appropriate changes to existing business processes. The possibility of remote scientific work on installations of this class has existed since the appearance of the technical possibility of realizing such access, and until recently it represented a specific highly specialized product (reducing costs and expanding access to equipment). In 2020, the social aspect was added: it could serve as health-protecting tool (so called scientific distancing).

The entire life cycle of megascience facilities can be divided into 2 related operation groups. First, the “engineering” task is the solution of applied problems arising during the creation, operation and modernization of the installation (the actual set of scientific equipment and its maintenance). The second, “scientific” task is the solution of scientific and practical challenges, setting up an experiment and analyzing the data obtained. While not regarding the “technical” side of organizing work on megascience-class facilities, we note the necessity to introduce a number of changes and/or corrections to some business processes in order to take into account the specifics of remote access in title documents.

In this work (carried out with the support of the RFBR grant No. 18-29-15015), we discuss some practices from different megascience facilities that can be implemented at International Center for Neutron Research on the basis of the high-flux research reactor PIK (ICNR PIK). ICNR PIK is now being constructed in the outskirts of St. Petersburg, Russia. It is a continuous flow type reactor and is intended for research in the field of condensed matter physics, nuclear physics and the physics of weak interactions, structural and radiation biology and biophysics, radiation physics and chemistry.

Its parameters are supposed to surpass the only analog in the world - the HFR reactor at the European Center for Neutron Research – International Institute Laue-Langevin. The facility will have 50 experimental stations fed by 10 horizontal, 6 inclined and 6 vertical experimental channels. It will possess 3 cold and 1 hot neutron sources. The volumetric energy release in the core will count up to 6,6 MW/l.

In our study we considered mostly possible cooperation with Republic of Belarus, BRICS countries, and European Union. This list, however, is not inclusive, and research infrastructure development projects are open for any participants. The most important thing here is the intention of the society and authorities to develop innovative infrastructure, and to participate in currently ongoing projects and/or initiate ones. Among Latin American countries, Mexico has the most consistent position on this issue. This country identifies several areas of cooperation: North America (USA, Canada), South America (Argentina, Brazil, Chile and Colombia), Europe (Germany,
Spain, France and the UK), and Asia (China, South Korea, India, Israel and Japan). Mexico has concluded about 100 bilateral and multilateral international agreements on scientific and technical cooperation with more than 25 countries, the EU and two multilateral organizations. In Brazil, interaction with foreign partners takes place within the framework of the Science without Borders (SwB) program, according to which Brazilian scientists have the opportunity to conduct research in the field of natural sciences and technology in leading foreign universities. Several directions can be mentioned: nanotechnology (e.g., CNPEM - Energy, LNNANO - Nanotechnology), fundamental physics (SIRIUS - The Brazilian Synchrotron Light Laboratory, LNLS). All other Latin American countries do not have specially formulated attitude towards cooperation in megascience field, but we expect this issue to draw much more attention in future.

2 Remote Access Mode for Megascience

It can be seen that remote access is a necessary element of many mega-science facilities, which has firmly entered into scientific practice. Our previous study revealed that remote access mode for megascience requires following adaptation of business process [9, 10].

First, it is necessary to reduce to a minimum the need for the physical presence of the customer at the facility where the unique scientific installation is located. For example, in the international project Borexino, there is an additional independent data collection system based on fast waveform digitizers. The complex allows to collect data remotely without the need for scientific teams to travel to the location of the detector.

Secondly, it is necessary to develop the appropriate digital infrastructure: it is required both the creation of new elements of the digital infrastructure (data centers, processing algorithms, etc.), and the integration of megascience installations into the existing e-Infrastructure facilities.

Thirdly, it is necessary to provide legal and methodological support for the operation of unique scientific facilities in remote access mode.

Fourth, it is necessary to provide engineering personnel to solve the related tasks of maintaining and ensuring the remote access mode.

The most interesting and challenging area of work that combines the fulfillment of these requirements is the development of digital infrastructure. The development of digital infrastructure gave rise to the new phenomenon – e-Infrastructure. This idea was first proposed and evolved in the EU, when digitalization of the scientific research was suggested as a method of organizing modern scientific research. The first step should be the unification of science, data collection systems and access to them. e-Infrastructure is thought to be the tool for implementing EU policies in science, when the achievements of the Internet, grid systems, cloud computing and databases are assembled in a new infrastructure. Europe started with the establishing of the open scientific portal EOSC (European Open Science Cloud), launched in 2016. Based on the EOSC work experience, the Go FAIR initiative is being prepared to put ideas and proposals related to digital science into practice [11].
To solve the applied problems of processing scientific information in the EU, a number of specialized data processing projects are planned: GEANT (management of scientific and educational network projects), EGI - Advanced computing for research (providing calculation options for CERN, EMBL projects), PRACE (providing computing power, 465 projects at the moment). In addition to the above, in order to codify and standardize digital infrastructure, within the framework of the Horizon 2020 program, the European Union launched the e-Standards project [12].

It is expected that the first consumers of e-Infrastructure will be representatives of the natural sciences, however, the greatest impact and the most significant results will be manifested in the field of humanitarian knowledge, which forces participants in the process to develop appropriate assessment methods and approaches today, simultaneously solving legal issues.

In this paper we present the dichotomy: phenomena of “digital twins” and “digital trace” (by analogy with digital footprint) from megascience installations. The first is a “digital reflection” (model), which carries the main features of a real scientific phenomenon. They can be used to work out the main experiments conducted on considered facility, and their interface allows to work on them in a remote mode and for unskilled users. This democratization of technology would provide easy access (including for non-specialists) to knowledge in technology and business without lengthy or expensive training. Called “citizen access”, this policy is already used in application development, in data and analytics systems, and in solving design and knowledge problems [13]. The implementation of such a policy leads to a gamification of science. For example, it is discussed that the world of “megascience” can be transferred to the Minecraft game shell [10].

For the case of a digital trace, there is a maximum collection of raw information and its subsequent multiple processing by new methods (i.e., this, in fact, it is a digital data lake). This raw data are subject to repeated processing at different points in time by different methods and algorithms (i.e., a scientific discovery is possible on old data that have been processed with new methods). This is similar to extracting information about people from their digital footprint. Here we face an intriguing challenge: if the digital footprint is now right or duty? For the case of human beings, recent trends show the obligation of people to produce digital footprint that can be later analyzed and commercialized. With scientific data we notice the similar trend (the more data - the better), and it is accompanied with new data handling algorithms development.

Both approaches are widely used in CERN. For example, in 1998 the MONARC project (Models of Networked Analysis at Regional Centers for LHC Experiments) was launched. It yielded in the concept of a hierarchy of data processing, modeling and analysis centers. There are currently 4 tiers of processing centers. At the bottom (zero) level there is the Tier-0 CERN Computing Center that deals with the primary reconstruction of events, calibration, permanent storage and archiving of the complete set of “raw” and simulated data. Then there are Tier-1 (13 centers), Tier-2 (about 170 centers), and, finally, Tier-3 (about 50 centers), represented by university clusters, or centers that provide resources on a voluntary basis, where physical data analysis is carried out [14, 15]. A similar policy for working with data was implemented in the case of the global neutrino network (GNN) [16]. This network can be interpreted as a distributed scientific infrastructure, its elements being separate facilities and
collaborations (for example, the Ice Cube collaboration with 47 organizations from 12 countries of Europe, America and Asia or the Dubna deep-sea neutrino telescope of multi-megaton scale).

The results of the work of the digital twin (in accordance with the FAIR principle actively implemented in the EU) have the same “rights” as the real installation. For example, ATLAS (CERN) uses data obtained by computer generated events (by Monte-Carlo method [13]) as an auxiliary tool during the “real” experiment.

3 Current Tendency and Conclusion

The described process fits into the general tendency in the development of scientific research systems, and the coronavirus pandemic acted only as a catalyst for the already existing trends. So, since the 2010s, the e-Science system is being formed and e-Infrastructure is being implemented [17]. At the same time, the digital infrastructure, which is an integral part of the e-Infrastructure, does not duplicate or replace the “physical” one. The idea of “digital twins” described in this paper also does not imply a complete replacement of “real” attitudes, but acts as a tool to support scientific research.

This approach leads to the active development of digital infrastructure, and its presence becomes a prerequisite for the functioning of any megascience facility. According to a number of experts, this process would lead to a universal withdrawal into the digital space. We argue, however, that this should be treated as the next stage in the development of unique scientific installations. The juxtaposition of the digital footprint and the digital twin is seeming: both of them are an integral feature of modern megascience facilities.

The next step would be to standardize the digital world, and the products it generates and/or possesses. From the point of view of scientific research, mega-science installations, translated into the format of remote access and acting in the form of digital twins or reproduced on their digital footprint, will lose their uniqueness and become a “common place” (a kind of a niche product). At the same time, a high degree of interconnectedness and inclusion would lead to the formation of a new supra-infrastructure.

The convergence of science and technology will bring together scientists from different countries and will facilitate the coordination of research and development aimed at overcoming global challenges, including the development of methods to prevent the spread of pandemics based on genetic research.

The format and principles of the new scientific digital infrastructure are not yet clear, but there is a consensus that scientific and technological progress is impossible without taking into account the interaction of man, society, science and technology. Therefore we could expect the adoption of a number of regulatory acts similar to the European Union General Data Protection Regulation [18].

Acknowledgments. This work was supported by RFBR grant № 18–29-15015.
References

1. Nurakhov, N.: The basic processes of creating a “megascience” project. In: International Conference on Integrated Science, ICIS 2019: Integrated Science in Digital Age, Batumi, Georgia, pp. 329–339 (2019)
2. Hallonsten, O.: History and politics. In: Big Science Transformed, pp. 43–98, Palgrave Macmillan, Cham (2016)
3. Crease, R.P., Martin, J.D., Pesic, P.: Megascience. Phys. Perspect. 18, 355–356 (2016)
4. Karlik, A.E., Platonov, V.V.: Conceptual foundations of the study of megascience as an organizational and management innovation. Innovation 10(228), 11–16 (2017)
5. Fotakis, C.: FP7 Interim Evaluation, Analyses of FP7 supported Research Infrastructures initiatives in the context of the European Research Area, Final Report 12 November 2010 (2010)
6. Lami, S.: Challenges and new requirements for international collaborations. Sci. Diplomacy 6(2) (2017). https://www.scencediplomacy.org/article/2017/mega-science-collaborations
7. Balyakin, A.A., Mun, D.V.: Formation of an open science system in the European Union Information and Innovations. In: Proceedings of the Conference “Sciencemetry and Bibliometry”, pp. 33–37 (2017)
8. Gartner’s Top 10 Strategic Technology Trends for 2020. https://www.pcmag.com/news/gartners-top-10-strategic-technology-trends-for-2020. Accessed 21 June 2020
9. Nurakhov, N.N.: Integrity of Innovation Management and INSO Inventory, p. 156. MST, Moscow (2010)
10. Nurbina, M.V., Nurakhov, N.N., Balyakin, A.A., Tsvetus, N.Yu.: Mega science projects for business. In: Ahram, T., et al. (eds.) Human Interaction, Emerging Technologies and Future Applications III Proceedings of the 3rd International Conference on Human Interaction and Emerging Technologies: Future Applications (IHET 2020), Paris, France, pp. 488–492 (2020)
11. Dutch Techcentre for Life sciences. https://www.dtls.nl/fair-data/go-fair/. Accessed 10 July 2020
12. Report on the EU funded eStandards project. https://ec.europa.eu/digital-single-market/en/news/report-eu-funded-estandards-project. Accessed 14 July 2020
13. Ay, C., et al.: Monte carlo generators in ATLAS software. In: Journal of Physics: Conference Series, CHEP 2009, Prague, Czech Republic, vol. 219 (2010). https://iopscience.iop.org/article/10.1088/1742-6596/219/3/032001
14. Grigorieva, M., Golosova, M., Ryabinkin, E., Klimentov, A.: Exabyte repository of scientific data. In: Open Data Systems. DBMS, Moscow (2015). https://www.osp.ru/os/2015/04/13047963
15. Klimentov, A., Kiryanov, A., Zarochentsev, A.: Russian lake of scientific. In: Open Data Systems. DBMS, Moscow (2018). https://www.osp.ru/os/2018/04/13054563
16. Global Neutrino Network (GNN). https://www.globalneutrinonetwork.org/. Accessed 30 June 2020
17. Zhulego, V.G., Balyakin, A.A., Nurbina, M.V., Taranenko, S.B.: Digitalization of society: new challenges in the social sphere. Bull. Altai Acad. Econ. Law (9–2), 36–43 (2019)
18. The EU General Data Protection Regulation (GDPR). https://gdpr-info.eu. Accessed 18 Nov 2019. Accessed 17 Mar 2020

Digital Twins vs Digital Trace in Megascience Projects 539