NATIONAL RESEARCH UNIVERSITY
HIGHER SCHOOL OF ECONOMICS

Meissner Dirk,Narkhova Anastasiia,
Plekhanov Dmitry

THE MEANING OF DIGITALIZATION FOR RESEARCH SKILLS: CHALLENGES FOR STI POLICY

BASIC RESEARCH PROGRAM
WORKING PAPERS

SERIES: SCIENCE, TECHNOLOGY AND INNOVATION

WP BRP 69/STI/2016

Meissner Dirk[1], NarkhovaAnastasiia[2], Plekhanov Dmitry[3]

THE MEANING OF DIGITALIZATION FOR RESEARCH SKILLS: CHALLENGES FOR STI POLICY

This paper gives an overview of the impact of the current digital evolution of S&T systems, in particular on STI policies and research skills development. Within this context digitalization demonstrates strong impact on human capital and tangible assets (infrastructures and information processing tools etc.) by means of enhancing sources of new knowledge. Hence, challenges arise on the useful and proper utilization of numerous opportunities which emerged from digitalization. What are the new challenges for S&T systems caused by digitalization and what mechanisms could be implemented for system adaptation to them? What are the new research skills that are a high priority to acquire both by individuals and system as a whole? The paper argues that STI policy thus far hasn't demonstrated sufficient responses to the changing requirements on researcher skills but remains at the infrastructural discussion.

JEL ClassificationO31, O38,

Keywords: Open access, open science, researcher skills, digitalization, STI policy

---

[1] National Research University Higher School of Economics, Institute for Statistical Studies and Economics of Knowledge, dmeissner@hse.ru

[2] National Research University Higher School of Economics, Institute for Statistical Studies and Economics of Knowledge, anarkhova@hse.ru

[3] National Research University Higher School of Economics, Institute for Statistical Studies and Economics of Knowledge,dplekhanov@hse.ru

Linking research: The Knowledge Triangle Perspective

The core mission of science of broadening boundaries of mankind’s knowledge and intellectual capacities did not change through centuries. Nowadays however science and knowledge creation process can significantly benefit from new advanced and ingenious digital tools for information gathering, analysis, storage etc.New digital era has a potential to translate research results into business practices in a better way eventually reaching a higher degree of commercialization and improving decision-making processes. The phenomenon of digitalization reveals itself through ubiquitous growth of information availability and production. Some estimations showed that in 2015 the amount of data produced worldwide was about 8 zettabytes[4] (Gantz & Reinsel, 2011; Meeker, 2016), and this number is expected to grow exponentially due to rise of wireless networks and connected devices. By year 2020 the amount of data worldwide is likely to be more than 16 zettabytes and can potentially rich 40 zettabytes (5 times growth!) (Turneretal., 2016; BVD, 2016).Todaydataarecomingfrom a multitude of sourceswhetheritisinternet, scientificdatabases, digitalrepositories, onlinefinancialtransactions, governmentdatabases, socialmediaandnetworks, surveysetc. For example “SKA’s [Square Kilometre ArrayTelescope] firstphase of developmentwillseetheequivalent of 3 terabytestransmittedeverysecondtothecentralprocessingcomputer. Sothat’sBigData.” (BigData - SKA Telescope, 2016). Thistrend interweavesallaspects of humanendeavoropeningnewinsights in science, datastructuringandmanagement, privacyandethics.

Thegrowingcomplexity of informationhasalsoaffectedallbusinessareasfrom e-commercetogame-applications. Themosthypedappin 2016 – ‘PokémonGo’ – wasinstalled morethan 100 milliontimesin a monthsinceitsofficiallaunchin July. Thisstatisticdoesnotcover “cracked” versionsdownloads. Thismeansthattheapplicationdevelopercompanyhastostoreinformationaboutalmost 100 millionaccountstoanalyzeconsumers’ behavior: thetimeclientsspendin the game, whatproductstheybuyetc. This requires excellent data analytical skills and strong business intelligence. Now we can say without any doubts that our world “...is ready for augmented reality”(Eordogh, 2016), whichmeans the emergence of a wave of new process changes, particularly within the way we do business and science. Not only companies, but other stakeholders (i.e. research centres, institutions, universities etc.) nowadays realize the need in data management, analysis and interpretation and widely look for experienced specialists able to integrate and extract value from data.

Simultaneously with business environment scientificprocesses are either undergoing significantchanges. From digital technologies they receive new opportunities for collaboration, dissemination of research results, increasing public engagement in science and higher quality of research evaluation schemes (Carayannis et al 2016, Sarpong et al 2016).Data and information are by tradition and nature the core of scientific, namely research, activities. It’s common practice that research begins with inspiration and ideas of phenomena to address leading to building first concepts and models to explain these. Next researcher formulae hypothesis which are tested through experiments finally leading to refined models and concepts. These principle research steps vary in shape and intensity between the science and research fields and issues dealt with naturally. Especially the form of experiments and data and information collection varies accordingly. However with increasing complexity of the basic challenge or problem addressed by research the amount or information and data required increases which call for improved data and information analysis algorithms.

---

[4] A unit of information, which is equal to 1024^7 bytes.

We consider the Knowledge Triangle (KT) model is useful for further explanation since it forms the stage for research-education and research-innovation linkages. We need to understand what involved stakeholders think regarding the changes of those linkages under digitalization and highlight recent developments of tools and instruments for making the KT systems sustainable. The special feature of the KT is that all stakeholders (government, business, universities and nonprofit organizations) are somehow involved into knowledge production and to a different degree perform research, innovation and education activitiesMeissner 2014, 2015, Meissner & Shmatko 2016).

Universities. It has been proven already that HEIs are quite heterogeneous and follow different development paths and perform various roles in regional and national economies, still their main task is to contribute with highly skilled graduates and competitive research outputs. HEIs usually act under strong pressure from government authorities trying to adopt the newest practices in education and fulfill multiple quality and efficiency evaluation requirements. Facing the trend of digitalization HEIs adopt new technologies very intensively: online courses, digital libraries, repositories, learning management systems, various software and equipment for experiments, etc. HEIs require digital infrastructures to store and disseminate knowledge, manage workload of researchers and to prepare information for investors, grant applications and other financial activities. HEIs are also a major consumer of off-the-shelf digital instruments produced by private sector.

Governmentinstitutions are one of the major players in the field of digital infrastructuresfor research and innovation comprising data centres, networks and digital government services (European Commission, 2014). The interest for STI policy stems from the need to increase economic growth through commercialization of research results and to optimize spending on science and technology. In several countries state-run innovation agencies e.g. KISTEP in South Korea, NISTEP in Japan directly participated in the establishment of e-infrastructures for government. In other countries national governments chose the methods of indirect support through financing projects of universities and not-for-profit organizations in the field of e-infrastructures (e.g. Isidore, Research Data Storage Initiative (RDSI)) or through creating environmental frameworks auspicious for development of digital infrastructures (legislation on the use of Big Data analytics).

Another group of stakeholders is represented by ***not-for-profit organizations***. Not-for-profit organizations are either established by university or government-affiliated entities or by the members of the scientific community. The main rationale behind the establishment of such organizations is to facilitate scientific research by designing digital tools for managing workflows, research data stewardship, scientific communication and collaboration. Several not-for-profit organizations provide free digital solutions for storage and dissemination of research information (e.g. Dryad, DuraSpace). The free digital products that they design considerably decrease the expenditures of universities and Public Research Institutions (PRIs) on research information management and allow researchers to invest more time to research tasks than to administrative duties. Certain entities were established or indirectly supported by national governments to facilitate creation of national digital platforms for research information management by designing the guidelines, standards for interoperability and disambiguation of data. One of the most important not-for-profit organizations engaged in this field are EuroCRIS and Consortia Advancing Standards in Research Administration on Information (CASRAI). EuroCRIS, a pan-European not-for profit organization, promotes cooperation between stakeholders in the field of research data stewardship and supervises the development of the Common European Research Information Format (CERIF), an international standard data model. CASRAI is an international not-for-profit organization engaged in the elaboration of unified terminology for the semantics and provision of interoperability between national standards. Such not-for-profitorganizations as Academia.edu, ResearchGate greatly contributed to the dissemination of scientific publications, creation of new channels for scientific communication and giving a new boost to the development of open peer review (Kintisch, 2014). Although the status of these companies is not-for-profit there are high chances that in the future they would extract revenues from analysing data generated by researchers. That is why certain scientists chastise these platforms for being parasitic on public educational infrastructure and for being a threat to Open Access, as they may narrow down the choice of digital instruments for publication and put virtual barriers for knowledge dissemination (Hall, 2015) We can argue that in the field of digitalization there is a fine line between not-for-profit organization and for-profit organization. Although the initial motive of not-for-profit organizations is to provide services for research information management free of charge, there is a high probability that in the future they will commercialise their services attracted by lucrative commercial opportunities.

Business stakeholders usually pay their attention to the development of technical solutions for digital infrastructures for STI policy is a very promising market. There is a variety of private companies that are actively engaged in the providing services for e-infrastructure. The most comprehensive digital platforms for research information management are Pure, Symplectic Elements, and Converis which are equipped with analytical tools enabling them to track, to analyse, and to reuse research results. Some other private companies were established under the umbrella of funding agencies or universities and mainly work with national governments on design of technical solutions for digital infrastructures (e.g. Data61 (Australia), ResearchFish (the UK)).

Obviously the stakeholder group interests are very divers but all put reasonable pressure especially on the knowledge producers, namely universities.

Recent developments on research skills

Global view on digitalization

Despite the numerous advantages of digital technologies, for the time being their full potential has not been fully exploited yet. There are technological and organizational challenges, as well as problems associated with human resources that impede deeper and more thorough application of digital technologies. There are *technological challenges*, which are mainly represented by standardization, disambiguation and interoperability problems.The use of different standards in establishing systems for data collection and analysis leads to fragmentation of efforts and, consequently, to lower efficiency. There are several international initiatives that targeted the provision of standards in research information management (e.g. CERIF, Dublin Core). They provide common grounds for collaboration on national or multinational level and can be used as ready-to-made solutions for designing of new digital infrastructures for STI policy needs. The existing standards need to be improved and its dissemination needs to be supported through national policies. There is a concern that the large number of standards will not contribute to better data governance, but, on the contrary, will result in greater fragmentation and repetitiveness of initiatives.

Digitalization and Open Science

Recent developments in digital technologies have expanded capabilities of e-infrastructures, data management platforms for scientific research and innovation. Through cloud technologies it becomes possible to store large amounts of research information, through scraping software and APIs[5]to connect fragmented pieces of data. By using semantic technologies we may efficiently disambiguate data and provide typologies of received information. As some studies pointout, data mining skills (incl. semantic text analysis) are in high value for all types of researchers: developers, creative-minded etc. (Hayes, 2016).E-infrastructures provide analytics services that have potential to contribute to better connections of inputs, outputs and outcomes of investments into scientific research and may track down the trajectories of scientific and technological developments. The potential area of e-infrastructure's application is formulation of STI policy which currently in predominant number of cases lacks evidence-based approach. Apart from policy needs, data analysis performed by e-infrastructures may be geared towards business needs. The data collected and analysed by e-infrastructures can be used for designing guidelines on R&D investments and analysing developments in frontier technologies.

The Open Science movement is targeted to making the scientific research free and available for the whole society (Budapest Declaration, 2002; Berlin Declaration, 2003). Free access to publications and other types of research results has a potential to increase the impact of scientific research to the benefit of research organizations, individual researchers and commercial companies. Nowadays, the research results are unequally distributed among different countries and institutions which lead to the poor quality of research and deprive the development of less well-established research organizations. Open Science, on the contrary, may provide less developed countries with a chance to exploit the findings of the high-calibre research (Lionelli&Prainsack, 2016). Apart from that, leaving out citizens from the participation in scientific research results in the public distrust of science and may cause backlashes against spending tax-payers money on science, since the large parts of the society are not aware of the social and economic benefits that research may bring about. Another social dimension of open science is ongoing trend of citizen science which may expand the capabilities of the scientific community in collecting and processing data, as well as in finding alternative ways of funding by close cooperation among the public, research organizations and governments.

There is another area to which Open Science may sufficiently contribute. Disclosure of research information and metadata if it is made in a right way has a potential to spur up the development of a new system of research impact evaluation. Traditional approaches of evaluating research results have many methodological flaws and are incapable of linking together research outputs, inputs, and outcomes. New metrics based on digital technologies powered by Open Science movement may sufficiently increase the quality of research funding decisions and may improve the system of incentives for researchers.

The origins of Open Science may be traced back to the Renaissance, but modern outlook of this movement was formed only recently (David, 2004). To advance further development of Open Science initiative it is required to provide legislative framework capable of solving privacy issues and intellectual property conflicts. Another serious challenge is represented by technological dimension: the development of digital tools is essential to manage research information and support of the further implementation of Open Science initiative.The advancements that have been made in the field of digital technologies are contributing to making science more transparent and inclusive. The impact of digitalization on science spans through various dimensions. It changes the way of how scientists work, communicate and disseminate the results of their research. It opens up new opportunities for the private sector by providing valuable information on research facilities and research projects (e.g. a British platform Konfer) or laying down the guidelines on R&D investments (e.g. a South Korean platform K2Base). Digital technologies

---

[5] Application Programming Interface

provide efficient ways for storing and processing information which leads to lower costs, more efficient supervision of internal activities and higher propensity to innovate. Governance of innovation and STI policy are also changing under the influence of digitalization: more detailed data on innovation activities is becoming available which contributes to exploring new areas which STI policy measures should address. Disclosure of research information spurred by Open Science movement gives a chance to use digital technologies for creation of new indicators and new statistical methods for research evaluation which may be used for assessing which policy measures shall be preserved and which discontinued.

Despite of the benefits that digitalization and open access can bring to the society, for the time being their full potential has not been fully exploited yet. There are technological and organizational challenges, as well as problems associated with human resources that impede deeper and more thorough application of digital technologies. Technological challenges are mainly represented by standardization, disambiguation and interoperability problems.The use of different standards in establishing systems for data collection and analysis leads to fragmentation of efforts and, consequently, to lower efficiency. There are several initiatives that targeted to provision of standards in research information management (e.g. CERIF, Dublin Core). They provide common grounds for collaboration on national or multinational level and can be used as ready-to-made solutions for designing of new digital infrastructures for STI policy needs. The existing standards need to be improved and its dissemination needs to be supported through national policies. There is a concern that the large number of standards will not contribute to better data governance, but, on the contrary, will result in greater fragmentation and repetitiveness of initiatives.

Research information is produced in different social and political contexts depending on a national or even on an institutional level, therefore the different concepts may represent the same things, or on the contrary, the same concept may stand for different subjects (Bosnjak&Podgorelec, 2016; Vancauwenbergh, 2016). Research information itself may be error-prone. That is why the disambiguation of data is an important milestone in ensuring high quality of research information analysis.

In order to provide opportunities for data sharing, research cooperation, reuse of research information, and verification of research results it is required to elaborate digital tools for interoperability of research information. Making data interoperable will sufficiently ease application procedures for research grants, minimise administrative burden on researchers and optimise the evaluation of impact of public and private investments into scientific research. There are certain solutions already at place (e.g. unique identifiers, repository plug-ins). However, their dissemination and therefore the impact of interoperability solutions are rather limited due to organizational and environmental imperfections.

While the technological problems seem to be solvable in the long run, the environmental challenges are far more complex and will require more sophisticated solutions. The developments in a field of digitalization of STI policy are shaped by several stakeholders: national governments, the private sector, universities, not-for-profit organizations, libraries, archives, and funders. The digitalization of STI policy is based on the interconnections among those actors. In order to understand the digitalization trajectory, we need to take into account political, societal, economic, and legislative contexts in which the stakeholders run their activities. At the policy level the major forces are the Open Science and Open Government initiatives which are targeted to the disclosure of information as well as to the inclusion of large parts of society to the benefits of free information exchange(OECD, 2015).

From economic perspective the major challenges for digitalization of science are represented by financial constraints. The costs of data collection and processing are very high. The establishment of e-infrastructure as well training of data scientists is associated with large financial investments which can be insurmountable for small and medium scale organizations. Therefore, governments shall guarantee a sufficient financial support to research organizations that cannot finance on their own the exploitation of digital technologies for research information management. Apart from that, it is advisable for block and project funding schemes to have several percent of their total amount designated for research information stewardship (European Commission, 2016). If this measure is taken, the disambiguation of research information on the latter stages of data analysis will require less time and fewer resources.

In the time of intensive international competition in science and technology, many countries implemented so-called Research Excellence Initiatives aimed at the support of limited number of leading research organizations and at the promotion of commercialization of research results. The decisions of evaluation comities to grant the status of centre of excellence to a particular institution can be under the great criticism. In order to identify the best practices that bring large economic impact and positive influence on workforce, and to ensure the required level of transparency in funding decisions it is required to deliver evidence based on data

As for legislative framework, some serious work has to be done in order to advance the application of digital tools for STI policy needs. There is a necessity to elaborate laws that would guarantee the extraction of data on researchers for policy monitoring and in the same time would ensure that the privacy of individuals is secured. Apart from that, the solid legislative framework should be established for managing of public sector data, as it may be an important asset for social and health studies on one hand and it may substantially expand the functional capabilities of e-infrastructures on another hand by using information of government databases for identification of socio-economic impact of research.

If technological, economical, and legislative problems of digitalization of STI can be solved in the long term despite of their complexity, the social problems are more challengeable and will require more time and resources. First of all, digital technologies are replacing traditional research methods, therefore, the set of the core competences that scientists should have is also changing. However, until now the level of digital literacy among the members of the scientific community remains insufficiently low. For example, the skills for computer modelling and simulation are still not very common. This should be addressed by policy measures. The efforts should be made in providing training programs tailored for the needs of digital literacy. The system of incentives for researchers should be also changed in order to make research activities stand in a line with the requirements of data interoperability and data disambiguation. For example, the unique identifiers like ORCID ID can considerably improve the quality of research information and the efficiency of research data and metadata processing. However, scientists are required to apply for these indicators themselves, therefore research organizations and governments need to design the system of direct and indirect support for implementation of this policy. Another problem that should be mentioned here is a huge shortage of data scientists which makes the further development of digital infrastructures for STI policy even more difficult (European Commission, 2016).

Generating new knowledge for regardless has been viewed one of the main missions of universities and research institutes. Whereas the amount of knowledge increased substantially over time the missions of universities and research institutes have changed considerably and numerous different types of these institutions have emerged. Namely during the last decade brought the emergence and widespread of the open innovation paradigm which lead to the perception that knowledge desirable for innovations is increasingly accumulated from different sources including universities and research institutes.

Accordingly these institutions began to change their models of knowledge generation and included the knowledge diffusion dimension to their missions and activities. Knowledge diffusion itself isn't a concept or approach but a rather broad term which is frequently understood as 'technology and knowledge transfer' or one of the two. The diffusion (or transfer) of knowledge (which is the main feature of technology) can many different shapes and institutions use a wide range of channels for this. Overall diffusion and transfer is characterized by a sender and a receiver which are linked together differently. In general there are the knowledge generating institutions, among them universities and research institutes, industry and the public sector which sets the framework conditions under which the linkages between the actors evolve and are used. This phenomenon was described as 'Knowledge Triangle' (KT). Within the KT the actors are searching for their optimal role and align their missions and activities accordingly. Currently we can observe that knowledge generating institutions are confronted with an increased request by the public sector and industry to make their knowledge more accessible using more sophisticated channels than in the previous times. This includes education and communication of research results to a broader audience and more near time. The digitalization movement experienced over the last decade provides additional possibilities for doing that but it also imposes new requirements and challenges which are dealt with in the succeeding sections.

Digitalization of Science and Open Science

Impact on research

Advancedutilization of digital instruments for research purposes and data manipulationsis usually referred as 'eScience' or 'digital science' (Bohle, 2013; Medeiros & Katz, 2016). This concept underlies a positive picture of the high-end 'digital organization of science', that incorporates above mentioned e-infrastructures, services and communications enabled by ICTs, open access and open data (Open Science), digital repositories, data storages and any other types of digital instruments.There is still no consensus or commonly accepted definition of digital science, because its interpretation often strongly depends on a discussed topic, whether it is a scientific problem, business issue or innovation intensity (Borrás & Edquist, 2013).Sometimes authors imply that digital science is equal to data scienceunderpinning the role of other operations of research process apart from straight data analysis via a computer. This paper argues that S&T environment is extremely dynamic: process of knowledge creation becomes more intensified and sophisticated; researchers gradually learn new things, concepts and elaborate outstanding ideas.

Earlier we defined the stages of scientific process: searching for scientific problem, planning, data collection, analysis, publication and impact spreading. The biggest impact of digitalization appears within the last three stages, because in that time researchers have to spend significant amount of time working with digital tools, gathering and cleaning data (i.e. the results of statistical surveys), preparing descriptive statistics, doing simulations, data mining etc. On the first stage of idea search creative skills still play the leading role. Utilization of digital tools on this stage might cause even negative effects on thinking processes, because they develop quite limited path for new ideas. During the planning phase the most important task is proper development of experiments (whether it is survey or a microorganism generation inside a laboratory), and extracting sufficient adequate data sets, which could really help to answer research questions or prove a claimed hypothesis. Simultaneously it is important not to make a biased choice, thus it might cause negative consequences for a study. The next steps of data collections and analysis remains almost the same for individual researchers in terms of basic procedures, but for policymakers now it is important to pay more attention to open access policies, IP protection and diffusion of digital literacy. The last two stages mainly refer to the quality of researchers' soft skills and imagination for results visualization and clear presentation for broad audiences. For policymakers the predictive power of these results would contain the highest value.

Impact on research skills

Intensive technological environment and high interest in Big Data led to a rapid increase of demand for digital researcher skills. Frequently they are understood as generic ICT literacy, rarely as professional IT (e.g. software development) and big data skills, but the distinction should be made due to the complexity of existing tasks, which researchers are trying to solve abusing data. Typically however, in literature digital skills are combined into 3 groups, none of which is dedicated purely to science: skills for innovation (OECD, 2011), skills for industry (Harris, Murphy & Vaisman, 2013) and digital competencies (simple e-skills) (Ferrari, 2013; European Commission, 2014b; Binkley et al., 2014, Hayes, 2015; Vuorikari et al., 2016; OECD 2016). All these groups are frequently discussed among policymakers in the context of general spreading of ICTs and rising digital competences of populations. This means, that specific digital skills for science receive less attention (Holtz, 2014; BDV, 2015).

Generally speaking the toolkit of digital skills for science includes the desire to synthesize new knowledge (i.e. through adoption of new technologies, data analysis) and make it shareable, to challenge existing assumptions, and remain open to risks and changes. Any type of economic environment whether it is industrial, knowledge-based or digital, requires appropriate conditions for these capabilities to flourish. Within the digital agenda of science, high priority is given to data-related skills and professions, for instance, data scientists and analytics, specialists in big data, IoT and artificial intelligence (AI)). These are people who "make discoveries while swimming is data" (H. Davenport &Patil, 2012). Data scientists are usually considered as those who transform data analytics into business practices, and their role in science is not as clearly determined.

We distinguish two main types of impacts caused by digitalization on researchers skills: direct (mainly on codified knowledge, i.e. access to information) and indirect (on intangible assets like the width of someone's audience etc.). ***Direct impact***states for the growing availability of information and data (through open access, open source, various digital platforms and government e-infrastructures). This brings new challenges in developing ethics, adapting data standards and designing new education and training policies and programs for a new generation of digitally high-skilled researchers and even populations. ***Indirect impact.***Apart from the actors, digitalization affects the process of knowledge creation itself, its external environment and all the other components. As it widens access for scientific knowledge [codified: articles, data, etc.] and opens new channels for faster and intensive cooperation [cloud storages, data sharing platforms, social networks etc.] the audience (scientific communities) reach also broadens, what brings to a number of important consequences. All requirements for research quality (both qualitative and quantitative analysis) are rising rapidly due to ubiquitous harsh competition, numerous efficiency evaluations and essential protection of intellectual property. Researchers are facing new challenges in expressing their ideas through publications, due to the lack of data sharing standards and unspoken willingness to disclose less to stay competitive in terms of bibliometric indicators.Digitalization also causes additional external pressure of 'openness', which require researchers to adapt and change their routines and behavior, but for the whole scientific community and system the process of innovative changes and application of new digital and "open-mindset" principles might is time-consuming.

Discussion and Conclusions

Digitalization of science isn't a new phenomenon however the full range of its impact on scientists and researchers remains slightly nebulous. Recently science is considered 'open science' which includes mainly open access and open repositories but also open teaching and education. Whereas at first sight these developments are reasonable and a logic consequence of technological development reality shows that scientists and researchers aren't fully prepared to make use of them to the fullest extent. So far little evidence is provided for the reasons of this. However one might assume that the research and personal skills required for using the digital potential fully is to different from the established long recognized skills for research and engineering but this is certainly to short thinking. A more plausible reason for not fully using the opportunities provided by the digital age at the scientist and researcher side is their motivation and the rules of the scientific community. Scientists are by tradition organized in communities which are mainly looking at publications. Now publications traditionally include results only but not data. Data are at least to some extend the 'currency' for scientists on which their research output, e.g. publications, is build. Therefore disclosing full datasets is not always common behavior within the scientific community. The scientific community requests publications to describe the approach towards obtaining data in full transparency to allow any third party repeating the work done (experiment or similar) but not the actual dataset. This contradicts open science at least to some extend and it will require a while for the scientific community to change its attitudes and behavior. In addition challenges towards open data repositories arise related to finance and maintenance of databases but also copyright and related obligations. The scientific community argues that there is increasingly administrative burden affiliated with data disclosure which is 1) preparing datasets to be accessible (coding, missing values, manuals for understanding databases among others) and 2) safeguards that the original data is not changed or manipulated and if this is the case who's liable eventually. One frequently neglected feature here is that scientists and researchers reputation strongly depends on the datasets and their quality, thus once a dataset is manipulated and based on this is the scientist or researcher is found neglecting scientific values and norms the career of the individual or even team is almost over. Within the scientific community it takes comparatively long time to build a reputation, it takes short time to destroy it in worst cases. Repairing a reputation even if the scientist isn't responsible for manipulated data is almost impossible as delivering evidence again takes long time. Taken these concerns together scientists and researchers remain rather reluctant towards disclosing more than the absolute necessary for the sake of their self-protection.Eventually it appears that professional skills are becoming less important than soft and tacit. Hayes (2016) showed, that communication skills are the most relevant (> 50-75%) for all types of researchers jobs (incl. developers, business managers, idea-generators etc.)

Progressing digitalization of science opens new possibilities for the scientific community to enhance cooperative undertakings in different forms and intensity. This becomes evident when analyzing the changing shape of linkages within the KT. Although many linkages, e.g. channels, require personal relations and related trust between the actors involved the digital development enables the channels to be applied and used more frequently and more intense. Trust as it was earlier considered the precondition for relationships remains in place but interactions multiply in terms of frequency and intensity thus becoming even more effective and efficient. Still this not an automatic mechanism, it requires more coordination and matchmaking given the digital environment than before. In addition it requires changing attitudes of all actors and different skill sets especially among scientists. In line with this open science, namely in form of open access, can be seen as a critical point. The critical nature of open access is hardly with the open and public availability of science's research results but it lies with the attitude towards making all research publicly available which is at least partially against the established routines of the scientific community. Moreover the digital age allows science to obtain more data and information and to store them but this also calls for an increasing skills base to process and analyse these data and information. One side effect in this regard is privacy especially in social sciences which aim to obtain and analyse different data related to individuals. This also sheds a light on the increasing complexity underlying the data intensive driven nature of science, the more personal and private data used for scientific endeavors the more complex research itself becomes because of the involvement of legal aspects. The latter are also important for the increasing pressure on science in the frame of KT which aims at fostering collaboration of different partners. Any type of collaboration is featured by a solid legal agreement but this is more challenging when it comes to cross border collaborations regardless if the actors are public or private. Consequently the scientist skills sets change in the light of digitalization and KT imposed requirements towards strong data analysis competencies in most scientific fields and understanding of the legal environment combined with the ability to communicate with legal staff in related matters. Eventually scientists are confronted not only with additional skills for information and data processing but also for administration surrounding their actual research work.

Against these developments there appears a strong need for public policy to refine and adjust the respective framework conditions for research but also for industry. During the last years intensive policy debates have emerged which mainly aim at supporting infrastructures for open science and open access. Still these debates and respective measures don't take the full impact of digitalization into account. While infrastructural measures and initiatives are an absolute requirement the soft factors such as research behavior and tradition, norms and values of scientific communities and research routines need also to be taken into account. Last but not least there remains a gap between monitoring and steering public research in form of performance evaluations at different levels and the digital movement. Researchers performance evaluations remain at publication and citation counting mainly with little attention paid to the open access and full data disclosures. Moreover publications as the perceived main research output are subject to restrictions for open access due to the financial models connected and researchers in many cases don't have budgets available for doing this. Moreover most open access journal suffer from indexation which in turn is essential for researchers to meet their ambitions and goals and performance indicators. One might argue that there is an urgent need for first movers to use open access and publish in less impactful journals but this is a luxury only the most prominent researchers might afford because the majority of researchers careers depend on publications in established impactful journals instead of open access (due to impact factors). Thus the danger arises that the new open access movement is dominated by the established recognized community building another closed network and / or the 'money buys publication' attitudes which is frequent in emerging countries and less qualified researchers in developed countries. Hence open access faces the challenge of low quality which has a long term impact on the importance and meaning of science. Therefore policy needs to rethink and align the policy measures in place towards considering the behavior and standards of research more actively.

Acknowledgement

The article was prepared within the framework of the Basic Research Program at the National Research University Higher School of Economics (HSE) and supported within the framework of the subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

References

BDV (Big Data Value Association),. (2015). *European Big Data Value Strategic Research & Innovation Agenda* (pp. 34-40). Brussels.

Big Data - *SKA Telescope*.(2016). *SKA Telescope*. Retrieved 18 August 2016from http://www.skatelescope.org/newsandmedia/outreachandeducation/skawow/big-data/

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., &Runmble, M. (2014, December 8). Partnership for 21st century skills. Retrieved from http://www.p21.org/

Bohle, S. (2013). What is E-science and How Should it be Managed?. SciLogs.

Borrás, S. &Edquist, C. (2013).The choice of innovation policy instruments. *TechnologicalForecastingAndSocialChange*, *80*(8), 1513-1522.

Bošnjak, A., &Podgorelec, V. (2016). Upgrade of a current research information system with ontologically supported semantic search engine. Expert Systems with Applications, 66, 189–202. https://doi.org/10.1016/j.eswa.2016.09.011

Carayannis E.G., Meissner D., Edelkina A. (2015) Targeted innovation policy and practice intelligence (TIP2E): concepts and implications for theory, policy and practice. The Journal of Technology Transfer, doi: 10.1007/s10961-015-9433-8

David, P. A. (2004). "Understanding the emergence of 'open science' institutions: Functionalist economics in historical context". Industrial and Corporate Change. 13 (4): 571–589. doi:10.1093/icc/dth023

Eordogh, F. (2016). *What The Pokémon GO Mania Says About Modern Society*. *Forbes.com*. Retrieved 18 October 2016, from http://www.forbes.com/sites/fruzsinaeordogh/2016/07/13/what-the-pokemon-go-mania-says-about-modern-society/#144731051e13

European Commission (2014).The digital agenda toolbox. Retrieved October 16, 2016, from http://publications.jrc.ec.europa.eu/repository/bitstream/JRC88896/ipts%20jrc%2088896%20(online)%20final.pdf

European Commission,. (2014b). *E-skills for Jobs in Europe: Measuring Progress and Moving Ahead*.

European Commission. (2016) Realising the European Open Science Cloud. Retrieved October 14, 2016, fromhttp://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none

Ferrari, A. (2013). *DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe*. Luxembourg: Publications Office of the European Union.

Gantz, J., & Reinsel, D. (2011). *IDC IVIEW: Extracting Value from Chaos*. Retrieved August 16, 2016, from http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf

Hall, G. (2015) What does Academia_edu's success mean for Open Access? The data-driven world of search engines and social networking. Retrieved October 18, 2016, from http://blogs.lse.ac.uk/impactofsocialsciences/2015/10/22/does-academia-edu-mean-open-access-is-becoming-irrelevant/

Harris, H., Murphy, S., & Vaisman, M. (2013). *Analyzing the analyzers*. Sebastopol, Calif.: O'Reilly Media.

Havas A. Izsak K., Markianidou P., Radošević S., (2015). Comparative analysis of policy-mixes of research and innovation policies in Central and Eastern European countries, GRINCOH Working Paper Series, Paper No. 3.12.*SSRN Electronic Journal*.

Hayes, B. (2015). Getting Insights Using Data Science Skills And The Scientific Method. Business Broadway.

Hayes, B. (2016). *The Most Important Skill in Data Science: Mining and Visualizing your Data.@bobehayes*. Retrieved 17 October 2016, from http://businessoverbroadway.com/the-most-important-skill-in-data-science-mining-and-visualizing-your-data

Hirsch-Kreinsen, H. &Schwinge, I. (2014). *Knowledge-intensive entrepreneurship in low-tech industries*.

Holtz, D. (2014). 8 Skills You Need to Be a Data Scientist. *Udacity*. Retrieved from http://blog.udacity.com/2014/11/data-science-job-skills.html

Kintisch, E. (2014) Is ResearchGate Facebook for science? Retrieved October 18, 2016, from http://www.sciencemag.org/careers/2014/08/researchgate-facebook-science

Medeiros, C. & Katz, D. (2016).eScience today and tomorrow. *Future Generation Computer Systems*, *56*, 523-525. http://dx.doi.org/10.1016/j.future.2015.10.016

Meeker, M. (2016). *2016 Internet trends report*. Retrieved August 16, 2016, from http://www.kpcb.com/internet-trends

Meissner D. (2015) Public-Private Partnership Models for Science, Technology, and Innovation Cooperation. Journal of the Knowledge Economy, doi: 10.1007/s13132-015-0310-3

Meissner, D. (2014): Approaches for developing national STI strategies. STI Policy Review, 5(1), 34-56

Meissner, D. Shmatko, N. (2016) “Keep open”: the potential of gatekeepers for the aligning universities to the new Knowledge Triangle. Technological Forecasting and Social Change, http://dx.doi.org/10.1016/j.techfore.2016.03.012

Meissner, D., Carayannis, E., & Sokolov, A. (2016). Key features of roadmapping for company and policy strategy making. Technological Forecasting and Social Change, 110, 106-108.

OECD (2011), *Skills for Innovation and Research*, OECD Publishing, Paris.

OECD (2015), "Making Open Science a Reality", OECD Science, Technology and Industry Policy Papers, No. 25, OECD Publishing, Paris.
DOI: http://dx.doi.org/10.1787/5jrs2f963zs1-en

OECD (2016), New Skills for the Digital Economy, *OECD Digital Economy Papers*, No. 258, OECD Publishing, Paris.

PokémonGo passed 100 million installs over the weekend. (2016). *TechCrunch*. Retrieved 16 August 2016, from https://techcrunch.com/2016/08/01/pokemon-go-passed-100-million-installs-over-the-weekend/?ncid=rss

Raju, J. (2014). Knowledge and skills for the digital era academic library. *The Journal Of Academic Librarianship*, *40*(2), 163-170. http://dx.doi.org/10.1016/j.acalib.2014.02.007

Sarpong D., Razak A.A., Alexander, E., Meissner D. (2016) Organizing practices of university, industry and government that facilitate (or impede) the transition to a hybrid triple helix model of innovation. Technological Forecasting and Social Change, in print, http://dx.doi.org/10.1016/j.techfore.2015.11.032

So You Think You Have Skills. (2012). *Science | AAAS*. Retrieved 12 October 2016, from http://www.sciencemag.org/careers/2012/09/so-you-think-you-have-skills

Vancauwenbergh, S. (2016).Governance of research information and classifications, key assets to interoperability of CRIS systems in inter-organizational contexts. Presented at the 13th International Conference on Current Research Information Systems, CRIS2016, Scotland, UK: Procedia Computer Science. Retrieved October 12, 2016, from http://dspacecris.eurocris.org/bitstream/11366/508/1/CRIS2016_paper_22_Vancauwenbergh.pdf

Vuorikari, R., Punie, Y., Carretero, S., & Van den Brande, L. (2016). *DigComp 2.0: The Digital Competence Framework for Citizens*. Luxembourg: Publication Office of the European Union.

Meissner Dirk
National Research University Higher School of Economics, Institute for Statistical Studies and Economics of Knowledge, dmeissner@hse.ru