Shaping digital transformation in science.

“Digital Information” Initiative by the Alliance of Science Organizations in Germany

Mission statement 2018 - 2022

Executive Summary

To equip scientists with the best possible information infrastructure they need for their research, that is the guiding principle behind the "Digital Information"[1] Initiative launched in 2008 by the Alliance of Science Organizations in Germany. With the Alliance Initiative, the Science Organizations are declaring their readiness to coordinate their basic political and technical decisions in the sphere of digital information, pool the resources they invest for this purpose and, if necessary, provide further funds. They are also reinforcing their determination and ambition to shape digital transformation in science and create a platform for joint action.

The previous period of cooperation as part of the "Digital Information" Alliance Initiative ended in 2017. Examples of results are the establishment of the two Alliance projects DEAL (Nationwide Licensing of Portfolios of major Academic Publishing Houses) and OA2020.DE (National Open Access Contact Point), the position paper "Research data at your fingertips" and statements and joint initiatives on copyright.

In view of the persistently dynamic changes which digitization entails for all areas of science, the need for a structure for networking, coordinating and taking action as exemplified by the Alliance Initiative remains as high as ever. The successful cooperation of Science Organizations in Germany in the Alliance Initiative will therefore be maintained for 2018 to 2022. The basis of the cooperation is explained in the present mission statement *Shaping digital transformation in science. "Digital Information" Initiative by the Alliance of Science Organizations in Germany 2018 — 2022.* The mission statement describes the principles and aspirations of the Alliance Initiative (Chapter 1), outlines the starting point and the thematic environment into which the Initiative fits in the third period of its cooperation (Chapter 2) and identifies the fields of action for digital transformation in science as well as the resulting potential priorities for the "Digital Information" Alliance Initiative (Chapter 3).

With respect to the subject of "Digital Information", it is the aspiration of the Alliance Initiative

- to shape and modify relevant operational structures and processes;
- to pick up on stimuli and bring them to fruition;
- to agree on joint strategic statements and policy papers and in doing so shape discussion on science policies;

[1] hereafter referred to as the: Alliance Initiative

- to draw up joint positions for science organizations on topics relating to digital information, and introduce and represent them in discussions of science policies;
- to promote the exchange of experiences, as well as liaise and communicate on current topics;
- to support and, if necessary, coordinate German participation in international initiatives;
- To collaborate in drawing up principles and handouts of relevance to all science organizations.

From 2018 to 2022, the Alliance Initiative will address new expressions of digitization. Since the transformation from analogue media to digital objects is now well advanced, genuinely digital phenomena are now leading the way. For example, they include the cross-border networking of research and teaching on digital platforms, new forms of digital publications in science and their evaluation as a research output, machine analysis and interpretation of large volumes of research data, or the significance of good scientific practice in the digital age.

As part of the Alliance Initiative, priority will be given to the following fields of action in order to address interdisciplinary topics such as Open Access, Open Science or research data with the appropriate degree of precision:

1. Scientific publication system
2. Digital tools — software and services
3. Digital collections of data and text corpora
4. Federated IT infrastructure
5. Digital learning, teaching and networking
6. Digitally qualified staff
7. Legislation for science in the digital age
8. Scientific practice

1. Mission statement and aspirations

To equip scientists with the best possible information infrastructure they need for their research, that is the guiding principle behind the "Digital Information" Initiative by the Alliance of Science Organizations in Germany. This means that publications, research data, services, software and tools are to be made available to researchers in broadly usable form and if possible without any legal, financial, technical or organizational barriers.

The Alliance Initiative helps to support joint responsibilities and objectives of science organizations as set out in their respective statutes. These are (1) to raise the performance and competitiveness of German science, (2) to exert a formative influence on European and international science and (3) to boost the recognition of science in society, business and politics. All fields of action in the digital transformation of science on which the Alliance Initiative is focusing, can be assigned to one of these overarching objectives.

The Alliance Initiative's focus lies on information infrastructures. Based on the definition of the Council for Scientific Information Infrastructures (Rat für Informationsinfrastrukturen - RfII), these are seen as "technically and organizationally networked services and products for accessing and receiving data, information and knowledge. First and foremost, they serve research purposes, are frequently the object of research and always have an enabling function." [2] From the perspective of the Alliance Initiative, the focus is on the benefit of information infrastructures for scientists. The design of information infrastructures is thus the product of weighing up what is desirable from a scientific viewpoint against what is organizationally and technically feasible as well as economically sensible. The Alliance Initiative sees itself as an enabling structure — both in a practical sense by working on common concerns shared by the participating science organizations, and also in the sense of supporting scientific discourse by pointing out the positive and any negative potential and the effects of digitization for science.

Giving appropriate form to digital transformation in science in the coming years will entail maintaining an overview of and evaluating current developments but also tackling acute problem areas. The Alliance Initiative sets the substantive priorities for issues where particularly large potential is identified for using the

[2] cf. Leistung aus Vielfalt, Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland [Performance from Diversity, Recommendations on Structures, Processes and Funding of Research Data Management in Germany], Göttingen 2016, p. A-15, http://www.rfii.de/?wpdmdl=1810

opportunities offered by digital technologies on behalf of scientific work with respect to the entire research cycle. Cooperation in the Alliance Initiative helps to adapt scientific publishing to the potential of digital technology, design access to and the use of large data and text corpora or ensure that all groups of scientific staff acquire digital competence. In addition, it is the Alliance Initiative’s aspiration to support and help shape dialogue on the level of science policy as well as discussions inherent in science with respect to the design of information infrastructures for science in the digital age. These include legislative processes with a high degree of relevance to science under digital conditions such as modifications to copyright law or developments such as the setting up of a National Research Data Infrastructure (Nationale Forschungsdateninfrastruktur - NFDI).

The Alliance Initiative works on several levels. It is a network for the management tier in the area of digital information right across all German science institutions. It offers a working structure to highlight matters of acute concern to digital information which are in their common interests and to discuss or initiate suitable options for implementing action. It aims to be the first port of call when it comes to coordinating the dialogue on digital transformation between science institutions and the political classes and general public within the purview of all German science institutions.

The Alliance Initiative sees itself as a complementary body to organizations with the responsibility to analyse systems or advise on them such as the Council for Scientific Information Infrastructures.

By continuing the “Digital Information” initiative, the partner organizations in the Alliance are aiming in the future to raise the performance of Germany as a place to conduct science by pooling expertise and resources and jointly working on and driving current and future issues. In particular, they are interested in coordinating their basic political and technical decisions in the area of information infrastructures, pooling their financial resources deployed for this purpose and, if necessary, providing additional resources.

2. Starting point

What forms the digital transformation of science and society will take in the coming years is barely foreseeable given the rapid pace of technological developments. However, it is already clear today that the international competitiveness of a country in the field of science will depend crucially on its infrastructure for digital information. The Federal Government has highlighted the issue in its Digital Agenda[3], the Federal Ministry of Education and Research has established the working group Digital Transformation consisting of three units[4] and the Council for Scientific Information Infrastructures has described the management and sustainability of digital research data as one of the major challenges facing science in the next few years[5].

Information infrastructures together with their content, research data and software will play a significant role in the next few years in the development and implementation of strategies on digital information in science. Major research questions of high relevance to society such as climate, nutrition, health or migration cannot be examined in expert groups spread all over the world and using data-intensive methods without digital information and networking. Separate research areas for digital information are developing in information science and data science for analysing social networks and usage data, for Artificial Intelligence, model simulation or for laws and ethics applicable to the digital world. In view of digital transformation, it is remarkable that the reputation system of science with its traditional performance indicators and metrics is focused on classical publishing; science in the digital age needs to adapt.

Nearly every form of scientific communication today is based on digital tools and data regardless of whether a particular discipline works digitally for the most part. In view of the need for dynamic and interactive means of presentation, data-intensive science shows that formats for communicating knowledge established over many centuries and exhibiting the characteristics of a "book" or "magazine article", are strictly limited in their ability to communicate. New forms of presentation offer new opportunities not only for research but also for communicating scientific content in teaching, for learning and for the presence of science among the wider public, and together these forms are contributing towards the development of a digitally enlightened society.

[3] https://www.digitale-agenda.de/
[4] http://www.bildung-forschung.digital/
[5] http://www.rfii.de/?wpdmdl=1998

For one thing, a steady rise in data volumes is to be observed, particularly in cutting-edge research, leading to a hundredfold increase in the volume of data over 10 years, akin to Moore's Law. And for another, unique data phenomena are occurring, whether as a result of new large equipment for measuring nature, billions of distributed, networked devices with the facility to analyse human behaviour or in the complex diversity of meanings in global, digital text corpora and aggregations of data. The last few years have also been marked by the penetration of science by new digital services in which start-ups play a crucial role alongside the major established players in the software and information industry and in publishing. This indicates that even the economic framework conditions for scientific information continue to pose major challenges. On the one hand, there is a risk that genuinely scientific functions will be monopolized or commercialized, e.g. the scientific reputation system — and on the other, new innovative markets are arising, e.g. in scientific publishing, which can be scientifically nurtured and exploited to make science even more powerful and to support the market's capacity for innovation.

The Alliance Initiative is facing up to these challenges by defining selected fields of action for digital transformation and deriving specific proposals for action (cf. Chapter 3).

3. Fields of action and priorities for collaboration

The effects and opportunities of Digital Transformation encounter the active players of the scientific world in many fields. They are shown in the following chart at the level of aggregation suitable for this document, and they are termed fields of action in the sections below.

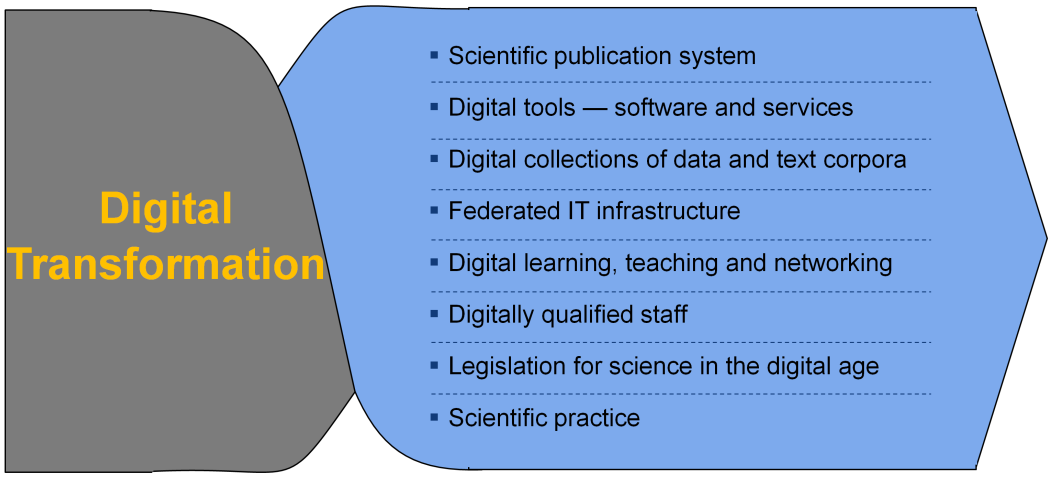

Fig. 1: Fields of action in Digital Transformation and the players in the scientific world encounter each other in numerous fields of action.

By playing an active role in these fields of action, the Alliance Initiative is helping to support the responsibilities and objectives of the science organizations as defined in their statutes (cf. Chapter 1). The fields of action form the thematic foundation for maintaining the Alliance Initiative. They are defined in relation to functionalities and processes in science, and not in relation to individual objects or contents such as research data or software. For the latter, no separate working groups are deployed — contrary to the previous working structure of the Alliance Initiative — rather they are addressed in several fields of action in relation to their specific scientific function.

The priorities for cooperation in the relevant fields of action lie in the areas where the joint interests of the Alliance organizations coincide.

The Alliance Initiative claims for itself the ability to act flexibly and react quickly from a stable structural base (stability vs. agility). It also claims to be able to act in principle in all the fields outlined above, if required. Against this backdrop, the working programme of the Alliance Initiative will be further refined in the period from 2018-2022, as and when needed.

F1: Scientific publication system

Field of action: The scientific publication system continues to expand in terms of volume and is characterized by a multiplicity of (also new) players with different interests. The primary objective to be pursued in this area is to exhaust the potential of digital publishing. This is achieved by ensuring Open Access to scientific content, exploiting the potential of the digital options and organizing publishing as a science-friendly service market on a sustainable basis. Derived objectives include establishing simple, comprehensive access to scientific information to support digital teaching and education, increasing the visibility of science from Germany, e.g. by achieving a high proportion of Open Access, and influencing global publishing by means of national solutions serving as examples, e.g. national master agreements for licences, services and business processes.

The transformation of publishing is being conducted on an organizational and economic level as well as a technological level. The term publication therefore also encompasses new formats such as the reusable or executable combination of text, code, research data and further media content (*enhanced publications*), optimized versions for certain types of content such as data, videos or software (*data journals, video publications, code repositories*) or updatable publications (*living publications*). Publications are a central element in the research cycle although work is published faster than was previously the case, and the phases of quality control, reception and commentary can overlap, maintaining transparency and openness.

Priorities for the Alliance Initiative: One central aspect of future work in this area is the stronger link between Open Access and licensing. The intention is to set standards, above all as part of Open Access transformation agreements and Gold Open Access framework agreements. Alongside the transformation of journals mainly from the areas of the natural sciences, technology and medicine (STM) which is to be implemented as part of the Alliance project for the nationwide licensing of the portfolios of major publishing houses (DEAL) and with the support of the National Open Access Contact Point (*OA2020.DE*), the establishment and support of novel business models and means of transformation relevant to further disciplines, will also play an important role. Already existing, cooperative funding approaches such as for monographs can be communicated and increasingly established in Germany. Of decisive importance for the entire transformation to Open Access are cost transparency and the sustainability of publication infrastructures. The tools of the transformation are located on an operational

level, and there is a need for adequate, efficient processes at institutions. Models for the self-hosting of content also form part of the overall concept for providing information. Open Access repositories, their relationship with institutional research information systems, their potential for Open Science and alternative supply routes are also a reality which must not be disregarded in the course of future conceptual work in this area. Continual dialogue on a policy level, above all with the Federal Ministry of Education and Research during the implementation phase of the Open Access strategy, is to support the action.

F2: Digital tools — software and services

Field of action: Software and digital services are essential today for scientific work which means that research and teaching are dependent on commercial, Open Source-licensed and individually self-created software solutions. To this extent, creating access to commercial solutions, obtaining and refining free solutions and facilitating the development of individual solutions are increasingly becoming indispensable infrastructure tasks. In terms of operational practice, it is helpful to distinguish between (i) individual research software such as simulation or control software, (ii) standard software such as Office software, mathematical software or visualization packages and (iii) online services such as search engines, file-sharing tools, Open Source code repositories, document servers, data repositories, computing and storage cloud services, identity services and social networks. The business models range from free-of-charge and free of any commitment (*open or libre*) via data usage ("*free*") and classically commercial ("*paid*") to institutionally funded (*sponsored*).

Priorities for the Alliance Initiative: In connection with standard software and online services, there are challenges caused by the individual interests of the commercial or non-commercial suppliers or bodies which must be appropriately addressed through licensing or usage agreements. With regard to federated or syndicated, publicly funded structures (e.g. German Research Network or National Research Data Infrastructure) which can support infrastructure and repository services, regular dialogue with the field of action "Federated IT infrastructures" must be provided for. In terms of research software which is often created in-house as part of fixed-term projects, adherence to the principles of "good scientific practice" (such as plausibility, reproducibility, transparency and quality control) must be supported and questions of re-usability and licensing (such as issues relating to copyright, patent and liability legislation) dealt with. With regard to

generic aspects of academic software development, there are already European initiatives such as *Knowledge Exchange* or *softwareheritage* with which it seems helpful to collaborate. Publication of the source code (together with associated parameters, documentation and workflows) of research software developed in many disciplines in accordance with wide-ranging and specific requirements, represents a type of scientific result *sui generis* which, in the same way as published research data, is increasingly to be seen alongside classical text publications and is to be considered in consultation with the field of action "Scientific publication system".

F3: Digital data collections and text corpora

Field of action: The increasing availability of sources for data and texts in digital form is allowing previously inaccessible research questions to be handled and software solutions to be created for new areas of application (e.g. speech comprehension, navigation in physical space, intelligent web searches). The indispensable basis for such developments is the availability of well-maintained data collections and appropriately curated text corpora with the right subject content, appropriate quality and up-to-dateness and transparent processes for how they came into existence. The benefit bestowed by modern analytical methods such as Machine Learning (ML), Text and Data Mining (TDM) and Artificial Intelligence (AI), often espoused with considerable hopes with regard to their business development and financial potential, depends on the availability of suitable data collections and text corpora to which they can be applied. Data collections and text corpora are not just the basis for research but also an end product created as part of publicly funded scientific work and commercial efforts and of equal importance to research and business.

Priorities for the Alliance Initiative: It is essential to promote access to the digital and digitized cultural heritage from libraries, archives and museums (such as primary literature, archive material, objects and artefacts), to digital research data (including high-quality, top-class data for AI training purposes) as well as to usage and operational data created outside of the research process but nevertheless of interest to the research. It is also important for any scientific evaluation that citation indices and reference data are freely available and usable, and that data collections are available for alternative metrics (*altmetrics*). In addition, progress is to be made in the distribution of data in accordance with the principles of Findability, Accessibility, Interoperability and Reusability (*FAIR data principles*).

From a methodological perspective, particular emphasis is to be given to harnessing the opportunities of Text and Data Mining (TDM). Legal questions, e.g. in connection with TDM, are handled in connection with the field of action "Legislation for science in the digital age".

F4: Federated IT infrastructure

Field of action: Scientific IT infrastructure is based on the interplay of basic infrastructure (technical basis of the internet), increasingly visualized, cloud-based basic functionality (storage, computing power) as well as genuinely scientific applications and data structures. Science makes the highest demands and sets standards for the necessary interaction of all levels of the IT infrastructure, as we are already familiar with from high-performance computers (HPC) and these days to be seen across a wide range of highly-specialized, data-intensive research in numerous information infrastructures for specific disciplines. The federation of information infrastructures is a highly topical subject both at a national level in the context of a national research data infrastructures and also at an international level with regard to the European cloud initiatives. It is the responsibility and business of science organizations to contribute to the success of national and supra-national networking structures by ensuring that the task of designing the contents of processes is driven by science and adheres closely to scientific practice.

Priorities for the Alliance Initiative: Federated scientific IT infrastructures are currently covered within several contexts: in Germany on an interdisciplinary basis (*NFDI*) and by specific discipline (e.g. *GFBIO*) and as national contributions for European infrastructures (e.g. *ARGO*, *ELIXIR*, *CESSDA*, *SHARE*, *DARIAH*, *EUDAT*, *CLARIN*) as well as in Europe as part of the European Cloud Initiative (*European Data Infrastructure*, *European Open Science Cloud*). Activities range from setting up commercially available, basic IT infrastructure to special, Open Source applications for specific subjects. With respect to the activities of the Federal Ministry of Education and Research (BMBF), the Joint Science Conference (GWK), RfII, the European Commission and others dealing with relevant transformation processes, the Alliance Initiative will identify those fields where there is a need for further action, and submit its own proposals for the design of such measures. The Alliance Initiative acts with the objective of initiating, preparing and supporting the communication and organizational processes of science organizations as part of

the political decision-making process, thereby contributing to the establishment of governance mechanisms commensurate with science.

F5: Digital learning, teaching and networking

Field of action: The effect that digitality has on learning and teaching and how science organizations can and should be networked with each other and with the public, represent current challenges facing digital transformation which will have a lasting effect on future generations. There are numerous public offerings for online media and online courses (Open Education Resources) — as well as commercial suppliers — in an agile market which leading US universities, in particular, are trying to exploit for themselves. Certificates for passing online courses tailored precisely to provide the specialized expertise currently required in science and business, reveal significant differences from university degree courses designed for the long term. This field of action tackles these challenges, starting from the question of how to teach digital content and make it generally comprehensible, thereby distinguishing itself from transformation processes in science inherent in the system which are covered in the field of action "Scientific practice" (p. 8). A separate field of action is devoted to the subject of "Digitally qualified staff" (p. 6).

Priorities for the Alliance Initiative: Intensive cooperation between the players involved in digital *learning and teaching* in Germany to support the effective positioning of the German science system. The objectives are to facilitate access to third-party material and simplify the placement of proprietary research, educational and course material, i.e. online courses, introductory videos and methodological videos or digital course material reserved for students. The focus here for many Alliance organizations lies on how content relevant to the research can be combined with digital learning and teaching. As the German Rectors' Conference is independently responsible for the subjects of studying and teaching among the Alliance organizations, activities with specific and unique relevance to these subjects are not to be covered within the Alliance Initiative but selectively agreed. Within the context of digital networking, the internet presence of science in Germany can be reinforced via expert platforms, encyclopaedias such as Wikipedia and social networks for scientists. This is also to support the creation of generally accessible learning opportunities for scientific facts, methods and perspectives which can help to make public debates more objective and develop the digitally enlightened society. This field of action is new in the Alliance Initiative and is first

to be the subject of more intensive preparation using methods such as the exchange of information.

F6: Digitally qualified staff

Field of action: There is a considerable need, presently not met, for digitally qualified staff and specialized digital expertise in science and industry. The spectrum of skills required begins with basic applications on computers and mobile devices, includes the qualifications of planning and implementing digitally based processes and projects and ends with the advanced application of methods of data analysis, algorithms, programming, modelling and simulation. Addressing this supply shortfall requires both long-term design through changes to the curricula of undergraduate courses and also short-term, agile programmes in further education through training, online certificates or other extra-curricular formats.

Priorities for the Alliance Initiative: Close collaboration between universities, non-university institutions and industry is required to ensure that requirements are coordinated which can be achieved, for example, by greater exchange of expertise in joint further training measures, but also through reciprocal shadowing or in consultations on recruitment strategies for highly qualified staff. In addition, outstanding digital skills in science, particularly in junior scientists, must be systematically promoted to cover the peak needs for digitally and analytically trained graduates and recipients of further education both in science and in industry. A close relationship is to be maintained between the Alliance Initiative and the relevant activities of the IT Summit (IT-Gipfel), particularly the working group on digital skills. Subject areas such as the scientific ethics of the digital world or good digital research practice are also to be closely coordinated with the activities in the field of action "Scientific practice" on the basis of current developments.

F7: Legislation for science in the digital age

Field of action: Science operates in the digital age in a space that besides disciplinary, technical and organizational dimensions also has a legal dimension. Along this line, the scope for action open to science is either restricted or facilitated. The effects of complex legal framework conditions with respect to scientists' sphere of action often have to be first analysed and communicated. Within the sphere of action of Digital Transformation, the legal framework limits or permits the flow and use of information, e.g. through laws governing copyright, neighbouring rights, data protection and personality rights. It also regulates the digital

service markets which science makes use of, e.g. through competition law, anti-trust law and tax law.

Priorities for the Alliance Initiative: As an interdisciplinary topic within the Alliance Initiative, the legal requirements must always be covered in the whole area of digital science and its progress. The previous procedure of responding primarily to current legislative initiatives and newly arising requirements from science or politics by means of position papers or handouts, for example, has proved effective and is to be continued. The aim on the one hand is to help shape political processes, and on the other to inform the players concerned of changes to the framework conditions in a comprehensible manner.

F8: Scientific practice

Field of action: Science has ever since been characterized by the fact that research results are always generally available, the search for findings is conducted on an open-ended basis and any result must be critically reviewed. These fundamental principles are not called into question by digital technology. Rather digital technology can support the process of obtaining scientific insights, its transparency and the general availability of scientific findings to a degree that was previously not possible. However, this is not happening to the extent that would be possible or even desirable. There are elements of present scientific practice that conflict with this aim. For example, the dominant reputation mechanisms (e.g. of indicators and metrics) act to stabilise the subscription-based publication system, thereby making the transformation to a system based on Open Access harder to achieve. The almost exclusively publication-based reputation system also provides no incentive for scientists to actively drive the establishment of novel services specific to their discipline, such as developing community-based review mechanisms using digital technology or promoting publication formats where the replicability of research results is inherent in the process.

The effects of scientific practice on the relationship between science and society must also be examined in this field of action. Reflecting on and questioning scientific practice on our own initiative — e.g. whether existing technical opportunities are being exhausted with respect to major scientific principles such as communicating results to the outside world, their susceptibility to critical review or even the dominant mechanisms affecting careers — this is a sign of science's ability to reflect on itself, handle criticism and take action.

Priorities for the Alliance Initiative: The contribution made by the Alliance Initiative to the subject of scientific practice can consist in identifying from the perspective of the information infrastructures how the potential of digitization can be used for the benefit of these fundamental scientific principles. And vice versa: to identify the causal relationships currently making it harder to exploit such potential, or preventing it altogether. As a result, scientific disciplines are to be supported in reflecting on certain aspects of scientific practice linked to digitization and actively helping to shape them. For example, this relates to questions regarding the system of indicators: What indicators are suitable for mapping scientific performance? What aspects of performance are to be considered? It relates to questions regarding the transparency of data and information: What is the attitude of the various disciplines towards the transparency of data, software and publications? What information infrastructures are necessary from the perspective of the disciplines? And it relates to questions regarding undesirable developments: What undesirable effects are caused by digitization, and how can they be counteracted? The Alliance Initiative cannot take the lead in organizing these discussions, as they belong instead to the genuine area of responsibility of the scientific disciplines. The Alliance Initiative can identify the supportive function of information infrastructures, as well as the positive and negative potential affecting scientific practice as a result of the digitization of science.

***

Imprint

The online version of this publication can be found at:
http://doi.org/10.2312/allianzoa.016

Published by

Steering Committee for the “Digital Information” Initiative of the Alliance of Science Organisations in Germany

Contact

http://www.allianzinitiative.de

Version

December 2017

Translations

This document is a translation of the German publication “Den digitalen Wandel in der Wissenschaft gestalten. Die Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen. Leitbild 2018 – 2022“ that can be found online at http://doi.org/10.2312/allianzoa.015

Licence