OPEN ACCESS

CrossMark

Grant Proposal

NFDI4Culture - Consortium for research data on material and immaterial cultural heritage

Reinhard Altenhöner[‡], Ina Blümel[§], Franziska Boehm[|], Jens Bove[¶], Katrin Bicher[¶], Christian Bracht[#], Ortrun Brand[#], Lisa Dieckmann[¤], Maria Effinger[«], Malte Hagener[#], Andrea Hammes[¶], Lambert Heller[§], Angela Kailus[#], Hubertus Kohle[»], Jens Ludwig[‡], Andreas Münzmay[^], Sarah Pittroff[ˇ], Matthias Razum[|], Daniel Röwenstrunk[^], Harald Sack[|], Holger Simon[¡], Dörte Schmidt[?], Torsten Schrade[ˇ], Annika-Valeska Walzel[¶], Barbara Wiermann[¶]

‡ Stiftung Preußischer Kulturbesitz, Berlin, Germany
§ TIB Leibniz Information Centre for Science and Technology, Hannover, Germany
| FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, Karlsruhe, Germany
¶ Sächsische Landesbibliothek - Staats- und Universitätsbibliothek, Dresden, Germany
# Philipps-Universität Marburg, Marburg, Germany
¤ University of Cologne, Cologne, Germany
« Heidelberg University Library, Heidelberg, Germany
» Ludwig-Maximilians-Universität, Munich, Germany
^ Paderborn University, Paderborn, Germany
ˇ Academy of Sciences and Literature, Mainz, Germany
¡ Verband deutscher Kunsthistoriker e.V., Bonn, Germany
? German Musicological Society, Kassel, Germany

Corresponding author: Torsten Schrade (torsten.schrade@adwmainz.de)

Reviewable   v1

Abstract

Digital data on tangible and intangible cultural assets is an essential part of daily life, communication and experience. It has a lasting influence on the perception of cultural identity as well as on the interactions between research, the cultural economy and society. Throughout the last three decades, many cultural heritage institutions have contributed a wealth of digital representations of cultural assets (2D digital reproductions of paintings, sheet music, 3D digital models of sculptures, monuments, rooms, buildings), audio-visual data (music, film, stage performances), and procedural research data such as encoding

and annotation formats. The long-term preservation and FAIR availability of research data from the cultural heritage domain is fundamentally important, not only for future academic success in the humanities but also for the cultural identity of individuals and society as a whole. Up to now, no coordinated effort for professional research data management on a national level exists in Germany. NFDI4Culture aims to fill this gap and create a user-centered, research-driven infrastructure that will cover a broad range of research domains from musicology, art history and architecture to performance, theatre, film, and media studies.

The research landscape addressed by the consortium is characterized by strong institutional differentiation. Research units in the consortium's community of interest comprise university institutes, art colleges, academies, galleries, libraries, archives and museums. This diverse landscape is also characterized by an abundance of research objects, methodologies and a great potential for data-driven research. In a unique effort carried out by the applicant and co-applicants of this proposal and ten academic societies, this community is interconnected for the first time through a federated approach that is ideally suited to the needs of the participating researchers. To promote collaboration within the NFDI, to share knowledge and technology and to provide extensive support for its users have been the guiding principles of the consortium from the beginning and will be at the heart of all workflows and decision-making processes. Thanks to these principles, NFDI4Culture has gathered strong support ranging from individual researchers to high-level cultural heritage organizations such as the UNESCO, the International Council of Museums, the Open Knowledge Foundation and Wikimedia. On this basis, NFDI4Culture will take innovative measures that promote a cultural change towards a more reflective and sustainable handling of research data and at the same time boost qualification and professionalization in data-driven research in the domain of cultural heritage. This will create a long-lasting impact on science, cultural economy and society as a whole.

Keywords

Research Data Management, NFDI, NFDI4Culture, Cultural Heritage, Art History, Musicology, Architecture, Theater Studies, Dance Studies, Media Studies, Digital Humanities

Consortium

Research domains or research methods addressed by the consortium, objectives

NFDI4Culture aims to create a user-centered, research-driven infrastructure for **research data on non-textual material and immaterial cultural heritage**. With its focus on cultural assets, the consortium covers a **broad range of research domains from musicology, art history** and **architecture** to **performance, theatre, film, and media studies**. The research landscape addressed by NFDI4Culture is characterized by strong institutional

differentiation. Research units in the consortiums' community of interest are often very small (right down to individual researchers) and comprise **university institutes, art colleges, academies** and **cultural heritage institutions** such as **galleries, libraries, archives and museums** (GLAM). At the same time, this community is very rich regarding its research objects, methodologies and its future potential for data-driven research: Digital cultural assets, practices and performances have become an essential part of our daily life, communication and experience and have a lasting influence on our individual and societal perceptions (Münzmay 2018, Hadjakos et al. 2017, Kailus and Stein 2018, Kailus 2017). **The long-term preservation and FAIR availability of research data from the cultural heritage domain is therefore fundamentally important, not only for future academic success but also for cultural identity, cultural economy and society** (Wilkinson et al. 2018). Although many relevant institutions working in the field already have appropriate technological means, experience and solutions at their hands, until now, the community lacks a coordinated effort for professional research data management and curation on a national level.

To fill this gap, **several academic societies** ranging from musicology and art history to architecture, theatre, film and media studies **have joined with the applicant and co-applicant institutions** of this proposal **in a unique effort to tie together the heterogeneous existing solutions** and shape a research data management strategy and work program that is **ideally suited to the strong needs** of the involved community of interest. NFDI4Culture will create a **new and sustainable infrastructure** that is based on a **thorough analysis of the existing research landscape.** The consortium will integrate, merge and enhance a **broad range of existing solutions** and **take innovative approaches** to research data management in a field relevant to science and society as a whole (Fig. 1).
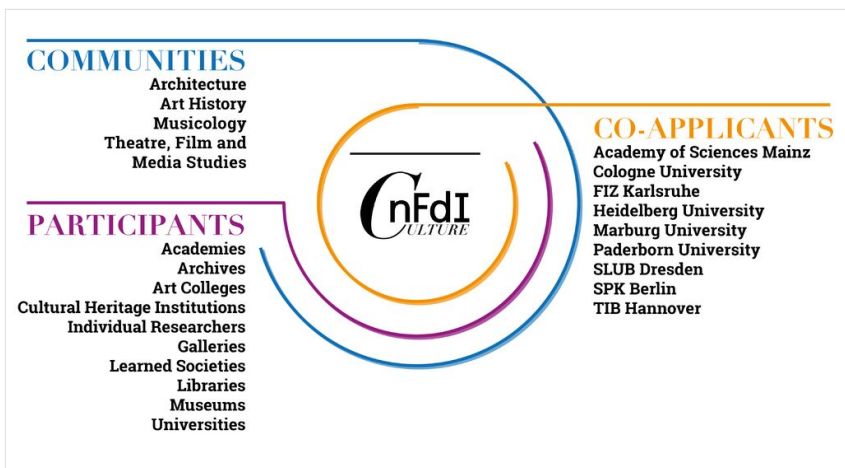

Figure 1. doi

Communities, Co-Applicants and Participants of NFDI4Culture

NFDI4Culture's community of interest produces a great variety of research data throughout the entire research process through close interaction with GLAM institutions and infrastructure facilities. **Up to now**, **data has been collected with a strong analog orientation** in mind (very often using digital sources with analog methods while aiming at print publications). In contrast, NFDI4Culture wants to establish a professional network in close cooperation with all participating partners from research, GLAM and infrastructure institutions, that will **promote a cultural change and conscious approach to research data management and data-driven research in line with the FAIR principles for all phases of the research data life cycle in the cultural heritage domain**: from data capture and enrichment of digital cultural assets to data analysis with software tools and enhanced (data-)publications, standardized and sustainable solutions for long-term archiving and ethically as well as legally safeguarded options for reuse of digital cultural assets, related metadata and procedural data. To this end, the following most important objectives shall be reached:

1. **Provide broadly-usable, sustainable, long-term accessibility to cultural heritage data**: Preserving the already existing digitized collections and procedural research data on material and immaterial cultural assets (2D as well as 3D representations of material cultural objects, audio-visual representations of immaterial cultural objects) and the data collections that are generated anew on a daily basis and making them permanently accessible across academic disciplines for today's as well as tomorrow's researchers and the interested public is of utmost importance. This requires the development of subject-adequate archiving and publication solutions as well as operating models for the long-term accessibility and reuse of digital material and immaterial cultural assets. Whereas practicable solutions in the area of 2D representations exist but need to be connected, the sustainable archiving of complex data types like 3D digital representations or annotated audio-video representations is still experimental due to the lack of common standards. The consortium will evaluate, merge and certify existing solutions and develop reference implementations for hitherto untreated data types within a federated infrastructure that exposes an easy-to-use interface to the outside for the users.

2. **Improve findability, interoperability and reusability of cultural heritage data**: Due to the highly diversified landscape in cultural heritage research, a huge amount of highly relevant but hitherto disconnected data collections in varying degrees of quality and levels of standardization exists. This also holds true for a broad range of specialized software tools for organizing, annotating, analyzing and visualizing data of material and immaterial cultural assets. The consortium will put the improvement of findability, interoperability and reusability of data, tools and services at its center. Domain-specific as well as overarching data standardization to improve the interoperability of data and software tools will be treated within a dedicated task area. The findability of data within NFDI4Culture's federated infrastructure will be ensured through the development of an easy to use open registry for all data collections, tools and services. A knowledge graph and a collaborative terminology service will integrate the data managed by NFDI4Culture on the semantic level and

connect it to the data offerings of other NFDI consortia as well as beyond the NFDI to important information infrastructures such as the data hub of the European Open Science Cloud and Wikidata (Div. Authors (n. d.) 2020). The reusability of research data covered by NFDI4Culture is thus guaranteed.

3. **Boost professionalization and training in Data Literacy and Code Literacy**: Since research processes are increasingly data-driven, it is of crucial importance to be able to critically assess and reflect on the algorithms and software instruments with which the analyses are carried out and the results achieved in scientific and social contexts. Just as the ability to critically assess sources has been part of the repertoire of researchers in the humanities and cultural sciences for a long time, Data and Code Literacy must also be included in the canon as new competences today. Eight professorships with research focuses in Digital Humanities, Digital Musicology, Digital Art History, Media Science, Data Literacy and Digital Methods in combination with four DH study programs and DH centers (Köln, Mainz, Paderborn, Marburg) as well as professional trainers from the cultural economy take part in dedicated task areas and will build a highly efficient network of training offers (Culture Research Data Academy, CRDA) for students and researchers in higher education as well as for employees in the GLAM sector.

4. **Foster knowledge exchange and enable innovations through inward-outward cooperation**: Openness for collaboration, mutual assistance, sharing of knowledge and technology have been guiding principles of the NFDI4Culture initiative from the beginning and will be at the heart of all workflows and decision-making processes of the planned consortium. In May 2019, a comprehensive *Working Paper* (NFDI4Culture Working Paper 2019) was published that laid open core topics for mutual discussion and exchange between NFDI4Culture and the larger NFDI community. NFDI4Culture is also one of the initiators of the *Memorandum of Understanding* (NFDI 2019) of NFDI initiatives in the humanities. It is also a signatory of the *Berlin Declaration* (*Glöckner 2019*) of 21 consortia regarding close collaboration on cross-cutting topics in the NFDI. The development of a trans-disciplinary national research data infrastructure through complementary consortia will enable the application of new research methods. In the area of material and immaterial cultural heritage such methods may be automated mechanisms for the enrichment of digital representations, crowd-sourcing approaches to increase metadata coverage and quality, semantic reasoning and knowledge discovery on federated cultural collections or new approaches in computer vision and artificial intelligence that result in multimodally enhanced (data) publications. Besides the provision of a sustainable research data management infrastructure for its community of interest, NFDI4Culture also sees its mission in the development of a cross-disciplinary methodological framework for innovative digital research.

Composition of the consortium and its embedding in the community of interest

The planned consortium consists of a **geographically, thematically and institutionally balanced network** of 9 co-applicants, 11 academic societies and, currently, 51

participants. The composition aims to **ideally represent the broad spectrum of different actors in the cultural heritage domain** as described above. The co-applicants comprise **four universities** (UHD, UMR, UZK and UPB), **three infrastructure institutions** (FIZ, TIB and SLUB), Germany's largest institution in the **GLAM sector** (SPK) and an **Academy of Science** (AWLM) covering the long-term foundational research aspects with particular focus on musicology and the fine arts (Fig. 2).
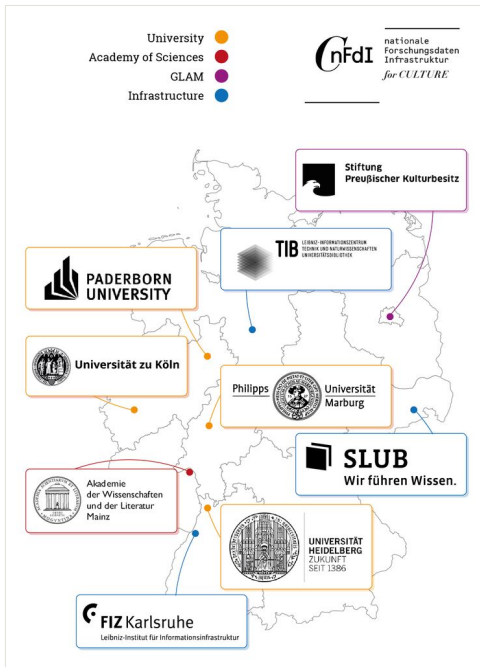

Figure 2. doi

Network of co-applicants

This group is **joined by 11 participating academic societies each representing one of the research domains** that together make up NFDI4Culture's community of interest: The Gesellschaft für Musikforschung (German Musicological Society), the International Association of Music Libraries, German section (IAML-DE), the Verband Deutscher Kunsthistoriker (Association of German Art Historians), the Gesellschaft für Medienwissenschaft (Society for Media Studies), the Gesellschaft für Theaterwissenschaft (Society for Theatre Studies), the Gesellschaft für Tanzforschung (Dance Research Society), the Bund deutscher Architekten (Association of German Architects), the Vereinigung der Landesdenkmalpfleger (Union of Regional Conservationists), the Rektorenkonferenz der deutschen Musikhochschulen and the Rektorenkonferenz der deutschen Kunsthochschulen (rector's conferences of the German universities of music and art) each of which represents 24 higher education institutions and the DHd Verband (Association for Digital Humanities in the German Speaking Area). **NFDI4Culture was initiated out of its research community by these very academic societies** and

conceptualized in close cooperation with the participating provider institutions. Therefore, NFDI4Culture is **highly embedded in its community of interest**. The academic societies will actively participate in the governance of the consortium and co-determine the operationalization of the planned work program together with the co-applicant institutions (Fig. 3).

Figure 3. doi
Participating academic societies.

Participating institutions

Akademie der Wissenschaften und der Literatur | Mainz (AWLM): With 35 long-term research projects (timeframe 12 to 24 years, currently reaching to the year 2040) and a special research focus in musicology and art history the Academy is one of the leading institutions in foundational research in the humanities, cultural studies and fine arts on a national as well as international level. A great number of highly skilled researchers work on musicological oeuvres including those of Beethoven, Brahms, Händel, Haydn, Schubert, Weber, Zimmermann, and others. AWLM also hosts long term art history projects dealing with rare and endangered cultural heritage objects from genres such as medieval stained glass. In view of the necessary administrative infrastructure for smoothly operating the consortium, the Academy provides long-standing expertise in the administration of funds and in scientific project management (currently the AWLM disburses funds for research undertakings in 11 federal states and also manages project funds from the DFG, BMBF and others). AWLM will provide NFDI4Culture with a highly efficient and goal-oriented

administration. In addition, AWLM has been a very active institution in the field of Digital Humanities and research data management for almost twenty years. The Digital Academy (DA), the Academy's research department for Digital Humanities, has a high impact on a national and international scale. The research activities of the DA focus on sustainable research software engineering and research data management in the humanities, current web technologies in cultural research contexts, and the application of Linked Open Data to open up new analysis and reuse scenarios. AWLM is one of the co-founding institutions of the DH master's program *Digital methods in the humanities and cultural studies* (JGU and HSM) and actively contributes with a DH professorship, regular lectures and international summer schools to the education and training of young academic professionals in a highly relevant area of research. With the dedicated support of the Ministry of Science, Qualification and Culture of the federal state of Rhineland-Palatinate, AWLM will found a center for digital musicological documentation within NFDI4Culture and commit to the consortium's administrative framework beyond the current NFDI funding period.

FIZ Karlsruhe – Leibniz Institut für Informationsinfrastruktur (FIZ): FIZ Karlsruhe makes significant contributions to the information infrastructure by supporting researchers in science, humanities and industry worldwide. FIZ curates and indexes very large amounts of patent information and research data from various sources. Nearly 300 employees develop and operate innovative information services and e-research solutions for precise research and intelligent analysis of these data. FIZ conducts applied research in close collaboration with academic and research organizations and acts as an experienced partner in national as well as European research projects. FIZ is engaged in several high-level working groups and committees dealing with information infrastructure and digital preservation aspects on a national and international level. The department e-Research (IEE) focusses on research data management, digital long-term archiving and virtual research environments. With a strong background in software engineering, it has participated in or led projects like National Hosting of Electronic Resources, RADAR (research data repository), German Newspaper Portal, TOPORAZ (digital space-time model for networked research) and Time Machine Europe. It oversees the operation and software development of the German Digital Library (DBB) and German Archives Portal. IEE brings extensive experience with software development in the Humanities and services such as RADAR to the consortium. Information Service Engineering (ISE) is a research department led by Prof. Harald Sack, covering Semantic Technologies, Knowledge Discovery, Ontological Engineering and Exploratory Search. ISE brings profound experience with the design, implementation and exploitation of ontologies, knowledge graphs, and Linked Data to the consortium. Intellectual property rights in distributed information infrastructures (IGR), led by Prof. Franziska Boehm, deals with copyright, IT (security) and data protection law on a German and EU level. Compliance with data protection and copyright laws are of high importance to the NFDI4Culture Community when it comes to the collection, storage and re-use of research data.

Sächsische Landesbibliothek – Staats- und Universitätsbibliothek Dresden (SLUB): With about 350 employees, SLUB is one of the largest and most efficient academic libraries in Germany. It serves its community both as one of the most important innovation

and coordination centers in the German library system, as well as providing a location for lively and intensive social and scientific exchange. With the help of forward-looking technologies, knowledge is produced, transparently networked, and made easily accessible. In its development of digital services (catalog, the digitization software Kitodo, etc.), the library consistently relies on open source products. SLUB coordinates the State Digitization Program for Science and Culture of the Free State of Saxony and the measures to safeguard Saxony's audiovisual heritage. In doing so, SLUB can draw upon the wide technical expertise of the Dresden Digitization Centre, and has, thus, been able to yield an average digitization volume of three million images in recent years. SLUB operates a productive long-term archive and possesses competence in LTA requirements for various media types. It is actively involved in the development of metadata standards (e.g. METS/ MODS, LIDO) and of standard data models. SLUB provides the research data support service for the Technical University Dresden (Excellence University) and is a founding member of the SaxFDM research data management network. With the Deutsche Fotothek (DF), SLUB operates one of the most important public image archives in Europe. With its Archive of Photographers, DF is committed to the preservation and publication of German photographic heritage. In its image database, DF publishes image media from around 100 partner institutions (research institutes, museums, libraries, archives). It also acts as an aggregator and content partner for DDB and Europeana. DF supports the expansion and design of research data infrastructure by offering innovative tools and services for image indexing, research, and analysis. Furthermore, it is home to DDB's Image/Photography Department, and also provides the spokesperson of AKBF (*Consortium for Art-Historical Picture and Photographic Libraries*). Founded in 1816, the Music Department of SLUB houses one of the most important collections of musical sources in Germany. Its holdings (textual sources and AV materials) are consistently directed towards the digital space. It holds a national RISM office and DDB's Sound Department. SLUB coordinates an average of 17 third-party funded projects per year with an average annual volume of €2 million. SLUB is tasked by the DFG with providing such services as the FIDs Kunst Fotografie Design (in cooperation with UHD/UB), and Musikwissenschaft (in cooperation with BSB). In doing so, SLUB offers the international research community key infrastructure components, such as the portal arthistoricum.net, musiconn.publish (subject repository), and musiconn.performance, a research data hub on cultural events.

Stiftung Preußischer Kulturbesitz (SPK): The Prussian Cultural Heritage Foundation is an internationally renowned cultural institution and an important player in the humanities and the social sciences. It includes museums, libraries, archives, and research institutes. Its collections have a universal character and document the evolution of human culture from its beginnings to the present in Europe and on other continents. The Music Department of SBB (one of the institutions of SPK participating in the consortium), for example, holds the largest music collection in Germany and includes autographs and manuscript copies of many composers such as Bach, Mozart, and Beethoven. An unmistakable influence on the Foundation's profile is the combination of art and culture with science and research. SPK supports digital research in many ways: Third-party funded projects with universities and non-university institutions stand alongside the machine-readable indexing of their collections and digitization with various procedures to

provide access for the specific requirements of researchers and the general public. Increasingly, data enrichment methods, which are also developed or applied within the Foundation, as well as data analysis methods are becoming more important. The organizational reorientation of SPK was manifested by appointing Chief Digital Officers by the establishment of a dedicated team in 2017. In NFDI4Culture, SPK and its institutions will contribute as an infrastructure provider between culture, research and science. Through a large number of national and international projects, it has rich experience in the organization and design of collaborative undertakings, especially in the object-oriented research field of text, image and moving image, physical and audio artifact. SPK is motivated by the opportunity to work closer with the NFDI4Culture research communities and to translate their requirements into better and transferable services. At the same time, there is an opportunity to better communicate and interlink the (basic) research carried out within the foundation and to contribute this potential to the consortium as a whole.

Technische Informationsbibliothek Hannover (TIB): The German National Library for Science and Technology (TIB) represents and operates a national research infrastructure facility for the provision of scientific information. TIB preserves and organizes information, data and knowledge in its target domains, such as architecture, and provides direct access to these large-scale information spaces through digital services, irrespective of time and place. With its vast collections and innovative services, TIB aims to support the complete lifecycle of research and the digitization of science and technology in general. As the world's largest specialized information center in its fields, TIB has outstanding expertise in developing, managing and preserving knowledge, particularly in key areas such as grey literature, big research data, vocabularies and ontologies, films and 3D objects, as well as patents and standards. Some examples of key services include the TIB portal giving access to more that 100 Million documents and research artifacts, the audio-visual portal AV-Portal comprising more than 15.000 scientific videos, license negotiation, Open Access and Digital Preservation offers, the research data management software Leibniz Data Manager or the collaborative OpenCourseWare authoring platform SlideWiki. In 2009 TIB founded the DataCite association with currently more than 130 international member organizations and provides the DataCite headquarters. In close cooperation with L3S research center from Leibniz University Hannover, TIB performs world-class research aiming to advance information, data and knowledge sharing in the digital age, for example with its Open Research Knowledge Graph. Research and development at TIB include Visual Analytics, Data Science, Scientific Data Management, Open Science, and Non-textual materials. TIB is increasingly active in the field of open culture, both in third-party funded projects and in community building. Partners are Europeana, Time Machine, Wikimedia, and other communities dedicated to free knowledge.

Universität Heidelberg (UHD): Founded in 1386 Heidelberg University is Germany's oldest university. It is a comprehensive university with 12 faculties covering a broad spectrum of subjects and one of Europe's strongest research universities with international reputation. Its success in all funding lines of the DFG's Excellence Initiative and in international rankings underscores its leading role and excellent reputation in the scientific landscape. As one of the university's central infrastructure units, the Heidelberg University

Library (UHD/UB) has developed electronic information infrastructures in pilot projects over the past two decades that have set standards for open science. These projects range from state-of-the-art digitization technologies and workflows, to innovative electronic publishing platforms such as Heidelberg University Publishing (heiUP), which publishes high-quality research work in several digital formats in Open Access, and also offers comprehensive research data management services. Within the framework of its FIDs Kunst, Fotografie, Design and Altertumswissenschaften, which are supported by the DFG, UHD/UB has for many years been taking on central tasks in the field of value-added infrastructure services for the art and classical studies community, in addition to the nationwide provision of literature and information throughout Germany. The focus here is on electronic publishing in Open Access, with innovative design of traditional formats playing a major role. In addition to aspects of informal communication and collaborative work, this also includes linking to other sources and integrating multimedia content or related research data. For this purpose, dynamic and collaborative publication services have been developed. Future cooperation within the NFDI will serve to strengthen these existing and well-used services within the specialist community, to further develop them in line with demand and to make them usable beyond the boundaries of the subject areas supported in Heidelberg. UHD focuses on electronic publishing (including the archiving of research data) and the development of workflows, business models and best-practice offers based on international standards and has great networking potential due to the high likelihood of subsequent use by other specialist communities. The predominant use of standard software under open source licenses, combined with the strong demand for Heidelberg's infrastructure offers by scientists, results in a broad range of opportunities for the use of developed services and technologies, both by subject-related disciplines within NFDI4Culture, and in other sciences and their NFDI consortia.

Universität zu Köln (UZK): The Faculty of Arts and Humanities of the University of Cologne stands, as one of the largest teaching and research institutions of the humanities in Europe, for outstanding diversity of topics in modern research and teaching. There is a lot of expertise in digital projects and research data management in many disciplines of the Faculty, however, especially in the area of art history and image/object-based research. Its image archive prometheus (UZK/prom) is very well established in communities that deal with objects of material cultural heritage. Prometheus covers the needs of these communities by providing research material through making available over 2 million objects from the fields of art history, archeology, architectural history, egyptology, theology, diplomatics, philosophy, ethnology and many more. More than 100 image collections are integrated from various institutions like museums, university institutes, research facilities, libraries and archives from all over Europe. UZK itself contributes five collections to prometheus (collections: Archeology, History, Art History, Media Culture and Theatre and the University Library). In addition to the integration and provision of the research material, prometheus develops and provides tools for image-based research and serves as a consulting service for projects, collections of image-based disciplines and sustainable software development. From 2001 until 2004 UZK/prom was funded by the BMBF. After transitional financing made possible by UZK, it became independently financed through its own license model in 2008. More than 160 institutions of the NFDI4Culture community are

using prometheus and there are currently also 8,000 active personalized accounts. In addition, the faculty provides a research-oriented data center (DCH) that addresses the needs of researchers in the dedicated humanities research data management. The Cologne Center for eHumanities (CCeH) is a member-supported teaching and research center at the Faculty of Arts and Humanities. In this capacity, UZK represents the interests in research data management of a large number of researchers who deal with objects of material and immaterial cultural heritage.

Philipps-Universität Marburg (UMR): UMR covers a wide range of disciplines and has a long tradition in cultural science, e.g. in art history and media studies. Its comprehensive research data infrastructure is well recognized for its cooperation with other research institutes and GLAMs, forming an ecosystem for digital research data preservation and accessibility with a notable focus on cultural assets. UMR has founded a local service center for e-research to support its scientists in all aspects of digitally assisted research. It has the lead within the joint project Hessische Forschungsdateninfrastrukturen (Hessian Research Data Infrastructure) which fosters a coordinated Hessian-wide research data support. With the projects FOKUS and TRUST, UMR has built up major competencies in promoting Data and Code Literacy. TRUST gained one out of five prizes in a nation-wide contest for digital skill development. UMR is one of 15 universities within the Data Literacy Education network driven by the Stifterverband. The German Documentation Center for Art History (DDK) is an internationally acclaimed research and service institute for art history and related disciplines. Its mission is to collect, index and make available image materials related to European art and architecture as well as to conduct research on the history, practice and theory of how visual cultural assets are passed on. It is further actively involved in designing workflow standards for the international museum community. For more than 30 years, DDK has been enabling and supporting the digitization of the inventory data of about 100 partner institutions (museums, offices for the protection of historic monuments, libraries, archives and research institutes) along with its own scholarly photographic collection. For scientific reuse, it publishes the data in its image databases and on other (research) web platforms. DDK has explored the application of standards and data quality management in dozens of projects and has provided expertise and recommendations for a sustainable data management. In current projects, DDK develops a systematic, FAIR-based quality assurance management of structured research data on objects of material culture. It also engages in the adaption of the GND (DNB) to meet the requirements of non-librarian cultural heritage and humanities communities. In the NFDI it focuses on standards and interoperability of data from various sources, as well as on facilitated use and reuse. In addition, UMR is establishing the Marburg Center for Digital Culture and Infrastructure (MCDCI) that combines scientific expertise, e-research infrastructure, local cooperation in digital humanities and competencies in cultural data studies. Cooperative professorships with the AWLM and strategically aligned professorships from UMR form the center of the digital humanities cluster.

Universität Paderborn (UPB): Paderborn University is the university for the information society. The strong foundation in computer science and its application, as well the importance of IT for a growing number of disciplines are the pillars for this claim. In order to

contribute to the scientific and technological development of the information society and to critically reflect these developments by taking into account the history, norms, and values of society, it complements the spectrum of "hard" sciences with the arts and humanities and to successfully learn from each other. Its mission includes a strong international and cultural presence, since the information society is decidedly global and should not remain a purely intellectual undertaking. In 2014, Digital Humanities were defined as one of the five key research areas of the University (with several interdisciplinary research institutes and many corresponding project activities). In 2019, UPB established the master study program Digital Humanities. For 15 years, the Musicological Department of UPB has been developing software tools for musical editing. Since then, the so-called "Edirom tools" have been used in many editorial projects. 10 years ago, the Edirom-Summer-School was established in order to promote the use of international standards like TEI and MEI and to provide students and participants from the academic or cultural community with all necessary skills around X-technologies and digital editions software. In 2014, the BMBF-Center for Music, Edition, Media (ZenMEM) was founded which is active in the field of tool-development and critical assessment of new tendencies in software development for the humanities. At the same time, ZenMEM is an important consultant for musicological projects and acts as an agency service for training and education. ZenMEM is the central contact point for digital musicology in Germany. Members are involved in the international TEI and MEI community and in developing solutions relevant to scholars beyond national and disciplinary borders. Besides its musicological research and services, UPB places a strong focus on cultural heritage in general. The competence center Cultural Heritage: Material, Immaterial, Digital serves as a central agency for the coordination of research in this area and provides education and services around methods in Digital Humanities. Since 2006, the center serves the Paderborner Bildarchiv with images of e.g. art, architecture, urban development, and sculptures which is searchable through prometheus.

To ensure optimal acceptance by its community of interest, NFDI4Culture integrates a broad spectrum of participants from all areas of the cultural heritage sector. With regard to the highly diversified research landscape already described above **this is the only approach that will lead to successful, demand-oriented service offers and real user participation**. For this reason, the scope of participants that contribute to NFDI4Culture's work program ranges from subject-specific international organizations such as RISM (musicology, working groups in 35 countries) or CVMA (art history, 17 member countries), international research institutions like the Hertziana (Rome) and DFK Paris to national GLAM institutions like DNB or GNM, smaller research institutions with high impact such as the Beethoven Haus Bonn (BHB) to universities (such as FAU, LMU and UdK) and archives from the public media sector such as the Deutsches Rundfunkarchiv (DRA). Besides this, NFDI4Culture has the support of numerous governmental and non-governmental organizations such as the **German Music Council** (Deutscher Musikrat), the German Commission for **UNESCO**, the **Union of the German Academies of Sciences and Humanities** as well as international organizations such as, **ORCID, DataCite**, the **Open Knowledge Foundation** and **Wikimedia**.

Analysis of the existing research and infrastructure environment and development of the planned consortium according to the users' needs

The envisaged research data infrastructure is based on a **thorough analysis of the existing research landscape**. At the time of writing, 76 relevant data collections with millions of records have been identified (with widely differing levels of standardization and quality assurance), that are already curated at one of the co-applicant or participant institutions. Five long-term repository and publication solutions for 2D, 3D and audio-visual cultural assets exist that are operated through the consortium partners. Regarding storage and publication infrastructures, many good solutions already exist in the area of 2D digital representations, documents and metadata, whereas the area of 3D digital representations and multimodal publications will need reference implementations that pave the way for standardization and interoperability. There are 51 research software tools by consortium members and 35 subject-specific information infrastructures (24 written or operated by members of NFDI4Culture) that need to be considered when dealing with research data of tangible and intangible cultural assets. This clearly demonstrates the broad technological potential on the one hand but also the great challenge on the other because the software quality and sustainability of those technical components is not yet subject to a defined process of quality assurance and monitoring. Solutions for this will be developed in a dedicated task area of the work program and in close cooperation with similar approaches in other NFDI consortia. With regard to teaching materials, the planned consortium can already fall back on a solid basis of existing training offers (33 by consortium members, incl. professorships). Nevertheless, coordinated teaching approaches to Data and Code Literacy (especially in the field of train-the-trainer offers) are still missing and will be developed through dedicated measures in the work program.

Means of communication and collaborative development of the consortium

The development of NFDI4Culture took place in a two-year collaboration between users and providers in the research domain of material and immaterial cultural heritage. The overarching needs and challenges of "Research-driven infrastructures in the humanities" were discussed early on in the NFDI process between representatives of academic associations, universities, memory institutions and infrastructure providers during three workshops conducted by CLARIN, DARIAH, the Union of the German Academies of Sciences and Humanities and the DHd Association in 2018. The workshops clearly showed that **research data in the humanities varies greatly** in terms of quantity, quality, dynamism, level of standardization, degree of abstraction, legal and ethical circumstances, and subject specific data types. Therefore, the **demands of the humanities in the NFDI can be best met** with a coordinated effort that takes a disciplinary as well as methodological approach. To reach this goal, a **network of cooperating NFDI initiatives with subject-specific, distinct areas of responsibility and complementary fields of action** was formed.

Through the **extensive national and international network of the participating academic societies**, NFDI4Culture has a very strong embedding in its subject-specific community of interest and **ideal means of communication between users and**

providers. The needs and demands of the users were gathered during dedicated workshops on research data management in art history, musicology and related disciplines in 2018. These workshops paved the way for the foundation of the consortium initiative in December 2018. **Integration, collaboration, networking, knowledge exchange and openness have been the guiding principles** from the beginning. With regard to networking and collaboration, NFDI4Culture has been a regular participant and active contributor to the NFDI Forum (March, June 2019) of the DHd Association (which will also dispatch a delegate to NFDI4Cultures' steering board and ensure tight relations to academic associations engaged in other NFDI consortia via the NFDI Forum of the DHd). NFDI4Culture also initiated several networking workshops of NFDI initiatives in the humanities (April, June 2019) and is one of the signatories of the *Memorandum of Understanding* of NFDI initiatives that has been the result of these workshops.

For the first time a coordinated initiative of this scale takes place in the field of material and immaterial cultural assets. Since 2018 it **has already created a very positive impact** on **and gathered strong acceptance** in the participating disciplines. Because NFDI4Culture follows a new and innovative approach for its area of concern, care has been taken to test the effectiveness and adequacy of the plans at a very early stage. **From February to October 2019 the consortium initiative tested its planned workflows in a proto-operational phase**. One example for this is the jointly founded **NFDI4Culture Coordination Office** (NCCO; 3 FTE provided by the (co-)applicants). The staff of the NCCO has been working closely with the designated spokesperson, co-spokespersons and representatives from the scientific communities and associations. A transparent and iterative process for participation was put into practice based on a **large-scale survey (May to August 2019)** that continuously collected, evaluated and structured information and feedback about existing infrastructure components, tools and services from the community of interest (Fig. 4).

The NCCO oversaw the publication of the **NFDI4Culture Working Paper** for the NFDI conference of the DFG **in May 2019**. Many of the topics introduced by NFDI4Culture's perspective (such as Data and Code Literacy, sustainable research software engineering, standardization and quality assurance and development of authority data) have become essential cross-cutting concerns for the NFDI at large (also cf. the *Berlin Declaration* from September 2019).

To make sure that the work program of the consortium really meets all the requirements of NFDI4Culture's community of interest, a **Culture Community Workshop** (CCW) was conducted in September 2019 during which the **work program with all measures was discussed openly and validated by representatives from the user community as well as from cooperating NFDI consortia** (Text+, NFDI4Memory, NFDI4Objects). The workshop also served as a test for the future yearly Culture Community Plenary (CCP) that will operate with the same methods of user participation. The extensive feedback collected during the CCW was integrated into the finalization phase of the work program.

Figure 4. doi

Development of the consortium.
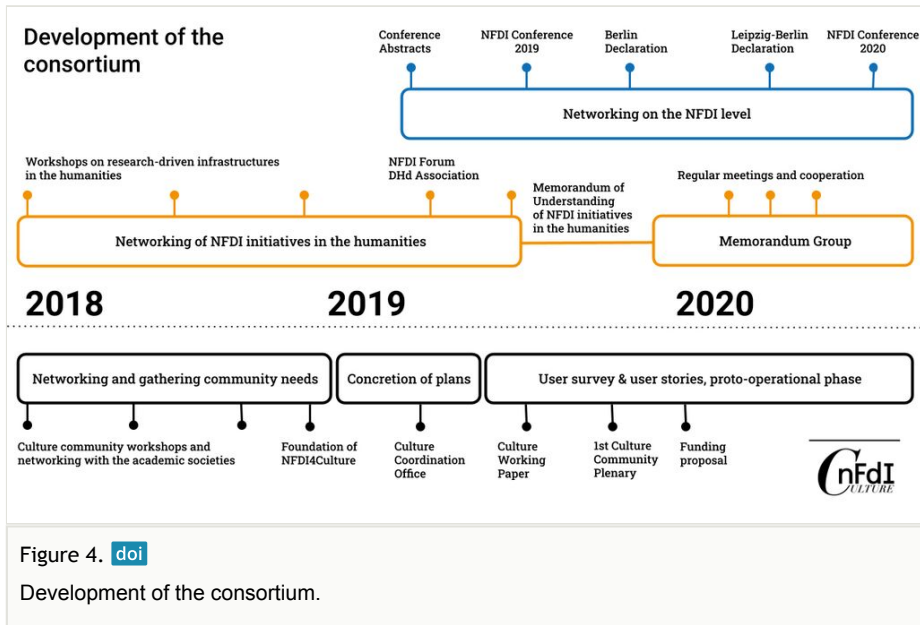
Interaction between users and providers

Much care has been taken to **create structures within the consortium that are evenly balanced between users and providers**. This starts with the composition of the group of co-applicants comprising four universities with particular research focuses (art history, musicology, theatre studies etc.), four infrastructure facilities (one from the GLAM sector) and one academy of science providing research as well as infrastructure. This also holds true for the designated (co-)spokespersons, each of which has long-standing **expertise with regard to scientific infrastructures** and is also an **expert researcher** in one of the subject domains covered by NFDI4Culture. The consortium's work program also provides for a wide range of participation opportunities for users. **Each of the six task areas will establish a specialized forum** in which users from the community of interest will closely and continuously work on the central topics together with experts from the provider institutions. Forums are organized and managed by the respective co-spokespersons of each task area. They do not only foster knowledge exchange – **they are long-term participatory structures** that will also **produce guidelines and recommendations** to the steering board and effectively **mirror the changing needs** of the user community. Forums will also **provide important monitoring functions** to the operationalization of the work program. Further interaction between users and NFDI4Culture also takes place on the level of consultation. Each of the task areas will implement measures for setting up specialized **helpdesks** (i.e. regarding digitization, data standards, sustainable software, archiving and publication, legal and ethical issues and training). The NFDI4Culture **information portal will provide a single point of contact** for the users from where inquiries and feedback are directed to the responsible helpdesk. NFDI4Culture plans two coordination offices, one administrative and one technical, that will oversee that the

inquiries are processed, documented and replied to in an appropriate time frame. The **Culture Community Plenary** will be a focal point of interaction between users and providers (cf. chap. Organizational Structure and Viability). During this two day event (planned each autumn) the NFDI4Culture Community and all organs of the governance body (the steering board, the advisory council, the task teams and the forum working groups) come together to discuss the current state of the consortium's work and decide on ideas and plans for the next year. This will also be the point where proposals for candidates and elections for seats in the governance bodies will take place and where new members may join the consortium and meet the community.

The consortium within the NFDI

Cooperation with other NFDI consortia

NFDI4Culture has established a **continuous and productive dialogue** with other consortia to **foster inward–outward cooperation in the NFDI**. NFDI4Culture's **work program has been discussed openly with representatives from other consortia** during the Culture Community Workshop and **advice** from this feedback process **has been integrated** during the finalization phase of the proposal. In the humanities, the consortium is one of the initiators and signatories of the joint *Memorandum of Understanding*. The other three signatories are the initiatives **NFDI4Objects, NFDI4Memory and Text**+. The *Memorandum* defines the responsibilities and modes of cooperation between the four partners with the **aim to create a meshed structure** with a complementary division of tasks that best addresses the diverse needs of the highly differentiated research landscape in the humanities. Additionally, NFDI4Culture is one of the 21 signatories of the *Berlin Declaration* on cross-cutting topics in the NFDI (Glöckner 2019). The consortium closely embraces the spirit of cooperation and the goals of the *Declaration* and will actively contribute to the cooperative treatment of topics such as metadata harmonization, provenance concepts, interoperability across research domains, legal and ethical aspects, terminologies and linked data services, and the development of training concepts for Data and Code Literacy.

NFDI4Culture **has planned an appropriate budget for cross area working groups as agreed upon in the *Berlin Declaration*** to allow for close collaboration with other consortia with regard to the cross-cutting topics (cf. TA7|M4). It is to be expected that it will take at least two to three years until the general framework of the NFDI is established. Nevertheless, there are already concrete arrangements between NFDI4Culture and several other cooperating consortia for collaboration on cross area concerns. The following list is not exhaustive and will be expanded as the NFDI unfolds: **MaRDI** (mathematics) and NFDI4Culture will collaborate in the area of modeling and standardization of research data. Both consortia engage in complex data types. In mathematics, these are purely ideal objects such as elliptic curves, PDE-based models or theorems/proofs. In research domains such as art history, musicology or media studies, these are cultural assets with their material, immaterial, conceptual, discursive and reception-historical properties. Both consortia will focus on symbolic object representations and ontology-based metadata

approaches. On the basis of this cooperation, MARDI and NFDI4Culture hope for good coverage of these research data categories and aim to evolve their work into a generic standard that is also useful to other consortia and the NFDI in general. In the area of training and qualification, **NFDI4Ing** and NFDI4Culture have identified Data Literacy, Code Literacy, and the provision of open educational resources as cross-cutting topics for close collaboration. Both consortia also intend to cooperate in the area of standardization and curation of 3D data types (e.g., CAAD models and other forms of 3D representations). Since image data plays an important part in **NFDI4Neuroscience** and NFDI4Culture, both consortia plan to cooperate in the standardization of image metadata and image formats and in the joint development of image analysis tools in the field of computer vision. **Text+** and NFDI4Culture intend to explore data standardization potentials at the interface between musicology and philology (taking MEI and TEI as a starting point).

Members participating in other NFDI consortia

Members of NFDI4Culture participate in the following NFDI consortia and commit to strengthening the institutional network within the NFDI across subject areas and research domains:

- **Astro-NFDI**: UMR (participant), TIB (participant), UZK (participant)
- **BERD@NFDI**: UZK (co-applicant)
- **DAPHNE**: TIB (participant)
- **FAIRMat**: UPB (participant), TIB (participant)
- **ForumX**: UZK (co-applicant),
- **MaRDI**: FIZ (co-applicant)
- **NFDI4Agri**: FIZ (co-applicant)
- **NFDI4Biodiversity**: UMR (co-applicant), SUBG (participant)
- **NFDI4Chem**: FIZ (co-applicant), TIB (co-applicant), UZK (participant)
- **NFDI4Earth**: TIB (participant), UZK (participant)
- **NFDI4Health**: UZK (co-applicant/participant)
- **NFDI4Ing**: TIB (co-applicant), SLUB (participant)
- **NFDI4Memory**: FIZ (co-applicant), HI (co-applicant)
- **NFDI4MSE**: FIZ (co-applicant)
- **NFDI4MobilTech**: FIZ (co-applicant), SLUB (participant), TIB (participant)
- **NFDI4Neuroscience**: UMR (participant)
- **NFDI4Objects**: SPK (co-applicant), UZK (participant)
- **PAHN-PaN**: TIB (participant), UZK (co-applicant)
- **Text+**: AWLM (participant), DNB (co-applicant), SUBG (co-applicant), UPB (participant), UZK (participant)

Contributions to the NFDI as a whole

NFDI4Culture already has solutions and experience in regard to several cross-cutting topics that could be of interest to other consortia and the NFDI as a whole. Its partners from archives, libraries and museums have long-standing expertise in the field of

metadata, the development and sustainable curation of authority data and data harmonization workflows. NFDI4Culture deals with complex data types (that consist of a combination of basic data types like semantically annotated 3D reconstructions of historic buildings, multimodal works of art like music, theatre or dance performances consisting of annotated video and/or audio streams, etc.). The consortium will develop standardization recommendations and reference implementations for **long-term archiving and publication solutions for complex data types** and will be glad to exchange knowledge and technologies with other consortia. The partners of NFDI4Culture are very much aware that the field of **rights management and ethical advice** for research data can only be solved in close cooperation on the NFDI level. Since legal expertise and data ethics are of primary importance for research in the field of cultural heritage as well as the NFDI in general, NFDI4Culture has planned dedicated measures in its work program, e.g. the establishment of a **legal helpdesk and an advisory panel on data ethics** (cf. TA5|M4, TA5|M5). It will also initiate a stakeholder process for the development of a legal framework for research data (cf. TA5|M5). These measures could also provide expertise to the NFDI as a whole.

Together with the participating universities and in exchange with other NFDI consortia, NFDI4Culture plans to implement **training and education** programs in Data and Code Literacy, two skills that will become indispensable for future research in any subject. NFDI4Culture also has a strong grounding in the field of **linked data and the semantic web**. Concepts for the **collaborative curation of terminologies** and the active onboarding and engagement of users from research as well as from citizen science and the **joint creation and community-driven curation of an overarching knowledge graph** on tangible and intangible cultural assets have been developed together with **Wikimedia** and could be an interesting methodological approach for other NFDI consortia as well. On a more general level, due to the close connection of its research domain with cultural politics and the cultural economy, NFDI4Culture intends to contribute the results and experiences with regard to the **development of adequate models for public/private partnerships** in the NFDI. The consortium can also share knowledge and best practices from its experience in integrating a diversified landscape of disparate research units into a consortium's network.

Expectations from the NFDI

NFDI4Culture defines **three levels of action in regard to cross-cutting topics**:

1. at the **level of the consortium** itself, pooling user needs from different research communities covered by the consortium,
2. at the **level of cooperation between several consortia**, sharing common methodological or technological expertise for the benefit of a larger user base and
3. generic solutions at the **level of the NFDI as a whole**.

With this differentiation in mind, NFDI4Culture has identified the following relevant topics that are aligned to the findings from the NFDI conference, the assessment of the NFDI

expert committee and the results of the DFG's NFDI governance workshop in August 2019 that need to be decided within the NFDI's collaborative framework:

- A **common legal framework** as well as a single legal entity for the NFDI to which consortia can join as members **while allowing to keep** their **internal structures** that have been carefully considered with respect to their community of interest.
- Joint development of a **Research Data Commons** as outlined in the *Berlin Declaration*. This cross-consortia infrastructure should include NFDI-wide solutions for a **common authentication and authorization infrastructure** (AAI) which must include a **flexible access management** for data sets that are subject to particular legal and ethical obligations. It should include solutions for **persistent identifier management** (PIDs).
- Joint **data federation solutions and federated discovery systems** across NFDI consortia and close collaboration in the area of **terminology services** and linked data.
- Joint development and exchange of experiences regarding **adequate operating models (especially for long-term operation, archiving and data publication)**. This also includes a **common NFDI certification scheme** for the services offered by the NFDI.
- Joint development of **new reputation schemes and academic credit systems** that will support a cultural change towards new forms of academic contributions and achievements (like data publications or research software).
- Joint development of subject specific as well as general **criteria for the sustainability and long-term provision of research software**.

International networking

NFDI4Culture is tightly interwoven into a broad international network of research communities, initiatives, institutions and infrastructures. **International interaction on the user level takes place through the networks of the participating academic societies**. The VDK is a member of the CIHA (Comité International d'Histoire de l'Art) and has good connections to the Swiss and Austrian associations for art historians. The GfTh closely cooperates with the European Association of the Study of Theatre and Performance (EASTAP), the International Federation for Theatre Research (IFTR/FIRT) and with Performance Studies International (PSi) as well as with the International Society of Libraries, Archives and Documentation Centers of the Performing Arts (SIBMAS). The GfM and its Fachgruppe Digitale Musikwissenschaft (Digital Musicology) has members from many different countries. In the Fachgruppe Freie Forschungsinstitute 56 institutions are represented, including the RISM Zentralredaktion and the internationally outstanding Edirom network. The GfMe is closely networked with the European Network for Cinema and Media Studies. In addition, there are working contacts to numerous other non-German professional associations such as SCMS (USA), Affecav (FR), AIM (PT), BAFTSS, MeCCSA (GB). The BdA is a member of the Architects Council of Europe (ACE). In addition, the BdA Federal Association together with its Polish sister association SARP annually organizes a German-Polish competition for up-and-coming architects. The DHd

association also has a strong international network. On the one hand through the language-bound definition of its area of responsibility, which per se includes not only Germany, but also Austria and (German-speaking) Switzerland. DHd is an associate organization in the European Association for Digital Humanities (EADH), which in turn is a constituent organization in the Alliance of Digital Humanities Organizations (ADHO).

International interaction also takes place on the participant level of the consortium. Several institutions have joined NFDI4Culture, each with an outstanding international impact, for example the KHI Florenz (Italy), the Hertziana (Italy), the DFK Paris (France) or RISM-CH (Switzerland). NFDI4Culture also has **participants that provide subject-specific, large-scale international research infrastructures**. For the field of musicology **RISM**, initiated and accompanied to this day by IAML and IMS, can be named as an example with its working groups in 35 countries across the globe. In the field of art history, there are scientific connections to 17 countries through the **CVMA**, while the network **PHAROS** (members from NFDI4Culture are UMR/DDK and the Hertziana), the International Consortium of Photo Archives, creates a digital research platform allowing for comprehensive consolidated access to photo archive images and their associated metadata worldwide. The **prometheus** image archive (UZK/prom) consists of an international consortium with 162 institutions, 35 of which are located in Europe and the US and already integrates up to 10% of research data from other European Countries. NFDI4Culture also interacts with international academic networks such as the **International Association of Music Libraries** (IAML, the German chapter is a participant of the consortium), the International Musicological Society (IMS) and the **European Federation of Academies of Sciences and Humanities** (ALLEA), representing more than 50 academies from over 40 countries in Europe. With AGATE (A European Science Academies Gateway for the Humanities and Social Sciences) the AWLM develops and operates a digital research information system with a European perspective that can connect information from foundational SSH research to the consortium.

Members of NFDI4Culture engage in numerous international standardization bodies for research data on material and immaterial cultural assets and will ensure mutual transfer of knowledge and work results between the consortium and those initiatives. Examples are the **MEI** standard (via the GfM, UPB, and AWLM), the **LIDO** standard, the **CIDOC CRM**, the **Iconclass** classification system and **IIIF** (via UMR/DDK, SUBG and TIB). Members of NFDI4Culture further engage in outstanding European GLAM initiatives such as Europeana, contributing to an open, knowledgeable and creative society. **Key providers in the field of international research** have expressed their strong interest for collaboration and their support for NFDI4Culture. Institutionally, this ranges from the **Getty Research Institute**, to **subject-specific research data initiatives** such as the **MEI community** and **Pelagios Commons**. Strong ties exist with important high-level international information infrastructure providers such as **ORCID**, **DataCite** and large-scale data-driven research projects such as **Time Machine** and international bodies such as the **Open Knowledge Foundation** and **Wikimedia**.

Organizational structure and viability

At the time of writing, the consortium consists of the nine co-applicants, 11 participating academic societies and 51 participants from universities, research institutions, GLAM institutions, and cultural heritage initiatives and associations. Together they represent the full spectrum of the diversified research landscape with regard to material and immaterial cultural heritage which is crucial for ensuring the acceptance of the consortium's services in its community of interest. All participants have expressed their explicit commitment and contributions to the consortium in letters of commitment.

NFDI4Culture implements four guiding principles:

1.   **lean decision-making**,
2.   **transparent communication** and **easy participation**,
3.   promotion of **innovative solutions** for the needs of users and providers, and
4.   **rapid adaptation of operationalization strategies** in response to (changing) needs in the community of interest.

These principles also serve as the four success criteria for the periodic evaluation of the governance that will take place every second year (for the first funding period: autumn 2022 and autumn 2024).

The planned internal structure of the **consortium consists of three levels**: academic societies, participants and co-applicants together form the **community level** with its **central representative organ**, the yearly **Culture Community Plenary**. On the operative level of the consortium, measures from the work program are realized by members of the community (co-applicants and participants). The **operative level provides participatory structures** for the community that are **anchored in each task area** (forums and task teams), **consultation structures for the users** of the consortiums' services (specialized helpdesks, consultation agencies, clearing agencies), and **overarching coordination structures** (the administrative and technical coordination offices) for the smooth running of the consortium. The **governance level** consists of three organs: the **Culture Steering Board** acts as the central body in all decision making processes of the consortium, the **Culture Spokesperson Committee** serves as the executive organ that oversees and takes responsibility for the proper and timely realization of the measures of the work program and the **Culture Advisory Council** that serves as an independent body that gives advice, assists with proposals for the resolution of controversial issues and connects NFDI4Culture's governance level to the governance level of other NFDI consortia and the NFDI as a whole (Fig. 5).

Communication, cooperation, responsibilities and decision-making

Culture Community Plenary (CCP): The CCP is a two-day collaboration and decision-making event that will take place each autumn. During this occasion, all members of the consortium (representatives of the co-applicants, participants and the academic societies) and all organs of governance (the steering board, the advisory council and the

spokesperson committee) and invited guests from national and international research and infrastructure come together to discuss the current state of the consortium, present results of finished tasks, validate the progress of the work program and shape the consortium's strategy for the next year. During the CCP, members of the community can present and submit proposals to the steering board for the realization of topics covered by one of the dedicated budgets that are planned for the enhancement of NFDI4Culture's service offerings (e.g. for data curation and rescue, tool and data service improvements, development of training offers and the support of research projects that apply for third party funding). Proposals always have to be substantiated with a reasonable need analysis and have to be prepared collaboratively between users and providers in the forums of NFDI4Culture.
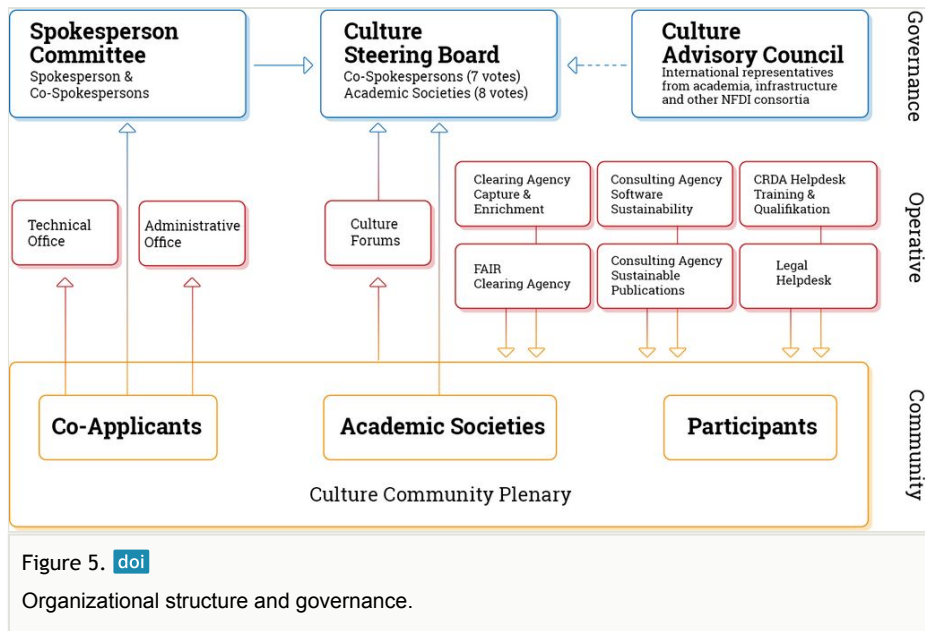

Figure 5. doi

Organizational structure and governance.

Culture Forums (CF): Forums are the key participatory structure of NFDI4Culture. Each of the six task areas provides a forum that is managed by the responsible co-spokesperson and the staff from the administrative and technical coordination offices. Forums are open structures in which providers (co-applicants, participants) and users from research work together on specific topics in each task area (data capture, development of standards, software sustainability, long-term archiving, rights and ethics, training offers etc.). Forums do not only provide a continuous knowledge exchange, they also fulfill a monitoring function and produce substantial and measurable output (e.g. in the form of guidelines, proposals for fund usage etc.).

Helpdesks, consultation agencies, clearing agencies: To guarantee optimal service brokerage and provide easy orientation to the users of NFDI4Culture, each task area will set up a subject-specific consultation structure which will be the starting point for inquiries. The analysis of the users' demands has shown that this content-driven, federated

approach will suit the needs of NFDI4Cultures's community of interest much better than a one-fits-all solution. Central consultation services of the consortium therefore consist in two clearing positions (for data capture and enrichment and implementation of FAIR data), two consultation agencies (for research tools and long-term data archiving and publication) and two dedicated helpdesks (for legal aspects and training offers).

Administrative and Technical Coordination Offices (ACO / TCO): To ensure the smooth operation of the consortium with regard to all administrative issues (reporting, contract management, financial management, documentation, etc.) and to provide an overarching and long-term technical perspective that reaches beyond single measures in the work program the consortium will set up two offices: the administrative and the technical coordination office. The staff of the two offices closely works together with the task area co spokespersons and the Spokesperson Committee. Both offices operate as distributed coordination structures and each of the co-applicant institutions has dedicated staff in both offices. The ACO lead rests with AWLM and the TCO lead rests with FIZ. ACO and TCO set up a tight coordination network for the consortium using weekly stand-ups moderated by the ACO and the TCO leads and report progress to the Spokesperson Committee at least once a month. This has proven to be very successful during the application phase with the NFDI4Culture coordination office (NCCO) being a prototype of this structure.

Culture Steering Board (CSB): The CSB is the central decision-making organ of NFDI4Culture. It consists of fifteen persons, six of which are co-spokespersons (representing the task areas and the providers), eight of which are delegates of the academic societies (representing the research domains and the users) and the consortium's spokesperson (representing the administration and mediation). The users have a majority of votes in the CSB. All decisions are taken by single majorities. The delegates (excluding the spokesperson) of the CSB perform their membership in the CSB for two years (in alignment to the planned evaluation of the governance structure in 2022 and 2024) after which a re-election (and possible restructuring) of the CSB can take place. For each period of office, the CSB elects a chairperson and gives itself rules of procedure. The spokesperson (which is a constant member of the CSB) has the responsibility to serve as a facilitator during all decision-making processes. He/she also has a vote in the CSB. The CSB has the responsibility to carry out regular performance assessments on all measures of the work program (based on the defined key performance indicators), prioritize tasks based on a needs assessment of the community, decide on the adaptation of short and long term strategies for the consortium, evaluate and decide on proposals from the community for the use of budgets for the enhancement of services (taking into account recommendations by the CAC), and give recommendations to the CSC for the use of unspent funds (e.g. resulting from maternity/paternity leaves or job vacancies etc.). The CSB also has the obligation to carry out a regular risk assessment for all task areas and give instant recommendations for the management of occurring risks to the CSC. The CSB will meet four times a year for two days. The co-applicants and academic societies have the obligation to ensure that the CSB is always in a position to decide and act. In case regular members of the CSB cannot take part in meetings, the party concerned must nominate a stand-in person for the meeting that will exercise the vote in the CSB.

Culture Spokesperson Committee (CSC): The CSC is the central executive body in NFDI4Culture's governance. It consists of all co-spokesperson and has the responsibility to coordinate, monitor and realize the implementation of the work program. The CSC bears the financial responsibility towards the DFG and ensures that funds are always used in compliance with the NFDI's funding policy. The CSC has to check and give approval to all financial decisions made by the other organs of the consortium's governance. The applicant institution will perform a final compliance check for the use of funds. In case a financial decision does not pass the check by the CSC and the applicant institution, it is returned to the CSB for reworking. The CSC sends six representatives (one for each task area) and the consortium's spokesperson to the steering board. The CSC conducts a monthly conference in which upcoming tasks are discussed and decisions on the efficient realization of measures are taken.

Culture Advisory Council (CAC): The CAC is a body of up to twelve independent counselors from national and international research and infrastructure and cooperating NFDI consortia (its exact composition may vary over time). The members are nominated and invited to the CAC by the CSB. The CAC advises the consortium on (long-term) strategic issues and can be requested to give recommendations on proposals from the community for the use of grants for the enhancement of the consortium's services. The CAC has full access to all reports and work results and takes part in the yearly Culture Community Plenary. It can also be requested to give counsel in controversial issues, always taking into account general developments and goals in the NFDI and beyond.

Reporting, guidelines for decision-making, conflict resolution

The ACO and the TCO will provide an efficient digital reporting system that is accessible to all members of the NFDI4Culture Community (members of governance bodies, co-applicants, participants). Utmost care will be taken that decisions are well prepared and substituted by information compilations beforehand, based on measurable demands as well as verifiable success criteria in the decision-making process itself and transparently documented for the community afterwards. In the preparatory phase, NFDI4Culture has already made very good experiences with low-threshold push/pull digital communication systems (messengers) beyond email and mailing lists. Setting up such an infrastructure for the community will be one of the primary tasks in the initial phase of the consortium. In case conflicts or delays of delivery occur on the operational level, the spokesperson and the co-spokespersons will have to be informed instantly and will take immediate and proper action to best resolve the controversy. The resolving of conflicts and mediation of controversial topics on the governance level is the responsibility of the consortium's spokesperson. He/she will look at the differing positions, mediate between the parties and decide which measures need to be taken to best resolve the conflict. For substantial controversies that might affect the progress of the consortium as a whole the spokesperson can request advice from the CAC and the NFDI directorate.

Disbursement of funds

As applicant institution of NFDI4Culture the AWLM will set up a cooperation agreement between the members of the consortium and ensure the implementation of the DFG's compliance rules on all levels of the consortium. Since the AWLM acts as a funding institution in the Academies' program of the Union of the German Academies of Sciences (federal and state funding), she has proven financial control mechanisms and established administrative workflows to ensure the appropriate use of funds during disbursement to co-applicants and participants. Regular financial controlling and yearly (external) audits will be implemented by the AWLM's ACO lead (sciences administration staff of the consortium) in close information exchange with the consortium's spokesperson. The AWLM will set up individual agreements for the transfer of funds using the sample cooperation agreement provided by the DFG (if applicable). This will also be applied to future cooperation between co-applicants and (possibly new) participants. Grants to participants will only be released after a proposal and evaluation process by the governance bodies (Community Plenary, Steering Board, Advisory Council) as described above has taken place and an official disbursement decision by CSB/CSC has been filed to the ACO.

Viability of structures

NFDI4Culture has included several measures in its work program to develop adequate business and operational models in the field of research data management on tangible and intangible cultural assets (cf. TA4|M5, TA6|M3 and TA7|M3). Since collaboration in the consortium will extend over a period of at least 10 years, it is to be expected that consortium partners may withdraw, or new partners may join the consortium. To this end, NFDI4Culture provides consolidated admission and withdrawal processes via the CSB. In addition, a mechanism will be implemented to review the overall structure of the consortium every two years (internal evaluation 2022 and 2024) in order to implement necessary structural adjustments in response to the changing needs of the users. Long-term viability and permanent acceptance in the community of interest will largely depend on the flexibility in the use of funds and the implementation of an innovation performance radar in response to the user feedback gathered through the many participation channels of the consortium (forums, teams, helpdesks, information portal etc.).

Operating model

NFDI4Culture strongly supports the creation of a single legal entity for the NFDI and will align its internal structures in the best possible way to this end. At the time of writing it is therefore not intended to transform NFDI4Culture into a legal entity of its own. For the time being and until all bodies of the NFDI have been set up and a common legal framework is in place, NFDI4Culture will select an operating model based on a consortium contract as this is best practice in current national and international consortium projects (cf. Collaborative Research Centers, Clusters of Excellence etc.). Special attention will be paid to the fact that no commercial exchange of goods or services takes place and that funds will only be transferred between non-profit entities or legal entities under public law

(contracted members of the consortium) with the sole purpose of promoting science and research and only by way of a genuine grant. As a strong advocate of the *Berlin Declaration*, NFDI4Culture looks forward to close collaboration with other NFDI consortia and the NFDI directorate on this topic.
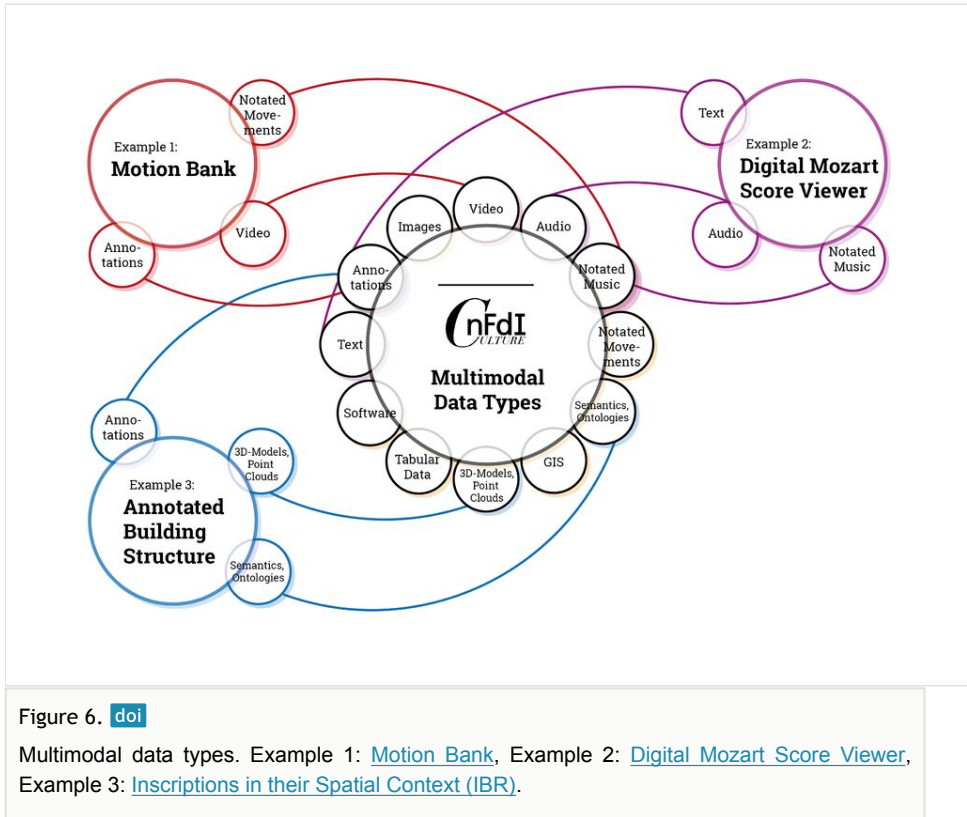
Research Data Management Strategy

Since the **tangible and intangible dimensions** of non-textual cultural assets have an intrinsic value that is never fully realized in a digital representation, the research data targeted by NFDI4Culture is subject to **specific requirements**: First, it is necessary to **differentiate between digital representations (reproductions) of cultural assets**, including their metadata, and **procedural research data** resulting from research processes on material and immaterial assets. Second, it must always be taken into account that **digital representations** of cultural assets **can become immaterial cultural assets in themselves**. This opens up new perspectives and innovative aspects for future research. Third, material and immaterial cultural assets are often **subject to complex legal and sometimes ethical conditions**. This also holds true for research data generated from such assets, which then requires thoroughly considering copyrights and personal rights (Klimpel 2015, Klimpel and König 2015) as well as questions of provenance during the whole research data life cycle. Fourth, research data as an outcome of research processes on material and immaterial cultural assets can be **closely linked to the cultural economy that uses the results** for economic purposes. In turn, the **cultural products created by the economic use** of this data **can themselves become the objects of research** among the communities addressed by the consortium. The users participating in NFDI4Culture do not only generate data *about* their research objects (through metadata or annotations), the **research objects themselves (usually in the form of digitized representations) become an inseparable part of the research data life cycle** as well.

Overview of data types to be managed

The consortium's RDM strategy deals with **complex multimodal data types on two levels**:

1. **digital representations of cultural assets**, such as all forms of **2D digital reproductions**, (e.g. of paintings, sculptures, sheet music, scores, but also of performed works, etc.), **3D models of cultural assets** (from artifacts to large man-made structures such as monuments, buildings or rooms) created by photogrammetric processes, laser-based 3D scanners or structured light systems or by (re-)construction using 3D modeling software, all types of **audiovisual data** on material and immaterial cultural assets (music, film, stage performances etc.), and

2. **procedural research data**, such as **graphic formats** (raster formats and vector formats including their image metadata, e.g. Exif, XMP, etc.), vector formats for the digital depiction of musical notations, **digital reconstructions of cultural assets, computer-generated structures** of objects or buildings on the basis of CAD/

CAAD or rendering programs, computer-based simulations of cultural spaces or artistic performances, **encoding and annotation formats** (e.g. XML-based formats for encoding music or annotating images, time tags or shape annotations for videos, etc.), **metadata and serializations of semantic models** (e.g. CIDOC CRM, IFLA LRM and other derivatives) for the description of material and immaterial cultural assets, exchange formats (e.g. IIIF, LIDO), authority data (e.g. GND, Getty: AAT, TGN, ULAN) (Fig. 6).

Figure 6. doi

Multimodal data types. Example 1: Motion Bank, Example 2: Digital Mozart Score Viewer, Example 3: Inscriptions in their Spatial Context (IBR).

Development of the work program, user involvement and acceptance

The work program is the **result of close collaboration between users and providers** during several extensive community workshops (August to December 2018; February, July, September 2019) and a **large-scale survey of existing RDM infrastructures** (April to August 2019). 51 institutions have expressed their interest in NFDI4Culture and submitted contributions to the consortium's infrastructure and services. The survey showed that there is great demand for a consortium dealing with research data on material and immaterial cultural heritage within the NFDI and that the consortium can draw on many mature (long-term) publication infrastructures, repositories and software solutions as well as on well-developed materials and expertise in the area of legal services and teaching/training. On the basis of the survey data, the **consortium partners have identified and formalized 28**

user stories (available at https://nfdi4culture.de/images/us/NFDI4Culture_UserStoryAll.pdf) **and validated them through close dialogue with researchers** from the participating communities and users from GLAM institutions. All user stories were aligned to the research data life cycle and served as the basis for creating the research data management strategy and the measures in the work program. In a second step, all suggested contributions were made eligible for efficient evaluation and collaborative measure planning by the consortium partners. In a third step, the **work program and all planned measures were presented and validated in an open community workshop** with participants from the academic societies and institutions, users from universities and GLAM and representatives from the neighboring NFDI consortia in the humanities (namely the initiatives NFDI4Objects, NFDI4Memory and Text+).

Against the background of the existing **multimodal data types** in the cultural heritage domain the consortium's partners have created a holistic RDM strategy that responds to needs from the level of data capture, collection and enrichment to the level of data analysis, data publication (taking into account complex ethical and legal situations), data reuse and training. This strategy is realized by measures in **six operative task areas** that can be **aligned to the different stages in the RDM life cycle**, allowing for reasonable overlaps and cross area flexibility. A **seventh task area contains the governance** and administrative measures and binds together communication and decision-making processes (Fig. 7).

Current/envisaged state of research data management in respect to the work program

Task area 1 (TA1), Data capture and enrichment: Since the 1990s, GLAM institutions have been intensively involved in the digitization of their holdings, often financed by third-party funds. Over the years, basic technical standards and metadata standards have been established. As a result, extensive holdings can now be comfortably used for research. Three deficits can be noted in the status quo: First and foremost, the holdings of large, infrastructurally well-positioned institutions were processed, while holdings of smaller institutions with the same scientific relevance are still not digitally accessible. In response to scientific needs a greater variety of techniques, for which generally accepted standards are still lacking, is increasingly being used. The work with digital representation has so far only led to the use of digital methods to a small extent. The **envisaged state** is that the exchange of know-how supports the digitization of cultural assets across the board. That standards are also developed and communicated for special techniques such as 3D digitization, computer tomography, thermography and multispectral photography. That the work with digital methods will be promoted through demand-oriented data enrichment measures and that new scientific perspectives and questions will thus open up.

Figure 7. doi

NFDI4Culture's task areas in alignment to the RDM life cycle.
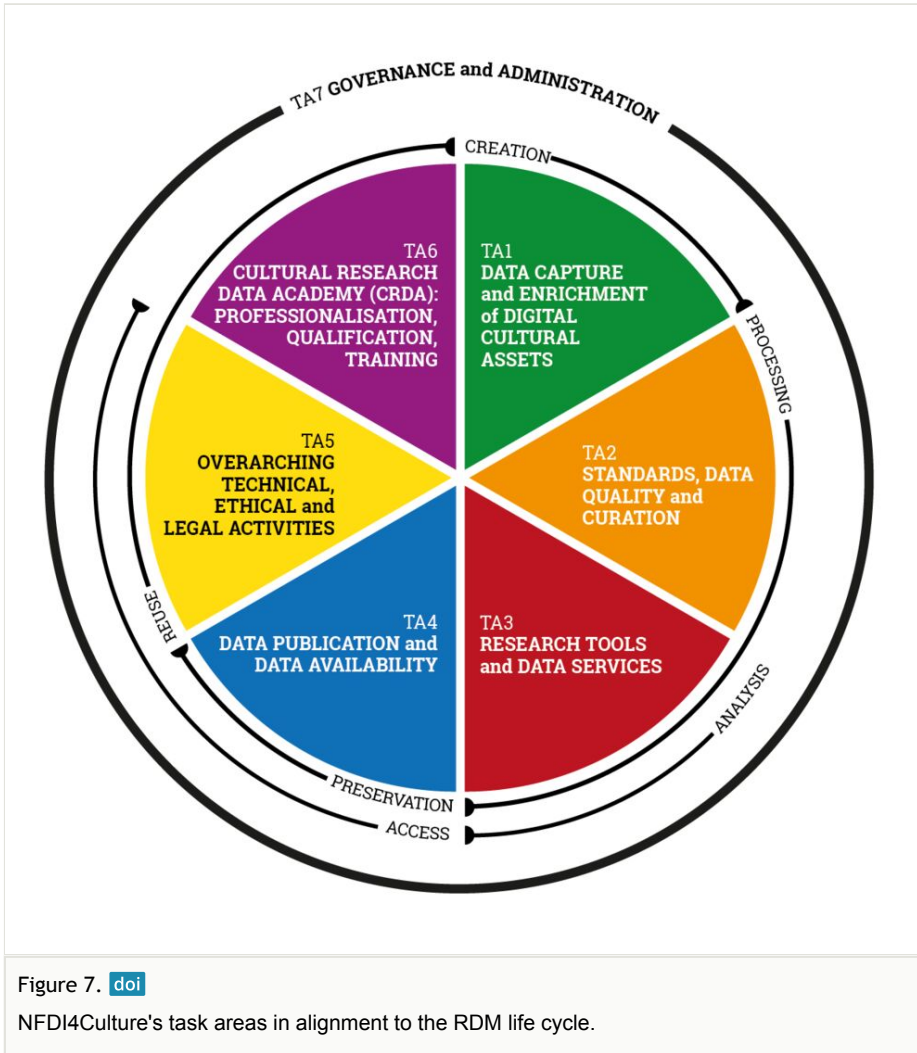
Task area 2 (TA2), Standards, data quality and curation: Sustainable digital research requires research data, which is findable, accessible, interoperable and reusable. To facilitate this, community-approved data standards, specialized and linkable authority data, data quality management, and curation processes are required (Aliverti et al. 2015, Bove 2018, Bracht 2016, Patton 2010, Simon 2007, Bove and Schmahl 2015,Rula and Zaveri 2014,Kailus 2018). In recent years, several standards relevant to NFDI4Culture have been developed: With **CIDOC CRM** (ISO 21127:2014) (Crofts et al. 2011, Stein and Balandi 2019), ICOM has published a formalized conceptual model to support the integration, access, and exchange of diversely structured information from the field of cultural heritage. **IFLA LRM** provides an entity relationship model for the (bibliographic) description of various characteristics of physical and conceptual entities. The internationally accepted **MEI** standard is used to express physical and intellectual characteristics of music notation documents (Veit and Richts 2018). Object-oriented metadata is increasingly exchanged

between repositories using **LIDO** (Coburn et al. 2010Knaus et al. 2019). **Authority data and vocabularies** such as the **GND** (GND-Kooperative and Kett 2017, Niggemann et al. 2017,van der Graaf and Waaijers 2014) of the DNB, **AAT** and **TGN** of the Getty Research Institute, and **Wikidata** (Poulter 2017,Association of Research Libraries / Wikimedia Foundation 2019,Schelbert 2017) function as central vocabularies in the NFDI4Culture domain. Standardized data allows **semantic technologies** to find and contextualize dispersed collections and to link the content of data sets and repositories in a significantly improved semantic way. In the highly diversified cultural heritage landscape mentioned above, the use of standards for data production is currently optional (at best). When, however, standards are applied, the degree of compliance varies, and the application is often carried out in an uncoordinated manner, dependent on local interpretation and use. The benefits deriving from the FAIR Principles and the resulting relevance for the introduction of quality management throughout the entire data life cycle are still insufficiently familiar. TA2 offers **measures to strongly encourage the use of these standards** as well as to **further develop them** on a national and international level. The consortium partners already play a decisive role in the exploration of the modeling requirements of the described standards in application contexts and research projects. They can draw from extensive experience in data mapping and data harmonization and are active in authority data development. The **envisaged state** is that standards are available that can match the specific needs of the scholars in the NFDI4Culture domain. Thanks to effective communication, they will be accepted and further developed within the community. A comprehensive data quality management will be implemented covering the entire data life cycle. Non-standard data sets, which continue to be of scientific relevance, will be upgraded or enhanced by individual data rescue measures.

Task area 3 (TA3), Research tools and data services: For the highly relevant area of research software and data services in the field of research on material and immaterial cultural assets there currently exists **no dedicated infrastructure** that would allow knowledge exchange and coordination for the specific requirements of the NFDI4Culture community. There are many institutions and DH-centers with expertise (like UPB/ZenMEM, DCH, UZK/prom, UMR/MCDCI, TIB, FIZ, mainzed) that offer support for sustainable development and operation of research tools but there is **no consulting agency** that covers the topics of the NFDI4Culture consortium in relation to the development, consolidation, operation and certification of sustainable, interoperable research tools and data services on the basis of the FAIR principles (Manovich 2011,Brett and Croucher 2017, Arendt and Taentzer 2013, Arendt et al. 2011, Röwenstrunk 2018). Within the DHd association, there is a **working group Research Software Engineering** (co-founded by the designated Speaker of NFDI4Culture), closely connected to the international RSE community, which pursues the goal of sustainable software development that can serve as a model (Czmiel et al. 2018, Schrade 2017). Regarding the **findability and interoperability** of research tools and data services, there exist several registries (like https://fairsharing.org or https://www.re3data.org/) but **no domain-specific overview** over the offerings exists that would be of impact to the research community of NFDI4Culture. A quite large number of domain-specific research tools and data services exist in the field of material and immaterial cultural heritage, but quite often they are not meeting the

standards with regard to quality and sustainability. Additionally, the mentioned institutions and DH-centers provide software development infrastructures but there is still a lack of documentation of workflows and guidelines for their efficient use (e.g. development environments, project management tools, version control systems, ticket systems, user feedback systems, test environments, continuous integration, continuous delivery, container systems, etc. which address this situation and are adapted to the respective contexts). Therefore, **the envisaged state of research data management by TA3** is that NFDI4Culture enables the researchers to find the relevant tools and data services for their research and obtain access to them. Researchers verbalize their needs and participate in the conception and (further) development of research tools and data services. Researchers and institutions which develop software obtain support through NFDI4Culture for the development, consolidation and certification of sustainable, interoperable research tools and data services through a single point of contact. This creates a low threshold in software-driven research and data reuse through (web) services in various environments (cloud, server, individual).

Task area 4 (TA4), Data publication and availability: A key task for NFDI4Culture and the NFDI as a whole is to provide reliable and sustainable services for the storage, publication and digital preservation of research data. In cultural heritage, **research data** very often **cannot be clearly separated** from the publication of research results. Rather both, the data and the forms of publication (be it conventional or digital) increasingly merge into **multimodal publications of complex research data and results**, in which a wide variety of resources from distributed storage locations form a multilayered publication, connected via hyperlinks or embedded in research software (Effinger et al. 2018b, Effinger et al. 2019). For example, these can be texts from art history (Effinger 2018), supplemented with images from the SLUB/DF and digital 3D models. In musicology, a combination of sound recordings and the notation of music in MEI can form such a complex data type. In the performing arts, it can be annotated video recordings of performances. These complex data types require appropriate publication environments with interfaces to discipline or format-specific repositories, including their connection to systems for long-term digital archiving, as well as intensive advice for researchers. Such services are currently only provided by few institutions and by no means cover all areas of NFDI4Culture. For the **envisaged state** of RDM, TA4 will therefore provide **re-usable reference implementations** based on the already established offerings in art history (arthistoricum.net) and musicology (musiconn.publish, musiconn.performance) as well as other building blocks such as media/rep/ (Repository for Media Studies), RADAR (Generic Research Data Repository) and the SLUBArchiv (Content Preservation), which allow institutions and projects to set up and offer similar complex data publication services quickly and easily. These reference implementations comply with the standards developed in TA2 with regard to metadata, data formats and interfaces and thus ensure compatibility with the NFDI infrastructure in general and particular the NFDI4Culture infrastructure. Many existing offers are isolated solutions and do not meet the requirements for FAIR data. The ubiquitous use of **unique identifiers** such as DOI and ORCID as well as the structured provision of metadata as **Linked Open Data** (cf. TA5) will allow for reliable referencing of (data) publications even beyond the boundaries of NFDI4Culture, thus

fostering **FAIR principles**. The complex data types employed in the NFDI4Culture community require **new or expanded concepts and services for digital preservation**, including adequate operating models. Objects distributed across several repositories represent a challenge for reliable long-term storage, as do relatively new data types (e.g. digital 3D models). In close cooperation with other consortia in the NDFI and international initiatives, TA4 will develop **solutions and offer services** for all data types relevant to the NFDI4Culture community.

Task area 5 (TA5), Overarching technical, ethical and legal activities: Up to now, the manifold research data collections on material and immaterial cultural assets are dispersed across institutions in varying states of access and quality. Hitherto, **no common data federation layer exists** that would allow uniform and overarching findability and access to digital representations of cultural assets and according metadata and authority data. At the same time, research data on tangible and intangible cultural assets is in many cases subject to **complex legal circumstances** – as are the cultural assets themselves (due to copyrights, related rights, rights of use, exploitation rights, domiciliary rights, personal rights, protection of cultural assets etc.; Hartmann 2013, Hartmann 2014, Hartmann 2017, Hartmann 2018, Klimpel and König 2015, Klimpel et al. 2017, Universitätsbibliothek Mainz 2019). **Critical aspects of data ethics must also be considered**, e.g. in cases where the cultural assets originate from a colonial past or where data publications could lead to the loss of the actual assets (due to looting or vandalism etc.; Alge 2019, Deutscher Museumsbund 2018). At the same time, highly relevant data hubs connecting structured data and terminologies from the cultural heritage domain already exist beyond science (e.g. in large scale information infrastructures such as Europeana or community driven projects such as Wikidata). It is essential for future methodological and analytical innovations and success in research on material and immaterial cultural heritage that these hitherto **disconnected data domains get connected in a standardized, quality assured, structured** and at the same time **collaboratively curatable way**. Thus, **for the envisaged state of RDM**, TA5 will provide efficient joint authentication and access solutions (AAI) which on the one hand enable digital research and on the other guarantee legal certainty e.g. with regard to privacy regulations (TA5|M1). A **comprehensive information portal will provide a single point of entry to the consortium** for the users and **connect** NFDI4Culture's information resources to international infrastructures such as **OpenAIRE and EOSC** (TA5|M2). The overarching **federation, connectivity, findability, accessibility, interoperability and reusability** of NFDI4Culture's information resources will be dealt with by the modeling and implementation of a **Linked (Open) Data platform** that overarchingly integrates structured data in the form of a **knowledge graph** and a **terminology service** that will be made **curatable for the users** from research and GLAM that participate in the consortium (TA5|M3). This way, **domain experts can continuously contribute and improve to the data federation layer** of NFDI4Culture and, over time, also connect it to the data offerings of other NFDI consortia and the cultural data hubs beyond the NFDI. Parallel to these central measures TA5 will offer regular advice and consultation to researchers in order to create **legal certainty for the providers and users** with regard to the existing legislation (Legal helpdesk, TA5|M4 and advisory panel on data ethics, TA7|M6). Furthermore, the **experiences and requirements by the stakeholders**

will be collected and passed on to legislators in order **to advise** them with regard to drafting adequate **legislation** (TA5|M5, stakeholder process for the improvement of legal certainty).

Task area 6 (TA6), Cultural Research Data Academy (CRDA): Professionalization, qualification, training: A holistic RDM strategy thoroughly includes elaborated competencies in handling digital research data, software tools and new methods (Martin 2018,Busch et al. 2018). Qualified and up-to-date competencies for reflective dealings with digital research data in the sense of **Data Literacy** are necessary throughout the whole research data life cycle (Bundesministerium für Bildung und Forschung 2019, Gesellschaft für Informatik e.V. 2019, Heidrich et al. 2018a, Ridsdale 2015). It is of decisive significance for researchers to also learn **to critically evaluate and reflect on** algorithms and **software tools** in scientific and social contexts (**Code Literacy**). With regard to the **current state of RDM referring to the scientific disciplines involved**, several institutions and initiatives have set up courses, trainings, workshops and labs, hackathons and carpentries, and dedicated professorships in order to foster the development of competencies in dealing with cultural research data (Heidrich et al. 2018a, Heidrich et al. 2018b). All those actors strive at supporting users in their ability to capture data, to code and/or to set up and guard analyzing and visualizing tools for cultural research data, to conceptualize data models and software architectures, to know about and how to apply CIDOC CRM, TEI, MEI, LIDO, IIIF etc. However, offers and knowledge on cultural data science are dispersed, although first steps towards networking were made with e.g. a registry for courses on Digital Humanities (cf. https://registries.clarin-dariah.eu/courses/) and e.g. the Parthenos network (cf. http://www.parthenos-project.eu/portal/trainingsuite). Also, training options for GLAM employees, for multipliers and for the citizen science community are more or less lacking, as well as acknowledged quality criteria or frameworks. At the same time, there is a great need – articulated by small heritage institutions as well as renowned scholars, e.g. whilst attending our Culture Community Workshop – for easily overviewing and participating in suitable and qualified qualifications and trainings.

Thus, for **the envisaged state of RDM concerning Data and Code Literacy**, the consortium will set up a Cultural Research Data Academy (CRDA) that allows for needs-based development of subject-specific frameworks, centralized options to enter these training options, to certify them by defined quality criteria compatible with the FAIR4S-framework of the EOSC and to quickly develop new and innovative training options. The CRDA will endeavor to reach a **stronger embedding** of Data and Code Literacy **in curricula** in the sense of **critical data studies and critical algorithm studies**. As NFDI4Culture has a **strong grounding in universities and professional associations**, there is a high potential for CRDA to ensure the continuous transfer of results to teaching. Thus, the planned measures address, firsthand, a close involvement of the scientific community in defining and sharpening needs and constantly monitoring and evaluating them. Second, quality criteria will be developed compatible with the FAIR4Sframework, as well as a specific framework and recommendations for content, didactics and methods. Third, existing offers, courses and trainings from a wide range of participants of

NFDI4Culture and beyond will be scanned and sorted and, if desired, certified. Based on a detailed evaluation of existing course portfolios, qualified offers for training will be made visible and accessible. New offers, e.g. a mentoring network and training courses for GLAM employees, will be set up and supported. Fourth, the CRDA consulting helpdesk will support users when conceptualizing own courses, in developing own skills or in organizational implementation; also, it will conceive suitable training offers by (co-)applicants and participants.

Data selection, quality management, community engagement

As an effect of an ever-increasing digital orientation within the subjects represented in NFDI4Culture, numerous data pools varying in size, quality, and research potential have been created in recent years. In the course of the preparations for the NFDI4Culture application, an initial overview of existing data sets was compiled as part of a comprehensive survey (cf. chap. Analysis of the existing research and infrastructure environment and development of the planned consortium according to the users' needs). The aim of NFDI4Culture is to allow for the long-term preservation of existing relevant research data, and to ensure that new data is generated according to specific data quality management criteria, which are yet to be set up. Within the framework of NFDI4Culture, the topic of data curation and quality management is essentially the responsibility of TA2, in cooperation with TA4. However, ultimately, this matter affects all tasks, reaching from data capture (TA1) to teaching and training (TA6). Measures are aimed at a NFDI4Culture specific development, concrete support, and broad communication of a data quality management system. The latter, in turn, ensures compatibility with the FAIR criteria for newly emerging data corpora, and at the further curation of existing data corpora, some of which must also be enhanced and transferred into sustainable repositories. In principle, counseling measures for the creation of new corpora and subsequent curating activities will cover the entire spectrum of subjects relevant to NFDI4Culture. For this reason, representatives of three different disciplines (art history, musicology, dance/theatre studies) will work together within the central FAIR clearing agency. In close exchange with the Forum (TA2|M1) and the research community, the FAIR clearing agency develops data quality management that covers the entire data life cycle, with quality criteria adequately matching the specific data found in NFDI4Culture. The FAIR clearing agency publishes QM-guidelines and supports novel techniques for data quality measurement. Further, it offers project-specific advice, including advice on applications for third-party funding. NFDI4Culture will also work towards ensuring that aspects of professional data management are increasingly taken into account as evaluation criteria for third-party funding decisions. This measure will strengthen the establishment of data quality aspects in the community in the long term.

With regard to the multitude of pre-existing data corpora in all NFDI4Culture relevant disciplines, the curating measures will be carried out in a selective manner. Firstly, the FAIR clearing agency supports all institutions that possess active data in the NFDI4Culture context in upgrading and securing their data corpora. This approach not only saves data, but also spreads knowledge about data quality management. Secondly, the FAIR clearing

agency itself administers concrete curating measures or initiates and supports measures to be executed within the framework of the individual projects in TA2|MA3. The selection of the data corpora to be processed in this context is carried out based on the scientific relevance of the data to be processed and in close consultation with the forum (TA2|M1), in which experts from the various NFDI4Culture disciplines are represented. Relevance is objectified, for instance, through existing peer-reviewing (e.g. within the framework itself, or from third-party funders), and through measurable subsequent use. A further aspect to be considered is the risk posed to relevant scientific data corpora in terms of unclear ownership or impending infrastructural obsolescence. Such data sets must be identified, and the risks addressed.

As of today, NFDI4Culture can already build on existing initiatives and collaborative infrastructures for quality management, data curation, and data backup. Thus, for example, profile and strategy of the AKBF, founded in 2004, are focused on the establishment of methods, techniques, and standards for archiving, indexing, and publishing photographic holdings. Furthermore, in conjunction with respective consultation services, the FID Musikwissenschaft curates musical event data and subsequently stores it in the central repository musiconn.performance. In addition to the activities of the FAIR clearing agency, these decentralized approaches, which enable close networking with the scientific community, must be further strengthened. Such approaches can disseminate knowledge of data quality management, increase acceptance of it, and, most importantly, allow further development of NFDI4Culture's specific quality criteria in accordance with the needs of the scientific community.

Monitoring and evaluation

Monitoring and continuous re-adaptation of user needs are thoroughly integrated in all task areas and in the governance of NFDI4Culture: Firstly, all task areas work together to drive the forums, as the basic structure to regularly gather and monitor user needs in a reflexive and participating manner. As part of the forums' meeting, a regular standardized user survey will be established and exceeded. Secondly and as a prerequisite, NFDI4Culture has developed detailed user stories which are continuously evaluated and adapted in order to recount users' needs and journeys alongside and across the research data life cycle. Thirdly, all task areas head at close monitoring of task area-specific user needs while being in contact with the community, e.g. when interacting with uses throughout helpdesks and consulting agencies. Needs, feedback and ideas emerging during those interactions will immediately be forwarded towards the ACO, which takes care of initiating suitable adaptations and innovations. To further ground monitoring and evaluation on a technical basis, all relevant task areas will contribute to technical reporting, e.g. via log analysis, traffic (TA3) and downloads. TA5 will take care of collocating all technical monitoring of NFDI4Culture services. Fourthly and lastly, monitoring and evaluation is an intrinsic part of NFDI4Culture governance, as the CSB is designated to exceed monitoring as a main task, especially via the scientific communities representing the users and participating in and holding the majority of the CSB. Also, external monitoring is integrated via the Advisory Council and via exchange and collaboration with other NFDI consortia.

Metadata standards

In order to meet the FAIR criteria, research data requires descriptive metadata. It ensures its findability and contextualizes it. Metadata contains information on accessibility, connectivity, and re-usability of the research data described by it, and, thus, organizes the academic communication with and about research data. For a successful communication process, metadata is standardized and is both machine and human-readable.

Domain-specific metadata standards

Metadata in cultural studies describes cultural assets and their digital representations as research data, as well as procedural research data on the basis of general and domain-specific syntactic or semantic feature groups. The described entities, thus, become addressable and relationships between them and other research data and metadata can be modeled. For research data of the NFDI4Culture community, general metadata formats are extended or differentiated by special features. The definition and sustainable maintenance of these domain-specific feature groups takes place, above all, in the respective disciplines and requires a certain degree of standardization. Domain specifics supplement, for example, generic authority data concepts (e.g. the GND), cataloging guidelines (e.g. the library guidelines Resource Description and Access), exchange formats (e.g. LIDO), vocabularies or ontologies, and reference models (e.g. CIDOC CRM, FRBR/er/oo, IFLA LRM).

Both above mentioned types of research data from the NFDI4Culture domain (representations of cultural assets and procedural research data) are described with metadata: For *digital representations of assets*, the Metadata Encoding and Transmission Standard (METS) has been established, which can be combined with other object and media type-specific description standards. Metadata for research data as digital representations of assets of the NFDI4Culture domain basically contain

1. identifier,
2. descriptive information on the cultural asset digitized,
3. technical information on data formats,
4. information on the technical environment of creation,
5. information on the contextualization of the production situation,
6. legal information on accessibility,
7. information on technical accessibility issues,
8. information on tools for displaying and using research data, if applicable, and
9. information on the internal structure of the digital object.

While the METS profile for digitized texts, which is maintained as part of the MARC standard by the Library of Congress, has already been a recognized standard for several years (METS/MODS), the adaptation of METS for NFDI4Culture-specific digital representations, for example of AV media (e.g. as METS/MODS for AV), 3D digitized objects (METS/MODS for 3D) or notated (and coded) music (METS/MEI) and images (METS/LIDO and IIIF/LIDO for 2D) is still partly pending. For *procedural research data*,

various metadata formats are available or will be developed. These include, for example, specific authority data, subject-specific vocabularies and thesauri, annotation formats such as TEI/MEI as XML specifications, Exif, XMP.

For NFDI4Culture's research data and metadata, some special features are to be considered: Metadata describing material and immaterial cultural assets itself can serve as research data. For this reason, metadata in the cultural heritage domain must allow the mapping of **blurriness and uncertainty**. Complementary or contradictory information also shall be captured in a human and machine-readable way. NFDI4Culture meta- or research data is characterized by **double performativity**: The special relationship of digital to analog object requires special attention in the metadata, which must take into account the specifics of the digital implementation as a performative act as well as the digitized source itself. In addition, metadata or research data in the NFDI4Culture domain does not have to describe a digital or physical source, or possibly cannot be materialized in a physical object (which in turn can be represented digitally), but deals with **events** that **only manifest themselves in a certain realization** of an "object" (as concept) in certain social, time- and space-dependent contexts or even in potentially infinite *options* of realization. These "**events**" can be documented with digital or physical testimonies, but they all will only describe certain aspects of the event and never the full realization. Sometimes digital or physical sources are even completely absent – and the metadata must be able to reflect this. Metadata in NFDI4Culture does not only describe research data, but also supports its **exchange** between different applications and re-use contexts, repositories and presentation platforms. Here, too, general and generic formats are extended by NFDI4Culture specifics. Especially relevant are IIIF, LIDO, MusicXML, MEI or AAF. Finally, metadata describing research data must be distinguished from metadata describing metadata. This "higher-order metadata" makes semantic statements about the information sources, authorships, and contexts of the metadata described and in particular contains essential statements about the quality of both research data and metadata described.

In order to ensure syntactic and semantic interoperability and reusability of research (meta-)data of the NFDI4Culture domain, it is necessary that a transdisciplinary standardization process takes place through the definition of quality standards and their application on the basis of appropriate criteria and associated procedures for data collection in a sustainable dialogue of both researchers and infrastructure providers. In the NFDI4Culture work program, this task is dedicated to TA2 "Standards, data quality and curation" (cf. TA2|M1, TA2|M2). In addition, established standards will be further developed and selected collections of data will be saved or enhanced (cf. TA2|M3). In addition, the measures of other task areas consider domain-specific (meta-)data standards in relation to their respective task spectrum: TA1 focuses on the further development of IIIF for domain-specific requirements, TA3 on the citation file format, TA4 on metadata standards for long-term archiving and availability (e.g. PREMIS, DataCite Metadata Kernel), TA5 on the further development of metadata in the context of Wikidata, Europeana data model, CERIF etc. – always driven by the question of what minimum of additional information is needed to grasp the particularities of the domain and still remain connectable to other research (data) areas. When developing NFDI4Culture specific metadata formats, NFDI4Culture can rely

on cooperation in an international context and community, such as IFLA, IAML, the TEI and MEI council, RISM international, and the CIDOC LIDO Working Group as part of ICOM.

Common metadata standards with respect to the NFDI

This interaction between generic and domain-specific metadata standards guarantees the compatibility of special formats with general formats. METS, e.g., as a generally accepted container format includes a manageable number of special categories and allows the integration of highly specialized formats in the descriptive metadata section. The further development of metadata formats of the NFDI4Culture domain also sustains data formats of other disciplines: the further development of AV metadata, which serves the long-term availability (e.g. METS/MODS AV) or the exchange of AV data (IIIF A/V) for example, can of course be reused as a general standard for all disciplines and domains dealing with AV formats. The ontology development will also pay attention to ensuring connectivity to ontologies of other disciplines and will foster NFDI4Culture knowledge graph as part of a general knowledge graph consisting of different NFDI graphs. In accordance with the *Berlin Declaration*, NFDI4Culture will participate in the harmonization of standards as a cross-cutting topic with other consortia and build on accepted and prevalent formats and authority data. Members of NFDI4Culture are already active in projects such as GND4C, DataCite, the Research Data Alliance and AAT and will contribute their expertise and preliminary work to NFDI4Culture and the whole NFDI community.

Implementation of the FAIR principles and data quality assurance

FAIR principles

The FAIR Data Principles form the overarching guidelines for the prioritization and orientation of all measures in the consortium. Since their publication in 2016, they have been regarded as a framework of quality requirements for sustainable research data management to be established in the context of a European Open Science Cloud (EC EGFD - European Commission Expert Group on FAIR Data 2018). The FAIR principles for ensuring findability, accessibility, interoperability and reusability are general recommendations for all research areas. Content community standards have to be respected in the description of research objects. It will be necessary to define how the FAIR criteria are to be implemented in relation to all phases of the research data lifecycle. As early as in the conception phase of a project, the requirements for interoperability and (re-)use must be taken into account.

The FAIR principles are being applied in all six task areas. **Controlling and coordinating FAIR-related activities lies primarily with TA2 Standards, Data Quality and Curation**: Providers and researchers will collaborate to achieve an agreement on open data standards, persistent identifiers, version and provenance management, interfaces, software and its documentation. This will **improve** the **findability** of the objects but also the **accessibility, interoperability and reusability** for research over the research data's entire life cycle. The **FAIR clearing agency** (TA2|M2) will coordinate the mediation and

implementation of the FAIR principles within the consortium and beyond. With regard to the selected standards, it identifies adaptation requirements and accompanies the further development of standards by requirements management. It supports data producers and repository operators in the implementation of the FAIR Principles and the assessment of the relevant data quality within the framework of an audit procedure. The consortium provides impulses for anchoring the FAIR requirements at the advisory and decision-making levels of research funding. The impact of the FAIR Principles is strengthened when their implementation becomes a quality feature for the funding of projects and institutions.

Additional technical, organizational and legal means to implement FAIR are present in other task areas: **TA1** aims at improving findability and reusability and at enhancing interoperability by identifying supportive structures and well-matched FAIR-oriented digitization standards adjusted to the needs of the researchers. In order to improve interoperability, it strives for a data discovery system that aims at rendering related information. In **TA3**, the code of research software will be rendered findable and accessible as well as interoperable, as it turns into research data itself. The development of the IIIF annotation server will enlarge access to high quality image data. All workflows and guidelines for using tools will be oriented towards FAIR principles. It is considered e.g. that versions and derivatives of software are made available permanently by PIDs. In **TA4**, existing offerings for professionally supported, sustainable and uncomplicated research data publications and long-term archiving will be significantly expanded, thus fostering findability, accessibility and reusability of cultural research data. In close cooperation with TA2, the Consulting Agency (TA4|M2) will support researchers in the publication of complex data types in accordance with the FAIR principles, e.g. via handouts and guidelines. The Core Trust Seal certification for repositories in TA4|M4 warrants, among other things, the accessibility and reusability of the published data. All data publication and archiving services will support the metadata schemas developed in TA2 and will uniquely register all objects via persistent identifiers. **TA5** will foster FAIR by setting up a uniform interface interoperable with various data formats, e.g. by implementing a CERIF-compatible HTTP interface. The implementation of the knowledge graph enables integrative access to distributed and heterogeneous research data in the consortium's domain, enhancing all four FAIR principles, e.g. by describing it as LOD. **TA6** implements FAIR through the FAIR4S-framework developed within the EOSC: While developing criteria and culture-specific frameworks within the Culture Research Data Academy, the FAIR4S framework will serve as a guideline and reference. Within specific trainings (TA6|M3) for employees at GLAMs, the FAIR principles will be a leading part of the curriculum. Advice within the helpdesk, e.g. when conveying suitable trainings and workshops, is permanently oriented towards offers compatible to FAIR criteria.

To ensure the **viability of the technical basis for the implementation of the FAIR criteria**, the consortium prefers openly licensed standards, authority data, vocabularies and ontologies. Their availability must be guaranteed in the long term. They are maintained and developed by an active expert community and are widely used. Standards themselves must be as FAIR as possible. The consortium also preferably relies on open source products with a substantial international developer community. With regard to the

implementation of FAIR, it focuses on services and databases which are generally recognized as relevant by the user community and which are already part of the long-term, sustainably maintained portfolio of data providers or which represent a subsequent development within a convincing overall strategy. As technical aspects are embedded in a wider context, the consortium will aim at strengthening awareness that FAIR principles are the core element of RDM in the field of cultural heritage.

Technical, organizational and legal means to implement data quality assurance

At the time of writing it is not possible to fall back on an already defined level of data quality in the domain of cultural heritage. The level must always be defined by taking into account the original intent of the users during data generation and the application purposes of the data. It is crucial to develop a uniform and, if necessary, scalable level of data quality to be achieved, while taking into account the specific requirements of the user community. Domain specifics such as the handling of uncertainty or of the dynamics resulting from continuous updates of the data must also be taken into account. The following strategies enable the implementation of quality requirements:

- Specification of areas in which standards, authority data, vocabularies and ontologies are underdefined for subject-specific concerns or show insufficient coverage. Execution or support of appropriate developments and adjustments.
- Development of a technically supported test procedure for different quality levels in data sets using already existing tools for testing data quality and FAIRness (within the framework of an audit procedure).
- Adaptation of repositories and cross-discovery systems in order to enhance re-use opportunities from improved data quality.

Services provided by the consortium

Against the background of NFDI4Culture's multidisciplinary research landscape with its strong institutional differentiation the overarching aim is the **documentation, evaluation, enhancement, operation**, and **further development** of existing and new services for RDM which are provided by the numerous members of the community. These include services for high quality digitization of research objects, coordination of standardization of data and metadata formats for material and immaterial cultural heritage, services and guidelines for ensuring research data quality, provisioning of infrastructure and help for sustainable research software development, information about, evaluation of, and guidance for usage of research tools and data services, platforms and authoring tools for publication of research data and results including long term preservation and persistent identification, access to NFDI wide structures and services, support in legal matters, and services for professionalization and continuous education in Data and Code Literacy and other topics of RDM. To make these services visible and available to the whole community will need an unprecedented effort that is highly enquired. All of the planned services are

organized in three distinct categories: **(1) consulting, (2) support** and **(3) development (including enhancement, merging and upscaling).**

The **permanent availability and longevity** of these services will be ensured by the co-applicant institutions and by the broad participation of a variety of institutions that are committed to the consortium. Their broad range of competencies cover needs in on every level, including data security. All participants have given **sustainability commitments** regarding their provided services. In addition, the consortium's organizational structure is designed in a way that new participants can be incorporated continually. This ensures not only a **close affiliation and continuous contribution** of the research field to the consortium but also **continuous development** of the consortium adapting to changing needs of the community. The consortium will adapt its structures to existing national, European and international infrastructures (e.g. OpenAIRE, EOSC, DataCite, RDA; Aliverti et al. 2015) in order to ensure **maximum connectivity**. The members of the consortium are actively involved in the development of international initiatives for data and metadata formats, which ensures a **high interoperability** of the planned services. The intended **strong activation and involvement of communities** in the planning, development and operation of services will lead to wide acceptance, usage, and participation of community members. NFDI4Culture will provide easily understandable documentation of its infrastructure, services, methods, and workflows and will publish software as well as guidelines under the terms of **open source and open access**. One of the target audiences of TA6 are GLAM institutions, which provide services to the broader public. Enhancing qualification within these institutions will lead to better integration and dissemination of the services.

A **participatory process** is constitutively for NFDI4Culture's approach to services. It is based on the close collaboration between the providers and the academic societies. In order to constantly expand and improve the services the consortium's work programme also schedules to **survey needs and requirements** of the community. In order to ensure a constant adaptation to the needs of the involved communities, the decision on suitable measures must therefore be kept flexible in parts. For this reason, **expert forums** are installed in all task areas (cf. chapter Interaction between users and providers) in order to enter into a close exchange with the community on changing needs.

Services in TA1

Digital representations of tangible and intangible cultural assets constitute an important basis of research in cultural heritage. The overarching goal of TA1 is to **align the production of digital representations with the demands of researchers**. To reach this goal the consortium will provide the following services to its users:

TA1|S1 Consulting by a digitization agency in questions of data capture and enrichment

TA1|S2 Provision of materials and information services on data capture and enrichment

TA1|S3 Development and provision of modules for automatic/collaborative data enrichment

TA1|S4 <u>Provision</u> of participation opportunities in automatic data enrichment for stakeholders

Services in TA2

The overarching goal of TA2 is to identify and **develop standards and quality criteria** in order to ensure the **implementation of FAIR principles** in the research area of the consortium. These standards will be incorporated in data services and research tools (TA3) as well as in data publications and archiving services (TA4). Their implementation will allow for cross-disciplinary reuse of research data in the NFDI and beyond (in cooperation with TA5). To reach this goal the consortium will provide the following services to its users:

TA2|S1 <u>Consulting</u> services by a FAIR clearing agency on research data quality management

TA2|S2 <u>Provision</u> of best practice materials on research data quality management

TA2|S3 <u>Provision</u> of reusable authority data offers (regarding data formats as well as data hubs)

TA2|S4 <u>Provision</u> of documented exchange formats, research and metadata standards

TA2|S5 <u>Development</u> of standardized exchange formats, research and metadata standards

TA2|S6 <u>Provision</u> of data rescue measures for selected uncared-for corpora (community-driven)

Services in TA3

The overarching goal of TA3 is to **identify, provide and develop relevant research tools and data services** along the needs of researchers. It will implement the standards defined by TA2 (when applicable) and provide interfaces to data publication and archiving services described in TA4. To reach this goal the consortium will provide the following services:

TA3|S1 <u>Consultation</u> agency for research tools and data services

TA3|S2 <u>Provision</u> of evaluation, analysis and certification of research tools and data services

TA3|S3 <u>Provision</u> of documentation and DOIs for existing research tools and data services

TA3|S4 <u>Provision</u> of "one-click solutions" for selected tools and services

TA3|S5 <u>Support</u> in the development of sustainable research tools and data services (guidelines)

TA3|S6 <u>Support</u> conception, setup, use, and operation of software development infrastructures

TA3|S7 <u>Provision</u> of a contact point to report needs with regard to software sustainability

TA3|S8 <u>Development</u> of selected research tools and data services (needs-based)

TA3|S9 <u>Development</u> of interoperable interfaces for existing tools (needs-based)

Services in TA4

The overarching goal of TA4 is the provision of adequate and interoperable **data publication and archiving infrastructures for all disciplines and complex data types in NFDI4Culture**. These services will implement standards set out in TA2 and will be based on appropriate operating models. To reach this goal, the consortium will provide the following services:

TA4|S1 <u>Consulting</u> agency focusing on long-term data publication and operation of repositories

TA4|S2 <u>Provision</u> of immediately reusable data collections for the participating disciplines

TA4|S3 <u>Support</u> for publication of complex data types, including adequate authoring tools

TA4|S4 <u>Development</u> of long-term preservation for ensuring permanent availability for all data types

TA4|S5 <u>Development</u> of reference implementations for data publication and repository services

Services in TA5

The overarching goal of TA5 is fostering the access and reuse of data through support services in the field of **rights and data ethics** and to **interface with further NFDI consortia** regarding NFDI4Culture technologies and services and vice versa. To reach this goal the consortium will provide the following services to its users:

TA5|S1 <u>Provision</u> of a single point of entry to the consortium via the NFDI4Culture portal

TA5|S2 <u>Provision</u> of information resources into European infrastructures (EOSC)

TA5|S3 <u>Provision</u> of authorized access to NFDI4Culture data via the common AAI infrastructure

TA5|S4 <u>Development</u> of a technical validation service for research data

TA5|S5 <u>Provision</u> of access to the planned Research Data Commons (cf. *Berlin Declaration*)

TA5|S6 <u>Provision</u> of collaboratively curatable LOD (knowledge graph, terminology service)

TA5|S7 <u>Consulting</u> and clarification for researchers in questions of data rights (legal helpdesk)

TA5|S8 <u>Support</u> and increase of legal certainty through a stakeholder process

TA5|S9 <u>Support</u> by an ethical framework developed in close cooperation with the community

Services in TA6

The overarching goal of TA6 is to build an infrastructure that collects, clusters and visualizes training and qualification opportunities in the field of (cultural) **research data management**. To reach this goal the consortium will provide the following services:

TA6|S1 <u>Consulting</u> in teaching and training offers for Data and Code Literacy

TA6|S2 <u>Provision</u> of systematized documentation of existing teaching and training offerings

TA6|S3 <u>Support</u> of institutions and teachers in the development of teaching and training

TA6|S4 <u>Support</u> for implementing digital training offers

TA6|S5 <u>Provision</u> of monitoring structure for development of competencies in RDM

TA6|S6 <u>Development</u> and provision of new training offers

Longevity of services

Many of the listed services are **already in operation** and have **funding independent** from the NFDI, thus **ensuring the longevity** of the offerings. The responsible infrastructure facilities have made **long-term commitments** for the operation of the services and the availability of the data. Continuous enhancement, development and acceptance are ensured by:

- discipline-specific **regular surveys**,
- **forums** which discuss and evaluate existing services on a regular basis and provide advice for the services roadmap, and
- **consulting agencies** which will collect additional user feedback and identify existing shortcomings in the service portfolio.

Work Program

Table 1: Overview of task areas

Table 1.

Overview of task areas.

Task Area	Measures	Responsible co-spokesperson(s)
TA1 Data capture and enrichment of digital cultural assets	**M1** Forum – Digital representations of material and immaterial cultural assets in research	TIB **Prof. Dr. Ina Blümel** *Architecture, Information Infrastructure* SPK **Reinhard Altenhöner** *GLAM, Information Infrastructure*
	M2 Observing and understanding the main pathways of digital cultural objects: digitization – enrichment – reuse	
	M3 Coordination of digitization centres for uniform quality levels	
	M4 Development of complementary environments for collaborative and automated enrichment of non-textual cultural data	
TA2*1 Standards, data quality and curation	**M1** Forum – Scientific requirements for standards and quality criteria for research data on non-textual material and immaterial cultural assets	SLUB **Prof. Dr. Barbara Wiermann** *Musicology, Library and Information Science* UMR **Dr. Christian Bracht** *Art History, Information Infrastructure*
	M2 FAIR Clearing agency for research data on non-textual material and immaterial cultural assets	
	M3 Further development of data formats and authority data in the area on non-textual material and immaterial cultural assets and data curation	
TA3 Research tools and data service	**M1** Forum – Sustainable software development for the field of NFDI4Culture	UZK **Dr. Lisa Dieckmann** *Art History, Computer Science* UPB **Dipl. Wirt.-Inf. Daniel Röwenstrunk** *Business Computer Science, Musicology*
	M2 Consulting agency for sustainable development and operation of research tools and data services in the area of NFDI4Culture	
	M3 Domain-specific registry of (existing) tools and services	
	M4 Needs-based (further) development of specific research tools and data services for research in the field of material and immaterial cultural assets	
	M5 Software development infrastructures for subject-specific project contexts	
TA4 Data publication and data availability	**M1** Forum – Data publication and data availability	UHD **Dr. Maria Effinger** *Art History, Open Access, Information Infrastructure* SLUB **Dr. Jens Bove** *Art History, Information Infrastructure*
	M2 Consulting agency for data publication and data availability	
	M3 Provision of subject-specific and re-usable reference implementations	

Task Area	Measures	Responsible co-spokesperson(s)
	M4 Integration of relevant data publication and digital preservation services into the NFDI4Culture infrastructure	
	M5 Establishment of digital preservation procedures for complex data types and development of operating models for digital preservation in NFDI4Culture	
TA5*2 Overarching technical, ethical and legal activities	**M1** Cross-area and cross-cutting infrastructural tasks	FIZ **Prof. Dr. Harald Sack** *Information Service Engineering* FIZ **Prof. Dr. Franziska Boehm** *Jurisprudence*
	M2 NFDI4Culture information portal	
	M3 Linked Open Data 4Culture	
	M4 Legal helpdesk	
	M5 Stakeholder process to improve legal certainty in the research context	
	M6 Advisory panel for data ethics	
TA6*3 Cultural Research Data Academy (CRDA): Professionalisation, qualification, training	**M1** Forum – Cultural Data and Code Literacy	UMR **Prof. Dr. Malte Hagener** *Media Studies* UPB **Prof. Dr. Andreas Münzmay** *Musicology*
	M2 Needs, CRDA quality criteria framework and recommendations	
	M3 CRDA portfolio with NFDI4Culture training offers	
	M4 CRDA helpdesk for consulting and support	
TA7*4 Governance and administration	**M1** Administrative Coordination Office (ACO)	AWLM **Prof. Torsten Schrade** *Digital Humanities*
	M2 Technical Coordination Office (TCO)	
	M3 Governance operations	
	M4 Coordination, cooperation, knowledge pooling for cross-cutting topics	
	M5 Culture RDM Kickstarter	
	M6 Dissemination, outreach and community enlargement	

Task Area 1: Data capture and enrichment of digital cultural assets

As research data, digital representations of cultural assets (either digital-born or created by digitization followed by subsequent processing and enrichment steps) provide an essential basis for scientific work. NFDI4Culture specific data types are, thereby, mostly non-textual 2D documents (e.g. scores, images, photographs), 3D models (e.g. monuments and musical instruments), as well as audio and video materials. Due to a plurality of different objectives and interests inherent in the scientific community, an ever-increasing **variety of technologies** is being employed: 3D digitization, thermography scanning, computed tomography, and multispectral photography have emerged alongside the more traditional practice of 2D scanning.

The extent to which digital representations are capable of **satisfying the needs** of the scientific community is **determined early on during the production phase of the data** and of its corresponding metadata. Particularly, given the permanent development of technical and conceptual methods, the possibility of reuse and reproducibility of research data can only be ensured through documentation of employed procedures and production conditions. In addition, an enrichment of digital assets with additional, content-describing metadata, or the use of techniques such as Optical Music Recognition (OMR), is particularly important for non-textual materials. This not only allows for an efficient search and navigation, but also **enables new methods of scientific interpretation** of relevant data corpora. Accordingly, **a regular exchange between producers of digital materials and their users is crucial (Aim 1).** Important questions in this regard are: How are scientific working contexts structured when using digital representations of immaterial or material cultural assets? Which characteristics must digital representations possess in order to meet the requirements of the scientific community? Which processes will guarantee a reliable production? And which documentations are needed to ensure trust in the production process? **Standardization for digitization tangible and intangible cultural assets is to be furthered with regard to the various technologies in use (Aim 2). The development and implementation of modules for a collaborative and automatic enrichment of digital representations of cultural assets is to be promoted (Aim 3).** To facilitate traceability and reusability of research data, information acquired in this enrichment steps have to feed the interoperable and reliable metadata layer (TA5|M3).

Risk analysis and risk management

SWOT analysis for TA1 is summarized in Table 2.

Cooperation with other task areas and cross-area dependencies

TA1 will be in particularly close exchange with **TA2** regarding the development, quality assurance and sustainability of data enrichment vocabularies and ontologies. Recourse will be made to the overview components developed in **TA3** for digitization measures, for example to the registry, and we will pursue the needs-based development of research tools for data enrichment. There will be close cooperation with **TA4** in the field of interfaces for digital reproductions / digital representations. Knowledge about data capture and enrichment of material and immaterial cultural assets is networked and contextualized in **TA5** especially through contribution to the knowledge graph. Moreover, the knowledge is drawn upon in close cooperation with **TA6** (CRDA) within the context of the development of Data and Code literacy and the networking of target groups/researchers. There is close exchange with **TA7** on governance and innovation development.

Measures

TA1|M1: Forum on digital representations of material and immaterial cultural assets in research

The extensive digitization activities of the past two decades have simplified access to research related objects for researchers. However, this so far has not yet led to a significantly increased application of **digital methods** in the NFDI4Culture context. The researcher's work with digital representations of relevant objects and the necessary means for his work have, so far, only been addressed on a marginal scale at best. The aim of the measure is, therefore, to promote self-reflective insight into the scientific use of digital representations, and to formulate requirements for digitization centers. Thus, questions arising within the scientific community, and regarding the practical use of digital representations and their corresponding (meta)data will be referred to the producers and processed in a dialogue with the latter. This marks the starting point for further methodological developments and new research questions.

Table 2.

SWOT analysis for TA1.

STRENGTHS creating OPPORTUNITIES	WEAKNESSES preventing OPPORTUNITIES
• The stakeholders forming NFDI4Culture are well connected with the communities and can act as accepted change catalyst especially for the production phase of data; • The breadth of partners and methodical skills allows a new level of clarification and community decisions	• A lack of engagement of the decisive stakeholders, leading to their rejection of resulting specifications; • The integration of competences is either unsuccessful or is not accepted; • Individual interests within the consortium could have a marginalizing effect and divert the focus from the actual requirements
STRENGTHS minimizing RISKS	**STRATEGIES avoiding RISKS**
• A preexisting communication and coordination process connecting the various communities and stakeholders ensures that the risk of neglect of developed services or unacceptable conditions (licensing, costs) is minimized; • Active observation and targeted support of relevant tools, in particular in the area of data enrichment, guarantees a high level of maturity in explorative and experimental products even during the selection stage	• Close communication with the scientific communities and early evaluations make it possible to meet the requirements of specialists; • A successful provision of tools and active data platforms are ensured through clear guidelines for licenses, short-step developments in a flexible organizational body, and binding agreements on the basis of a clear decision-making channel within the consortium

Overview of tasks in TA1|M1

TA1|M1|T1 Implementation and operation of a forum on the use of digital representations of material and immaterial cultural assets in research: Gathering of relevant requirements for the scientific use of digital representations and data enrichment.

TA1|M1|T2 <u>Creation of material on digital representations for researchers</u>: Introductory and in-depth material on the use of digital representations in accordance with the forums' priorities.

TA1|M1|T3 <u>Organization of forum events on digital representations:</u> Introductory and in-depth events such as workshop discussions on "exhibitions" or "marketplaces" within the framework of other large events in order to bring together producers and science and to initiate new developments. Specific cases and project proposals are crucial.

TA1|M1|T4 <u>Establishment of a single point of contact:</u> In order to coordinate, channel and distribute the digitization requests made by researchers it is central to identify the appropriate procedure and take into consideration aspects of interoperability and reusability.

Key value propositions for TA1|M1: The exchange within the community via the forum creates clarity about established and new methods and standards of both digitization and research with digital objects **(KVP1).** The adaptation and further development of established methods and community standards as well as the creation of new ones are carried out according to the needs of the community **(KVP2)**. This is supported by cooperation with infrastructure and producers **(KVP3)**. Triggers for new issues emerge based on new methodical approaches and the lively discourse surrounding them **(KVP4)**.

User stories getting added value: 3, 5, 13.

TA1|M2: Observing and understanding the main pathways of digital cultural objects: digitization – enrichment – reuse

A comprehensive understanding of the status quo of **how the digital content of relevance to cultural studies and GLAM** (that is, both digitized objects and born-digital objects) **is captured, enriched and (re)used** is the ultimate goal of this measure. Monitoring of the output of digitization centers is automated as far as possible and complemented with regularly updated manual descriptions and estimations. The same is done for large-scale efforts of indexing (e.g. in order to make objects findable in discovery and viewing environments of GLAM institutions) and enrichment. Where and how does the description and annotation of objects mainly take place, which parts of the data are actually used for discovery and display in online usage environments? Where is further enrichment and contextualization (e.g. by means of setting up virtual collections, by traceable references in scholarly and non-traditional online literatures) most typically situated? Moreover, the enrichment process is examined with regard to the following issues: Which digital reproductions are enriched? Where? By whom? How is this carried out? Which tools and standards are used? The aim is that in 5 years' time a profound understanding will have been created about the available data, its use, as well as the tools most typically used for data enrichment. A better understanding of the characteristic life cycles of digital objects from the cultural field with regard to the succession and meshing of data production, enrichment and reusability should uncover recognizable patterns to be taken into consideration.

Overview of tasks in TA1|M2

TA1|M2|T1 <u>Developing a monitoring infrastructure:</u> The aim is to gain a comprehensive picture of the actual scope of digitization at regular intervals which includes the description, enrichment and reusability of the data sets by digitizing institutions and third parties.

TA1|M2|T2 <u>Source analysis:</u> Quantitative analysis of the figures from the monitoring infrastructure with a focus on digital reproductions as well as on metadata and differences between data providers; qualitative analysis in order to understand motivation, potentials and problems; overview of producers.

TA1|M2|T3 <u>Annual survey of research community and usage of digitization centers.</u> Compilation and evaluation of usage statistics; Identifying highly relevant development needs for the scientific use of digital representations and data enrichments.

TA1|M2|T4 <u>Needs-based preparation of the digitization process:</u> Digitization, indexing and enrichment of specific, individual inventories can be planned which possess a high centrality and broad impact to NFDI4Culture's community of interest.

Key Value Propositions for TA1|M2: Data providers / digitization centers obtain a clear picture of how others undertake the digitization and enrichment process and which standards are used. Among other things, this enables them to avoid the duplication of work and optimize their services **(KVP1)**. Researcher who use and produce research data know the scope and form of the data available to them and can encourage the creation of new data **(KVP2)**. At the same time, developers of data enrichment tools obtain an overview of the standards used, can adapt their tools and identify gaps for new functionalities **(KVP3).**

User stories getting added value: <u>5</u>, <u>6</u>, <u>12</u>, <u>14</u>, <u>16</u>, <u>20</u>, <u>23</u>, <u>24</u>, <u>26</u>, <u>27</u>

TA1|M3: Establishing uniform quality levels for NFDI4Culture specific digitization techniques

The DFG's *Practical Guidelines on Digitization* (DFG-Praxisregeln zur Digitalisierung) offer an established standard for conventional 2D scanning of predominantly textual sources, for the generation of associated metadata, as well as for long-term archiving. In regard to the creation of digital copies of specific objects (e.g. slides, photo negatives, audio, video, 3D objects – e.g. musical instruments), the definition of a standard has yet to be negotiated. Moreover, standards for digital representations using special techniques such as thermography (for watermarks), multispectral photography (for faded manuscripts) etc. still need to be developed. M3 analyses the requirements resulting from M1 and M2 and will develop FAIR-oriented standards in consultation with relevant digitization centers at libraries, museums etc., and in interdisciplinary exchange with technical fields. The measure also directly promotes the reusability of digital representations of audio-video materials by standardizing IIIF/AV manifests.

Overview of tasks in TA1|M3

TA1|M3|T1 <u>Networking and cooperation of digitization centers</u>: working groups with a strong focus on types of media or methods, such as 3D, audio, video, AR/VR, 2D, thermograph or computer tomography.

TA1|M3|T2 <u>Identifying needs for standardization for the digitization of cultural assets</u>: Standardization of specific materials, methods or types of media such as watermarks, photographs of buildings, musical instruments, digitization of previous image layers, including activities; documentation of processing steps (technical provenance); documentation of the reasons for selection/digitization; Improving the stability and resolvability of citations and references to digital reproductions.

TA1|M3|T3 <u>Regular survey and analysis of development needs in the digitization field</u>: Development of a **digitization index** which compiles the digitization requests and activities of the different digitization centers; establish uniform quality levels which can be addressed within or outside NFDI4Culture.

TA1|M3|T4 <u>Standardization of IIIF manifests for audio-video materials</u>: Further development of the standard in cooperation with the IIIF A/V Technical Specification Group with the aim to extend the advantages of interoperability, which IIIF provides for images, to A/V and to enable interoperable access to temporal and spatial segmentation and transformation of content.

Key Value Propositions for TA1|M3: The quality of digitization from digitization centers is enhanced, and, thereby, aligned with the specific needs of the research community **(KVP1)** .

User stories getting added value: 3, 5, 13.

TA1|M4: Development of complementary environments for collaborative and automated enrichment of non-textual cultural data

In order to make a broad range of rich cultural data sets more accessible to the public, the measure aims for the **collaborative annotation of digital multimedia objects** of potential relevance for cultural studies exploiting the combination of crowdsourcing with machine learning, focusing on the intricacies of annotating images and film. Initially, state-of-the-art feature extraction methods for images and audiovisual data are collected and matched to appropriate sets of openly available digital collections. These collections, to which automated annotation methods are already applied or can be applied on a large scale with moderate effort, are fed into public online environments like Wikimedia Commons. Thereby, M4 makes the outcome of artificial image recognition research, that would otherwise be lost or gone unnoticed by the public, usable and available to larger online communities. Corresponding with the goals of **TA2** and **TA3**, modern web standards (e.g. W3C web annotation standard) based tools are assembled as a toolbox for developers and infrastructure service providers working in the field. The tools are aligned to the **TA3** IIIF annotation server. Toolbox and use case implementations focus on layers for expert

communities as well as for laymen and (limited access) layers for the classroom, allowing to interact with the objects and annotations and ultimately create digital spaces for research, education and discourse. This measure is aligned with Wikimedia GLAM initiatives such as Structured Data on Commons project and **TA5** in general, which helps bringing the potentials of linked open data to environments like Wikimedia Commons. Several tools and use cases are made available for active and sustainable use, in close alignment with community needs and for keeping up with collections, methods and tools.

Overview of tasks in TA1|M4

TA1|M4|T1 <u>Community Engagement:</u> Continuous assessment of NFDI4Culture stakeholder needs and use cases in automated and collaborative annotation of images and movies, with special consideration given to the joint events of the NFDI4Culture partners.

TA1|M4|T2 <u>Collecting and matching openly available image and movie collections</u> with methods from artificial image recognition research. Implemented as a focused contribution to TA3|M3 and TA3|M4. Setting up a modular, reusable tool set for user interactions with annotations in images and movies on image detail / movie scene level, supporting multiple layers of annotations following IIIF and W3C web annotation standards, in synchronization with TA3|M4|T2

TA1|M4|T3 <u>Complementing the media viewers and metadata display in Wikimedia Commons</u> with additional user annotation elements; Analysis and assessment of user interactions with automated annotations, ultimately feeding back data and analysis to the AI image community, as well as to image and film annotation tool developers.

Key Value Propositions for TA1|M4: Researchers on the topic of machine-assisted and automated annotation, extraction and inferencing obtain training data to improve their algorithms and tools **(KVP1)** including rare insights into the outcomes of their research in the context of large scale real world applications, beyond limited user test scenarios **(KVP2)**. Scientists as the users of research data tools are involved in needs-assessment and developments processes and are decisive for judging the relevance and applicability of tools in science **(KVP3)**. Citizens and open culture enthusiasts in general are involved in needs-assessment and development processes and contribute to the relevance and applicability of tools **(KVP4)**.

User stories getting added value: <u>6</u>, <u>20</u>, <u>24</u>, <u>25</u>, <u>26</u>, <u>27</u>.

Key performance indicators

Aim 1: Regular exchange between producers and users of digital material

- **KPI 1.1**: Based on a periodical qualitative survey of satisfaction, competence and demand in the research community which at the same time takes into consideration different types of material, creation and generation methods (realized through M2)
- **KPI 1.2**: Periodically updated statistics which compile information on the relevant digitization centers and evaluate the demand for various properties (human and

machine access, demand for digital reproductions of specific types of material, creation or enrichment methods, etc.)

- **KPI 1.3** Medium-sized and small institutions are enabled to realize digitization projects according to their (users/researchers) needs in cooperation with infrastructure partners.

Aim 2: Expanding standardization through coordination between producers

- **KPI 2.1**: Existence of corresponding written agreements one year after the consortium's work has commenced
- **KPI 2.2**: Evidence of continual further development and maintenance through transparent information and proven community work
- **KPI 2.3**: Execution of interoperability tests and reusability checks within the framework of M2

Aim 3: Provision of modules for automatic and collaborative enrichment of non-textual research data, including pilot use cases

- **KPI 3.1**: As a contribution to TA3|M3 and M4, a set of digital image and film collections of relevance for cultural studies is identified and related to potentially applicable AI methods
- **KPI 3.2**: After four years, a total of over 100 additions or amendments have been made by more than 10 different users each month in collaborative annotation environments.

Task Area 2: Standards, data quality and curation

Quality-assured research data provides the basis and instrument for researching material and immaterial cultural assets. It must be characterized by reliability, completeness and reusability and be described by metadata ensuring traceability and availability. Until now, there has been a lack of recognized quality-management procedures for research data on material and immaterial cultural assets which first and foremost consider their **domain-specific double performance**. Likewise, deficits and desirables in the field of specialized authority data and vocabularies for the requirements of the NFDI4Culture community have been identified. In order to address these deficits needs for sustainable data quality will be identified and measures prioritized (**Aim 1**) in close consultation with the NFDI4Culture community. A FAIR clearing agency promotes the development, definition and application of quality standards for NFDI4Culture research data and develops guidelines for data quality management. (**Aim 2**). In five years, quality-assurance processes for research data on material and immaterial cultural assets have been established (**Aim 3**). In constant exchange with researchers, existing data formats, authority data, and specialist vocabularies are optimized for scientific use or newly developed if currently unavailable (**Aim 4**). Numerous stakeholders who possess relevant expertise will be involved in these measures: in close cooperation with researchers, they will define scientific requirements, provide new impulses, participate in validating and communicating results, apply them in research projects, and act as multipliers. The

stakeholders gathered in the forum, among others, steer and bundle these developments. This open process with a strong dialogue basis and productive exchange ensures close alignment with the needs of science.

Risk analysis and risk management

SWOT analysis for TA2 is summarized in Table 3.

Table 3.

SWOT analysis for TA2.

STRENGTHS creating OPPORTUNITIES	WEAKNESSES preventing OPPORTUNITIES
• Standards acknowledged by both individual researchers as well as data providers are continuously and sustainably developed in accordance with the FAIR principles; • Relevant experience from prior projects regarding data quality is bundled; • Existing international networks of participating research data services can be integrated	• Due to a lack of awareness, information, and support in the community, non-standardized data corpora are created. They, at best, are only findable with great effort, are neither compatible nor reusable, and are not available long term
STRENGTHS minimizing RISKS	**STRATEGIES avoiding RISKS**
• In the rapidly developing field of research, close contact to the community enables new needs to be quickly dealt with in order to avoid non-standardized solutions; • A lack of awareness for data-quality measures is dealt with through extensive community work; • The acceptance of consolidated and universally applicable data-management guidelines is strongly supported by the consortium; • The development and enhancement of authority data and specialist vocabularies in terms of quality and quantity boosts acceptance and compatibility, and reduces the number of isolated solutions	• Data standards, authority data and vocabularies which were evaluated and developed in NFDI4Culture will be transferred to the wider community through cooperation with existing and new projects; • Within the forum, the variety of work necessary to improve the quality of research data for NFDI4Culture will be collected and prioritized together with the community; • The recourse to NFDI's overall structure prevents a subject-specific pillarization; • The guidelines on research data quality become a criterion for research funding decisions, thus enhancing their significance

Cooperation with other task areas and cross-area dependencies

An active exchange with **TA1** ensures that the requirements of researchers with regard to data capturing and the technical standards involved are considered when the FAIR roadmap is developed. In the field of data enrichment, there is close coordination with regard to the recommended measures for primary data and their consideration for the data quality management. In exchange with **TA3** one of the focuses is on the evaluation of tools

for measuring the FAIRness of data. This ensures that researchers dealing with data quality measurement and enhancement (e.g. authority data and special vocabulary) are supported with concrete information on available tools and services. This also initiates the creation of quality management tools according to the needs that have been identified. The close exchange with **TA4** focuses on the requirements of the repositories for the implementation of the FAIR principles in the development of quality requirements. The catalogs of criteria for data quality and quality assurance to be developed in TA2 are applied by **TA4** to the appropriate repositories and, thus, ensure reusability of the research data stored in them. In accordance with the envisaged Core Trust Seal certification, the catalogs of criteria demonstrate the suitability of a repository for standards-based and interoperable data services. TA2 lays the essential foundations for the knowledge graph to be modeled in TA5 and identifies the LOD corpora necessary for it. In addition, the exchange with **TA5** ensures that current legal developments regarding action-relevant aspects in the subject areas of TA2 are considered. The exchange with **TA6** explores training opportunities for both young and experienced researchers that also pertain to the TA2 subject areas (FAIR principles). Finally, the exchange with FAIR activities planned in other consortiums such as Text+, NFDI4Memory and NFDI4Objects ensures a comparison of the content and organization of the further development of authority data and standards.

Measures

TA2|M1: Forum on scientific requirements for standards and quality criteria for research data on non-textual material and immaterial cultural assets

A stable and lively expert forum with representatives from the NFDI4Culture Community **discusses, sharpens and evaluates the work area's ongoing tasks**. Moreover, it **formulates and prioritizes additional standardization and curation needs**. These lead to concrete tasks during the funding period, or to subsequent applications for third-party funding. The close exchange with scientists guarantees that the specific applications and domain requirements are taken into consideration in the development, consolidation, operation as well as the certification of standards-based and interoperable data services on the basis of the FAIR principles.

Overview of tasks in TA2|M1

TA2|M1|T1 Establishing and operating a forum for quality management of research data for non-textual material and immaterial cultural assets: Regular conducting or updating of needs assessment and of the fulfillment of research needs in order to evaluate the results and progress of the measures in this task area and to identify data repositories of particular scientific relevance (with TA4). Continual dynamic process to outline future tasks and manage ideas for their implementation (M5).

Key Value Propositions for TA2|M1: The measures of the task area are the result of the specific needs of researchers, scientific institutions and professional bodies **(KVP1)**. Specific indications for the research-oriented improvement of data quality and the improved orientation of the usage possibilities for the providers of research data flow into the

prioritization and sharpening of specific measures **(KVP2)**; dialogue between users and data providers is improved **(KVP3)**.

User stories getting added value: 2, 6, 7a, 8, 9, 10, 11, 13, 14, 16, 17, 18, 21, 25, 27.

TA2|M2: FAIR-Clearing agency for research data on non-textual material and immaterial cultural assets

Criteria for measuring the quality of research data are developed, and a **catalog of measures for quality assurance** is drafted, taking into consideration domain-specific requirements. The aim is to **establish a quality assurance process** for producers and providers of research data with prototype implementation on selected data sets and repositories. This also comprises the **preparation of an audit** of data quality for later transfer into a comprehensive certification process. Parallel to that, a **support and consulting center** is set up and maintained in order to support researchers and data providers in issues related to data quality and standards in all stages of the data life cycle. Researchers at various starting points with differing needs or prior knowledge can obtain individual flexible **support** for their projects. Where possible, reference is made to suitable offers of the other task areas.

Overview of tasks in TA2|M2

TA2|M2|T1 Development of a FAIR framework: Drafting criteria to measure the quality of the diversely structured research data connected to material and immaterial cultural assets. Development of criteria for the concrete implementation of the FAIR principles, drafting recommendations for ontologies and reference models with particular consideration given to CIDOC CRM, METS/MODS-AV, METS/MEI, TEI-MEI, TEI/IIIF. Evaluation of tools to measure the FAIRness of data and the enhancement of data quality (with TA3). Development and offer of an audit procedure for data providers and repositories. Development of a monitoring procedure in line with the criteria stipulated by the FAIR Data Maturity Model WG (Research Data Alliance).

TA2|M2|T2 Advice and services for data curation: Approach at the planning stage of a project, enabling the selection and implementation of appropriate standards and strategies to integrate quality assurance measures in its data management plan, continual integration of community impulses and requirements into the development of the FAIR roadmap. Conducting data qualifications and migrations in order to transfer particularly relevant but isolated data into sustainable structures or to optimize data which is relevant for large groups of users.

Key value propositions for TA2|M2: Researchers and data providers obtain concrete support in quality management when creating, curating and making data available **(KVP1)**; concrete measures to improve data quality are identified **(KVP2)**; researchers and data providers can apply recommended standards and procedures in the context of their respective initial situation and interests **(KVP3)**. The quality and reliability of data offers can be assessed **(KVP4).**

User stories getting added value: 2, 3, 5, 7, 8, 10, 13, 15, 16, 17, 21, 22, 27, 28.

TA2|M3: Further development of data formats and authority data in the area on non-textual material and immaterial cultural assets and data curation

Established standards for metadata and authority data with high potential for reuse are conceptually enhanced to match the needs of NFDI4Culture's community of interest. Initially, the following will be considered: authority data from the **GND** for (multimedia) works of art, authority data for various versions and expressions of musical works (Bicher and Wiermann (2018), Bicher 2018, Wiermann 2018) the documentation format of **MEI**, the exchange format for object-related data of material cultural heritage **LIDO.** This boosts interoperability and findability of research data from the NFDI4Culture context. **Authority data** for material and immaterial cultural assets, the identification and referencing of which can be considered a primary challenge for all disciplines addressed (works of music, architecture, media arts, visual and performing arts), can be suitably addressed as **Linked Data**.

Within the interdisciplinary context of NFDI4Culture, the transitions between the various sectors of art, music, film and performing arts and the resulting requirements for an authority dataset can be comprehensively and precisely examined. Here, the previous conceptual work conducted over the past years and the ongoing DFG project GND4C can be built upon. The GND, which has until now been insufficiently defined with regard to object categories relevant in NFDI4Culture, will be extended by subject-specific requirements. In the area of **documentation and exchange formats**, established and commonly used standard formats are further developed in accordance with the growing number of application contexts and in line with the advancing technical possibilities of the repositories. This pertains to formats like **MEI** as the basis of source-based musicological projects (catalogs of works, scholarly music editions, genetic textual criticism, contextual in-depth exploration, etc.) and **LIDO** as the CIDOC CRM-based standard exchange format for metadata on objects of tangible culture. One of the data curation tasks deals with good-practice solutions for the **linking of highly specific and comprehensive vocabularies** with overarching **international reference vocabularies**.

Within the framework of a pilot project, a transferable solution emerges for connecting project-specific vocabularies to the multilingual AAT of the Getty Research Institute as the standard vocabulary for the indexing of art and architecture as well as to the GND. This aims to **establish an editorial workflow** for the continuous and preferably reciprocal updating of data corpora. Activities of data curation should also apply to existing scientifically valuable data that does not yet meet the requirements of quality management according to FAIR standards. This includes both actively maintained data sets and older sets that are no longer curated. These curation measures are intended to establish a quality level in sync with the current admission criteria of quality-managed repositories. The established reference process of data curation will be applicable to additional data sets. Against the background of a highly dynamic research field, **further format development tasks** e.g. for new object categories and the transfer of developed concepts into concrete

research data management will be defined over the course of the project funding period in coordination with the forum.

Overview of tasks in TA2|M3

TA2|M3|T1 Development of authority data for new art forms and multimedia works within the framework of the GND (e.g. conceptual art, media art, performance and installation). Prototypical generation of authority data for works of art on the basis of pilot databases from participating institutions with a relevant research focus. Close cooperation with M4|T1.

TA2|M3|T2 Development of an authority data model for various versions of musical works: Enhancement of the data model for authority data for musical works to enable the development of an expression level. Particular consideration given to the international connectivity of the results. Training and counseling services for data producers and providers.

TA2|M3|T3 Continuous adaptation of the data format MEI: Collection, bundling, documentation, and implementation of new format requirements resulting from new research questions. Creation of additional mappings. Coordinating the international connectivity of the developments with the MEI board

TA2|M3|T4 LIDO development and enhancement: Creation of application profiles for improved support of the creation of "rich" metadata from various research-relevant documentation contexts. LIDO schema and extension of LIDO terminology published as LOD in collaboration with TA5. Adaption of METS and/or IIIF for 2D-image-specific LIDO metadata.

TA2|M3|T5 Cross-linking of subject-spcific vocabularies with international reference vocabularies: Development of an editorial workflow for the integration of inadequatly connected vocabularies into the AAT. Factual implementation of the editorial workflow for mutual consecutive updating of the reference vocabulary and the local vocabulary in collaboration with TA5. Creation of a good-practice guide to transfer the process to other projects and data sets.

TA2|M3|T6 Format development, data curation and data rescue: Concrete tasks are to be defined with the forum (addressing e.g. event data, endangered data corpora).

Key value propositions for TA2|M3: Object-like visual-media works and work stages can be found and linked in a stable way thanks to their nuanced addressability (**KVP1**). References can be made to other entities (**KVP2)**. Comprehensive documentation and exchange formats enable cultural assets to be presented, addressed, found, and retrieved using current semantic web technologies. (**KVP3**). Results of the model project on the mapping of subject-specific and reference vocabularies are transferable to other contexts independently of their direct benefit (**KVP4**). A reusable process of data curation is established and applied in practice **(KVP5)**. Additional tasks which are yet to be defined according to the demands and priorities communicated through the forum will address current scientific needs (**KVP6)**.

User stories getting added value: 1, 2, 3, 4, 5, 7, 8, 10, 12, 13, 14, 16, 19, 20, 21, 26, 27, 28.

Key performance indicators

Aim 1: Involving the expert community in needs assessment and prioritization

- **KPI 1.1**: A thematic expert forum is set up after year 1
- **KPI 1.2**: Workshops and/or networking meetings take place at least once a year
- **KPI 1.3**: At least 15 researchers from the NFDI4Culture research domains participate in the forum
- **KPI 1.4.**: In year 3 at the latest, the forum submits coordinated proposals for additional tasks for improving quality management or its prerequisites

Aim 2: Creation of data quality guidelines (FAIR roadmap) and establishment of a quality assurance process

- **KPI 2.1**: After 3 years, data quality management guidelines for at least 3 research-related application scenarios have been drafted
- **KPI 2.2**: An audit procedure for data quality management has been established and sample applications have been made in at least 5 cases

Aim 3: Communication of the FAIR roadmap and advice

- **KPI 3.1**: Each year at least 15 projects are advised by the FAIR clearing agency
- **KPI 3.2**: After 3 years, the guidelines for data quality management serve as the basis for research data planning in at least 10 projects/project proposals
- **KPI 3.3**: After 4 years, users and providers have successfully evaluated the service

Aim 4: Research-oriented adaptation of data formats and authority data including crosslinking of specialist and reference vocabularies

- **KPI 4.1**: An (internationally recognized) authority data format for the presentation of new art forms which satisfies differentiated scientific requirements has been developed in the GND
- **KPI 4.2**: A central, internationally applicable authority data format exists for various versions of works of music and is used in various contexts of musicological research (Muscat, MEI, etc.)
- **KPI 4.3**: MEI possesses sustainable documentation including decentralized format extensions
- **KPI 4.4**: LIDO possesses published terminology in its central semantically evaluable core
- **KPI 4.5**: An updated LIDO schema version is published
- **KPI 4.6**: The forum has evaluated the interlinking of specialist vocabularies with reference vocabularies as a good-practice solution

- **KPI 4.7**: After 5 years, at least two additional projects have implemented the pilot project solution of connecting subject-specific vocabularies to reference vocabularies
- **KPI 4.8**: Curation of 2 art historical, 2 musicological and 2 data sets from media or theatre studies, in years 3-5

Task Area 3: Research tools and data services

Within the context of research into material and immaterial cultural assets, **research tools** and **data services** are instruments with which to generate, analyze, administer, distribute, publish, and make available research data in the long term. Therefore, they are crucial to research data management. At the same time, software (code, documentation) is itself **research data and thus the subject** of an RDM strategy. TA3 is aimed at users as well as developers and providers of research tools and data services. Researchers can benefit from the services of NFDI4Culture in **finding relevant tools and data services for their research** and in obtaining access to them (**Aim 1**). Researchers **formulate their needs and participate** in the conception and (further) development of research tools and data services (**Aim 2**). Researchers and institutions which develop software **obtain support through NFDI4Culture for the development, consolidation, and certification** of sustainable, interoperable tools and services (**Aim 3**). This allows for a simplified **use of relevant research tools and data services** in various environments (cloud, server, individual) (**Aim 4**).

Risk analysis and risk management

SWOT analysis for TA3 is summarized in Table 4.

Cooperation with other task areas and cross-area dependencies

The close exchange with **TA1** focuses on the (further) **development, quality assurance** and **sustainability** of data enrichment tools. Similarly, the close exchange with **TA2** also serves the (further) **development, quality assurance** and **sustainability** of tools to measure quality, to enhance database quality, and to connect to authority data and vocabularies. The close cooperation with **TA4** is aimed at developing a **concept to safeguard software and services** and to realize their sustainable operation. In the field of cross-cutting components, there is a constant exchange with **TA5**. Thus, **enabling impulses** from the other TAs or cross-consortium initiatives regarding the topic of sustainable software development, certification and software development infrastructures to be taken up or introduced. In the area of rights, a close exchange and a cooperation with **TA5** is relied on in order to stay **up to date with regard to legal developments** regarding research data on material and immaterial cultural assets. Young scientists' capability and their development of competencies includes knowledge and skills relating to the topic of sustainable software development. Accordingly, a constant exchange of knowledge with the experts of the Cultural Research Data Academy in **TA6** will occur to ensure that the

topic is given appropriate consideration in the modules and curriculum templates for teaching as well as in professional development offers.

Table 4.

SWOT analysis for TA3.

STRENGTHS creating OPPORTUNITIES	WEAKNESSES preventing OPPORTUNITIES
• Knowledge about the diverse software solutions is used to develop a detection system for low-threshold access to available subject-specific research tools and services; • Close ties to the expert communities exist. They serve the further development of needs-based research tools and services together with the participants of the forum; • The consortium can draw on many years of experience in sustainable development. Paired with an increased willingness to professionalize software development a lean, efficient advisory body for sustainable software is facilitated; • We can use our experience in low-threshold provision of software, enable greater accessibility and a much broader use.	• Due to lacking visibility of software, existing solutions are overlooked in funding proposals or research processes; • We miss the chance for a greater acceptance of subject-specific software solutions because software is developed regardless of actual needs; • Due to a lack of know-how regarding the development of sustainable software, we miss out on the opportunity to professionalize software development; • We miss the chance for appropriate technologies to be available for different usage scenarios (e.g. cloud, server in the local data centre, individual) because some software types are difficult to install and operate
STRENGTHS minimizing RISKS	**STRATEGIES avoiding RISKS**
• We can use our knowledge of diverse software solutions, our experience in low threshold provision and the close ties to the community to minimize the risks of being unable to reuse software; • We can use our contacts in the software industry to develop software packages externally if necessary. Subsequently, we can minimize the risk of unfinished (part) projects due to a lack of staff; • Through our close ties to the community we minimize the risk of developing software which does not meet the needs	• A consistent examination and evaluation of relevant software reduces the risk of producing redundant software; • Recourse to external software developers potentially prevents unfinished projects due to a lack of staff; • We consciously involve researchers in the development of software solutions to avoid developing software which does not meet the needs of the community; • The comprehensive use of technologies which facilitate simple use and access of software counter the risk of a low reuse of software

Measures

TA3|M1: Forum on sustainable software development for the field of NFDI4Culture

The forum serves to manage the entire task area. A **consistent involvement of the user community** in the work of TA3 enables feedback and suggestions to be directly obtained. Thus, insights into the systematic and needs-based design of the work program, e.g. the prioritization of the development of research tools and data services specific to NFDI4Culture, can be gained to meet researchers' needs. The forum ensures the active

participation of all participants, planning of measures that are needs based, and networking within NFDI4Culture. Additionally, it supports the **procurement and dissemination** of the offers into the scientific community. After one year, the forum is set up and has started work. Researchers of all subject communities participating in NFDI4Culture are integrated (cf. KPI 2.1, 3.1, 3.2, 3.4, 3.5).

Overview of tasks in TA3|M1

TA3|M1|T1 Coordination and organization of the forum and networking within the NFDI at large (cross-cutting topic).

TA3|M1|T2 Elaboration of recommendations for the development of research software in line with FAIR principles: guidelines with recommendations on sustainability and quality assurance, coding guidelines, documentation of existing workflows.

TA3|M1|T3 Development of a strategy for the long-term availability of research software and services: concept development in close coordination with TA4 (hosting and digital preservation) to safeguard software and services and realize their sustainable operation.

TA3|M1|T4 Elaboration of recommendations for the development of interoperable interfaces between data repositories and tools: Recommendations (in coordination with TA2).

Key Value Propositions for TA3|M1: Users can actively participate in the discussion surrounding the sustainable and interoperable development and operation of software and thus formulate their individual requirements and needs **(KVP1)**. Users receive a place for exchange with experts and can thus validate their own approaches in projects **(KVP2)**.

User stories getting added value: 6, 7, 18, 19.

TA3|M2: Consulting agency for sustainable development and operation of research tools and data services in the area of NFDI4Culture

The development and coordination of a **consulting agency** for sustainable development and operation of research tools and data services serves to support researchers and institutions on a variety of levels. In particular, assistance is provided regarding specific requirements of the subjects gathered in the NFDI4Culture consortium towards development, consolidation, operation, and certification of sustainable, interoperable research tools and data services on the basis of the **FAIR principles**. Accordingly, a **strategy for the long-term availability** of research software and services is provided. The agency will work on the basis of the recommendations developed in M1. After five years, NFDI4Culture offers are coordinated with the community, are familiar with it, and are used by it (cf. KPI 2.1, 3.1, 3.2, 3.4, 3.5).

Overview of tasks in TA3|M2

TA3|M2|T1 Setup of the consulting agency (T3-5 in coop. with the forum; cf. M1).

TA3|M2|T2 <u>Expert service and advice</u> on the basis of the recommendations from M1 and in close exchange with TA4, TA5 and TA6 and as an external point of contact.

TA3|M2|T3 <u>Recommendations for the development of research software</u> in line with FAIR principles: guidelines with recommendations on sustainability and quality assurance, coding guidelines, documentation of existing workflows.

TA3|M2|T4 <u>Development of a strategy for the long-term availability of research software and services:</u> concept development in close coordination with TA4 (hosting and digital preservation) to safeguard software and services and realize their sustainable operation.

TA3|M2|T5 <u>Recommendations for the development of interoperable interfaces</u> between data repositories and tools: Recommendations (in coordination with TA2).

Key Value Propositions for TA3|M2: users receive recommendations and advice on the sustainable and interoperable development and operation of software **(KVP1)**. Users receive advice on the long-term availability of specific research tools and data services **(KVP2)**.

User stories getting added value: <u>6</u>, <u>7</u>, <u>18</u>, <u>19</u>.

TA3|M3: Domain-specific registry of (existing) tools and services

The development of a **registry of (existing) research tools and data services** for the research on material and immaterial cultural assets serves two main purposes. To provide an easy **overview** over the offers which are of relevance to researchers in terms of research data management. Further, it serves to enable researchers to **access** these offerings with a **one-click solution**, whenever possible. This should for example **prevent parallel and dual developments**. In addition, the registry should list data repositories which correspond to the criteria previously developed in TA4. All included metadata will be issued under a **CC0 license.** All tools and services will receive a **unique identifier** enabling referencing. After five years, NFDI4Culture Community is familiar with and uses the registry and a (multi-stage) certification has been developed for software (cf. KPIs 1.1, 1.2, 1.3, 1.4, 3.2, 4.1, 4.2).

Overview of tasks in TA3|M3

TA3|M3|T1 <u>Development of a data / metadata schema and structure:</u> Coordination with the needs of other work areas (cross-cutting TA4, TA1, TA2) as well as with the registry architecture of other NFDI consortiums, provision and application of "one-click solutions" which enable simplified access to the registered research tools and services.

TA3|M3|T2 <u>Development of a step-by-step certification process:</u> criteria for sustainable and interoperable software (based, among others, on the TaDiRAH taxonomy), multi-stage certification on the basis of the catalog of criteria, transform certification workflow to an established regular operation.

TA3|M3|T3 <u>Continuous recording, maintenance and evaluation of data:</u> Documentation of available research tools / data services and assignment of PIDs down to the level of individual versions, additional recording of certified research tools and data services, connection with registries of other consortia and with services such as https://fairsharing.org/ and https://www.re3data.org (in collaboration with TA5).

TA3|M3|T4 <u>Continuous evaluation and curation of the content</u>, including continuous detailed examination and annotation of selected research tools and data services on the basis of the catalog of criteria.

Key Value Propositions for TA3|M3: Users receive an annotated overview of accessible research tools **(KVP1)**. Researchers who develop software can document their research tools via the registry, if necessary, upgrading them with certification and, thus, promoting their subsequent use **(KVP2)**.

User stories getting added value: 2, 6, 8, 15, 16.

<u>TA3|M4: Needs-based (further) development of specific research tools and data services for research in the field of material and immaterial cultural assets</u>

The measure is aimed at the (further) **development of needs-based specific research tools and data services**. They are to be provided due to their high relevance for an extended circle of users and beyond their original (and time-limited) project context. Within the framework of the forum in TA3|M1, the **needs** of the communities **will be identified** and **suggestions for funding** of specific measures will be developed, evaluated and proposed to the steering board (T1). Besides project funding and support for further development of existing research tools and data services (including in particular the implementation of **interoperable interfaces** (API), **completely new developments** are possible – provided that their **necessity has been clearly and sufficiently justified by the communities**. Various measures have been identified as possible recipients of funding but are yet to be negotiated by the NFDI4Culture community. For instance, the musical and performative representations in paintings collected in RIdIM could be connected to other repositories of paintings. The ICONCLASS seeAlso Widget (I-SAW) could be generalized and technically improved to connect a wide variety of repositories through iconographic metadata. Or an AI-based approach for analyzing music notation on sheet music could be improved to achieve better results in Optical Music Recognition. However, the **need for improvement** within three specific services has **already been established**: Firstly, a **collaborative annotation of image stocks and individual images** on the basis of IIIF (T2). Secondly, the **extension of the search and display options in image archives** and the inclusion of additional media types in image archives (T3). And thirdly, the use of a **platform for validation and conversion of data formats** which are relevant in NFDI4Culture and their connection to relevant research tools and data services (T4). After 5 years, NFDI4Culture has developed proven routines for the (further) development of tools and services and optimized the targeted development of research-related tools in close cooperation with the community (KPI 2.1, 2.2, 2.3, 3.3, 4.1, 4.2).

Overview of tasks in TA3|M4

TA3|M4|T1 Development support for the community: Concept and workflow to assess the needs, assessment and decision-making concept for proposed projects in the forum, advice and support (travel, workshops, etc.) when applying for third-party project funding, conception and implementation of (further) development projects.

TA3|M4|T2 Further development of a IIIF annotation server: Further development of the quality-assuring editorial and registration workflow as well as the rights management for fine-grained administration and dissemination of annotations as micro-publications.

TA3|M4|T3 Further development of search access and extension of the media formats in image archives: Inclusion of additional formats such as 3D, video, or audio, extension of search access to image archives (like image-, location- or time-based), integration of image similarity search.

TA3|M4|T4 Further development and extension of a conversion platform (OxGarage) for data formats: Integration of existing conversion services and tools, Implementation of missing validation and conversion services, connection of quality assurance tools in collaboration with TA2, connection of data enrichment services in collaboration with TA1.

Key Value Propositions for TA3|M4: Users can carry out targeted thematic image search in 3D and video thanks to extended search options (**KVP1**). Users can evaluate editions of music scores in various formats and also visualize the results within their own edition (**KVP2**). Users receive adequately developed research tools due to the close coordination with developers (**KVP3**).

User stories getting added value: 2, 4, 6, 12, 14.

TA3|M5: Software development infrastructures for subject-specific project contexts

Research software is developed in the most diverse contexts in the area of digital humanities: from individual researchers to agile teams or in cooperation with professional software companies. The aim of this measure is to develop **workflows and guidelines** for the use of **software development infrastructures** (e.g. development environments, project management tools, version control systems, ticket systems, user feedback systems, test environments, continuous integration/delivery, container systems, etc.). These can address the above-mentioned complex development scenarios and are adapted to the respective (standard) contexts in order to make the production process more effective and valuable. Offers must be **consolidated** and if necessary **modified, extended, or coordinated with each other**. Moreover, in agreement with the Cultural Research Data Academy in TA6, **skills** in handling software development infrastructures are to be imparted. The recommendations and infrastructures developed within the context of NFDI4Culture must be coordinated as cross-cutting topics together with TA5 and other NFDI consortiums. After five years, the offers and recommendations of NFDI4Culture are coordinated with the community, researchers who develop software are familiar with and use them. (KPI 2.2, 3.1, 3.4, 3.5 und 3.6).

Overview of tasks in TA3|M5

TA3|M5|T1 Workflows and guidelines for the use of software infrastructures for different project configurations: Evaluation, collection, coordination with the forum, development to support modular workflows, coordination with other NFDI-consortiums (with TA5).

TA3|M5|T2 Development of infrastructure modules: Adaptation of existing services and infrastructures to the requirements of researchers in the consortium as well as on the NFDI level (AAI etc.), identification configuration and development of (missing) infrastructure components.

TA3|M5|T3 Maintenance and servicing of software development infrastructures for the consortium (in accordance with the needs from the community of interest).

Key Value Propositions for TA3/M5: Users obtain information and advice on the suitable compilation and use of an infrastructure for software development **(KVP1)**. Users obtain access to the components of an infrastructure for software development which are appropriate for them **(KVP2)**. Users obtain access to versioned publication services specifically for the research data type *software* **(KVP3)**.

User stories getting added value: 6, 8, 18, 19.

Key performance indicators

Aim 1: Researchers find and can access the relevant research tools and data services

- **KPI 1.1**: Submitted research tools, data services and repositories from NFDI4Culture are recorded in the registry (TA3/M3). Criterion: After 2 years 60%, after 4 years 90%.
- **KPI 1.2**: Awareness for the registry. Criterion: After 4 years 80% of the users of our communities know the registry's offers (survey).
- **KPI 1.3**: 10% of the search queries in the registry lead to a share of the software DOI.
- **KPI 1.4**: 80% of registry visits include a detailed view.
- **KPI 1.5**: (Number) of research tools tested and reviewed. Criterion: Each year, 25 research tools are tested and reviewed according to a catalog of criteria.
- **KPI 1.6**: More than two thirds of the forum members approved the catalog of criteria for the evaluation of individual research tools and data services.

Aim 2: Researchers formulate needs and participate in the conception and (further) development of research tools and data services

- **KPI 2.1**: At least 5 researchers from NFDI4Culture participate actively in the forum. At least 10 participants from the other communities are integrated.
- **KPI 2.2**: 2 annual workshops are held on specific topics (individual research tools, data services, etc.)

- **KPI 2.3**: In the forum, 25 researchers' needs have been identified and discussed in a participatory manner and transitioned into internally or externally funded projects.

Aim 3: Researchers/institutions obtain support in the development, consolidation and certification of sustainable, interoperable research tools and data services:

- **KPI 3.1**: Use of the recommendations and guidelines for sustainable development and operation of tools and services is documented by at least 300 downloads or 30 citations.
- **KPI 3.2**: 40% of the research tools and data services verified in the registry achieve the lowest level of the certificate at least.
- **KPI 3.3**: At least 3 development projects have been realized in cooperation with participants using the "tool & service enhancement budget" through a proposal and governance process.
- **KPI 3.4**: Counseling and the mediation of experts is used at least 150 times (TA3/M2)
- **KPI 3.5**: 40 percent of research tools listed in the registry are located in a repository operated or recommended by NFDI4Culture.
- **KPI 3.6**: 80% of survey responses from our specialist communities who had indicated that they were developing software, respond positively to being familiar with our offers for software development infrastructures.
- **KPI 3.7**: 50% of survey responses among our communities of experts confirm that they use the services of the consulting agency.

Aim 4: Researchers can use the relevant research tools and data services in different easily employed environments (cloud, server, individual)

- **KPI 4.1**: 15% of the research software listed in the registry can immediately be used from within the registry ("one click").
- **KPI 4.2**: The average use of software increases by 50% after installation of a one-click solution.

Task Area 4: Data publication and data availability

There is an increasing demand in the arts and humanities for professionally supported, sustainable and uncomplicated (ideally open access) **publication of research data**, especially for complex data types which are typical for the NFDI4Culture communities. This includes solutions for **image, 3D and AV material** which are (re-)used, analyzed or annotated for research questions, for **source code** developed within the context of specific research questions, and for data from interviews, observations, statistics, measurements, mass evaluations or experiments in **various formats**. In most cases, data publications include combinations of these data types. The services of NFDI4Culture ensure that the **FAIR principles** and aspects of digital preservation as well as the often-complex **legal framework** of intellectual property law, performance rights and licenses are taken into consideration. This vision is pursued by developing services that enable researchers to publish complex data types in an **interlinked and versioned** form with guaranteed

visibility, sustainability and citeability **(Aim 1).** Complementary, researchers from NFDI4Culture's research domains are easily able to access reusable, citable and permanently available data for their research questions **(Aim 2)**. To continuously align the service portfolio with the demands of researchers from NFDI4Culture's community of interest, researchers, professional bodies and scientific institutions regularly formulate their **needs** and participate in the **conception and (further) development** of data publication services and systems for the digital preservation of research data **(Aim 3)**. To achieve these aims, publication services must be optimized in terms of standards, interfaces, usability and workflow support. At the same time, use and awareness of existing services must be increased. Last but not least, services must be adapted or extended to meet the specific needs of subject areas within NFDI4Culture which do not yet benefit from comparable offers. Specialized Information Services (FIDs), for example in art history (arthistoricum.net) and musicology (musiconn.publish), which are already well established in the research community, should act as stable and reliable partners in the newly emerging NFDI network of digital services. Accordingly, they may act as a catalyst for the development of adequate structures in areas where these services are still missing by drawing from their many years of know-how.

Risk analysis and risk management

SWOT analysis for TA4 is summarized in Table 5.

Cooperation with other task areas and cross-area dependencies

Cooperation with **TA1** will be of particular importance regarding (further) development, quality assurance, and sustainability of data enrichment tools for inclusion in data publication infrastructures and systems for digital preservation. Cooperation with **TA3** serves the development of a concept to secure software and services and to realize sustainable operation, as well as to register publication services and repositories in the registry developed in **TA3**. Close exchange with **TA2** is planned regarding the formulation of criteria aimed at an improved sustainability of research data. At the same time, these criteria serve the preparation of the pursued Core Trust Seal certification for central services and repositories. The ongoing development of legal aspects relating to the management and reuse of research data requires close cooperation with **TA5**. This will enable TA4 to support the operators of publication services as well as researchers and editors in these complex issues. A constant exchange of information with the Cultural Research Data Academy in **TA6** is planned. It is thereby ensured that the modules and curriculum teaching templates as well as the exchange with professional bodies and other users is constantly up to date regarding publication and storage possibilities. For cross-cutting components, an exchange with TA5 takes place in order for sustainability impulses from other work areas or cross-consortium initiatives to be picked up upon or integrated.

Table 5.

SWOT analysis for TA4.

STRENGTHS creating OPPORTUNITIES	WEAKNESSES preventing OPPORTUNITIES
• Existing sustainable and accepted offers for data publication and digital preservation can serve as blueprints for similar offers in subject areas lacking such resources; • Participation and needs-based services are ensured via established connections to researchers in the individual communities; • Experiences with services in art history and musicology can be used in other subject areas (performing arts & media studies); • A growing demand by researchers for digital publication forms leads to increased acceptance of Open Access and Open Data.	• High support efforts have neither been scaled nor financed so far and are thus not secured; • Functional but cost-intensive services, especially in digital preservation, have not yet had an impact throughout the community; • The difficult legal situation is massively hindering the transformation to complex data publications in Open Access; • Many researchers do not yet consider complex data publications to be equivalent to a conventional publication.
STRENGTHS minimizing RISKS	**STRATEGIES avoiding RISKS**
• Needs-based offers may be derived from the existing publication services; • Reference implementations can serve as models for further services and minimize a lack of standardization and interoperability; • Appropriate automation procedures reduce the outlay of publication processes; • Targeted advice as well as funding guidelines promoting the Open Access minimize potential rejection by researchers; • Extensive communication with researchers reduces the risk of undesirable developments and ensures the wide acceptance of services.	• Close cooperation with participants and users allows for efficient resource allocation and evaluation; • Drafting of business and funding models for central offers in order to secure sustainability; • Adoption of ongoing developments, such as digital preservation of 3D objects, to the specific requirements of tangible and intangible cultural assets; • Consulting services explain to researchers the need for digital preservation and mediate appropriate services to them.

Measures

TA4|M1: Forum on data publication & data availability

M1 will establish and operate an expert forum for "**Data Publication & Data Availability**" mainly regarding complex data types in the area of tangible and intangible cultural assets with the aim of entering into close exchange with the researchers and institutions of the NFDI4Culture subject areas which support research. Involvement of the user community at large and the professional bodies in the work of the task area can generate feedback and suggestions regarding the additional systematic and needs-based design of the work program. The forum ensures the active participation of the user community as well as of professional bodies, the needs-based orientation of the planned measures and networking within NFDI4Culture, but also with the NFDI at large. It supports the dissemination of service offerings within the research community.

Overview of tasks in TA4|M1

TA4|M1|T1 <u>Establishment and operation of the forum:</u> Implementation of 5 online questionnaires to evaluate the service portfolio and quality, to obtain expertise, as well as to optimize potential prior to face-to-face meetings. Holding 5 face-to-face meetings/ workshops for the users with evaluation of online questionnaires and drafting new proposed measures for the work area as a decision paper for the management board.

Key Value Propositions (KVP) and relevance for research community: Participation – the forum provides an interface to users. Researchers, professional bodies and research institutions can feed their subject-specific interests into the task area's work program and help with the design and provision of needs-based services **(KVP1)**. The dialogue between users and data providers is improved **(KVP2)**.

User stories getting added value: 5, 6, 16, 23, 25, 26.

TA4|M2: Consulting agency for data publication and data availability

The consulting agency supports researchers and institutions in all questions relating to the specific requirements of NFDI4Culture subject areas and provides recommendations on ensuring the publication and digital preservation of their research results. The agency enables researchers and infrastructure providers to publish complex data types in compliance with the FAIR data principles. The consultation service will closely cooperate with the FAIR clearing agency planned in TA2 and the consulting agency for sustainable software in TA3.

Overview of tasks in TA4|M2

TA4|M2|T1 <u>Setting up and operating the consulting agency:</u> Expert service in close exchange with other task areas (cross-cutting) and as an external point of contact, consulting for researchers at all stages of the publication process and with regard to data curation, reuse and (further) processing of third-party research data, mentoring and supervision of service providers.

TA4|M2|T2 <u>Recommendations and support:</u> Preparation of handouts for the researchers and service operators: recommendations for researchers in the phases of planning and publication of data, support for repository operators (software solutions, exemplary reference implementations, implementation of FAIR principles, curation services, expense and cost estimates), including the use of nestor materials, recommendations for digital preservation of complex data types, support of the Cultural Research Data Academy.

Key Value Propositions (KVP) and relevance for research community: Data Literacy – users receive subject-specific advice and recommendations **(KVP1)**. Infrastructure providers receive advice and support in setting up and operating publication and archiving offers **(KVP2)**.

User stories getting added value: 2 , 4, 5, 6, 7, 8, 9, 10, 13, 15, 17, 19, 21, 22, 24, 27, 28 .

TA4|M3: Provision of subject-specific and reusable reference implementations

All interested institutions of the disciplines represented by NFDI4Culture should be able to set up their own subject-specific services on the basis of reference implementations. The latter will cover fundamental publication and digital preservation needs, be fully compatible with the NFDI and will thus guarantee interoperability. After two years, first productive reference implementations (software packages) for data publication and data preservation will be available for all NFDI4Culture disciplines to download.

Overview of tasks in TA4|M3

TA4|M3|T1 Requirement specification and identification of appropriate data publication tools and services: Catalog of requirements (interoperability with regard to MD, API, AAI, curation) for services for the data categories and disciplines specified in the use cases, enabling researchers, for instance, to evaluate services; identification of state-of-the-art services as basis for reference implementations; Identification of missing services.

TA4|M3|T2 Gap analysis with regard to aspects such as customizability, data reusability, versioning, license and rights management, persistent identifiers, ontology-based semantic publishing, annotation support and implementation of metadata and API standards.

TA4|M3|T3 Implementation of the necessary improvements: Creating reference implementations by consolidating and optimizing the identifies services (T1) based on the gap analysis (T2).

Key Value Propositions (KVP) and relevance for research community: All researchers have access to interoperable, sustainable and trustworthy offers in the area of data publication and digital preservation (including software) **(KVP1)**.

User stories getting added value: 2, 3, 7, 15, 16, 18, 19, 21, 22, 27.

TA4|M4: Integration of relevant data publication and digital preservation services into the NFDI4Culture infrastructure

The aim of the measure is to adapt all relevant data publication and digital preservation services in the area of NFDI4Culture to the minimum requirements regarding interoperability (cf. TA2) in close cooperation with their operators and to integrate them into the NFDI4Culture infrastructure. Both NFDI-wide guidelines and international standards are taken into account. The services and the data provided through NFDI4Culture are findable for researchers and are perceived as trustworthy. Subsequently, these services increase the visibility of researchers and of their publications.

Overview of tasks in TA4|M4

TA4|M4|T1 Impact-raising optimization of relevant services: Improved visibility in search engines and with aggregators thanks to data enhancement; authorized, comprehensive linking with ORCID-IDs; OpenAIRE compatibility.

TA4|M4|T2 Improvement of the interoperability of relevant services: Implementation of metadata standards API specified in TA2 and TA3; use of authority data vocabularies for data generation/curation specified in TA2.

TA4|M4|T3 Implementation of persistent identifiers for relevant services: Assignment of persistent identifiers (e.g. DOI, handle, URN) for data publications on a sufficiently granular level; use of authoritative data identifiers for entities like persons, locations and events (e.g. GND, AAT); setting the course for LOD, semantic web applications and knowledge graphs (in collaboration with TA5).

TA4|M4|T4 Certification of the data publication and digital preservation services: Core trust seal certification for central repositories (cf. https://www.coretrustseal.org/).

Key Value Propositions (KVP) and relevance for research community: Easy access to trustworthy data publication and archiving services is achieved **(KVP1)**. Researchers can easily access NFDI4Culture data via standardized interfaces or store their own data and network it across system boundaries **(KVP2)**.

User stories getting added value: 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 21, 22, 23, 24, 25, 26, 27, 28.

TA4|M5: Establishment of digital preservation procedures for complex data types and development of operating models for digital preservation in NFDI4Culture

The aim of this measure is to develop and implement processes for the digital preservation of data types within the scope of NFDI4Culture that have been lacking up to now. This includes in particular 3D data, annotations and complex composite data types. To this end, concepts and implementations are to be developed to the maturity of a prototype and underpinned with business and operating models tailored to requirements.

Overview of tasks in TA4|M5

TA4|M5|T1 Development of methods for digital preservation of complex data types: Take stock of existing approaches at the level of the NFDI and the international level, taking into account "born-digital" cultural assets and the quality criteria for digitization developed inTA1.

TA4|M5|T2 Development of prototype implementations: Digital preservation service for multimedia, networked and versioned publications; digital preservation for 3D objects; digital preservation for annotated AV media.

TA4|M5|T3 Development of needs-based operating models for subject-specific data preservation services in established areas (image, AV) and in new areas (3D, annotations, complex data types); definition of specific service offerings, service categories, contractual aspects, pricing of services; drafting of operating and cost models with the aim of a uniform calculation basis for the chargeable provision of subject-specific data preservation services as the basis of financial planning within the framework of NFDI4Culture governance.

Key Value Propositions (KVP) and relevance for research community: Even complex data types can be preserved for the long term and thus be safely cited **(KVP1)**.

User stories getting added value: 2, 5, 6, 7, 8, 9, 10, 13, 15, 16, 17, 27, 28.

Key performance indicators

Aim 1: Researchers can publish all relevant data types in NFDI4Culture with guaranteed visibility, sustainability and citeability

- **KPI 1.1**: Number of data publications doubled (based on registered DOI)
- **KPI 1.2**: At least one subject-specific or adequate generic publication service exists for each NFDI4Culture discipline
- **KPI 1.3**: At least one data preservation offer exists for each subject-relevant type of media
- **KPI 1.4**: Increase in the number of "trustworthy" repositories by 50%
- **KPI 1.5**: Costs of enhanced publications are halved

Aim 2: Researchers from NFDI4Culture disciplines have easy access to reusable and citable data for their research questions

- **KPI 2.1**: 100% increase in the number of retrievals of published data
- **KPI 2.2**: 200% increase in retrievals of published data via APIs such as IIIF (use of applications)
- **KPI 2.3**: Number of cited data publications is tripled (e.g. calculated using CrossRef Cited-by)

Aim 3: Researchers, professional bodies and scientific institutions formulate their needs and participate in the conception and (further) development of subject-specific data publication and digital preservation services

- **KPI 3.1**: At least 5 researchers from each participating discipline actively contribute in the forum
- **KPI 3.2**: successful implementation of annual workshops on specific topics (e.g. individual publication services or digital preservation procedures)
- **KPI 3.3**: Verifiable needs of researchers are identified and discussed through participation in the forum, with at least 3 of them being annually transferred to projects which are either funded internally or by third parties

Task Area 5: Overarching technical, ethical and legal activities

The main objective in TA5 is to **guarantee an overarching and comprehensive technical infrastructure layer** for the entire consortium which is **closely intertwined with legal and data-ethical aspects**. **(Aim 1)**. The manifold research data collections on tangible and intangible cultural assets that are addressed in TA1 to TA4 as well as the tools and services of the consortium associated with it (e.g. the registry in TA3|M3) will be made accessible in TA5 by way of a uniform interface to the World Wide Web. This will happen

with measures on two levels: by implementing an overarching NFDI4Culture research information system (portal) which can be edited by the staff from all task areas (content management based) and by the modeling and implementation of a collaboratively curatable Linked Open Data service comprising an overarching knowledge graph and a terminology service based on the standardized vocabularies developed in TA2 **(Aim 2)**.

At the same time, diverse ethical and legal issues relevant for researchers and institutions that provide or use data are to be taken into account. Tangible and intangible cultural assets are subject to specific and sometimes complex legal situations which they pass on to the digital data generated from them. Legal insecurity, complex declarations of rights, as well as legal and ethical provisions which do not fully meet digital scientific needs have two main causes:

- in many places a legal framework for the scientific handling of digitized and digital cultural assets is either missing or is overly basic. TA5 aims to form arguments from the scientific world with which legislature, parliaments and governments on the national and EU level can work towards drafting legislation which takes into account the needs of science, authors and exploiters equally
- A stakeholder process will be initiated with data providers from inside and outside the cultural heritage domain.

Consensual agreements, memorandums of understanding and similar approaches will enable a reliable, possibly even automated, rights management **(Aim 3)**. With the implementation of a legal helpdesk and the preparations for a competence center, accessibility and reusability in line with the FAIR principles is achieved. This, above all, affects the areas of copyright, property and personal rights as well as the treatment of culturally sensitive objects **(Aim 4)**. At the same time, specific aspects of data ethics are included so that a legally reliable and ethically trustworthy framework of action is developed and made available to researchers and data managers **(Aim 5)**.

Risk analysis and risk management

SWOT analysis for TA5 is summarized in Table 6.

Cooperation with other task areas and cross-area dependencies

TA5 works together with all other task areas on cross-cutting topics (AAI, technical validation, legal and ethical aspects). In particular, close cooperation will be carried out with **TA1** on the semantic enrichment of data, with **TA2** on standardized ontologies and vocabularies for cultural heritage data, with **TA3** with regard to the overarching registry of tools and services and with **TA4** in the area of semantic publishing solutions. The results of these collaborations will be integrated on the modeling level into the NFDI4Culture knowledge graph and terminology service. At the same time, TA5 provides content management services via the portal for all other task areas. **TA6** will use the portal for bundling and presenting the training offers of the Culture Research Data Academy.

Table 6.

SWOT analysis for TA5.

STRENGTHS creating OPPORTUNITIES	WEAKNESSES preventing OPPORTUNITIES
• Visibility of information and resources can be increased through bundling and subsequent presentation in a comprehensive portal; • Establishing an LOD platform enables linking and interconnection of heterogeneous, distributed information systems. A uniform access point enables combined search on the information resources provided the consortium; • Bundling of ethics and rights-related inquiries prevents redundancies and creates efficiency and transparency; • Thanks to its extensive network, NFDI4Culture can involve ethics experts in a manner enabling the construction of a data-ethical framework which also has an impact outside the consortium	• Lack of network/bundling until now and a resulting insufficient transfer of knowledge regarding rights and ethics-related issues; • To date, no clear point of contact to solve legal and ethical issues; • To date, no consistent and bundled cooperation with stakeholders; • To date, insufficient communication of offers, information, and resources in the field of material and immaterial cultural assets; • To date, a lack of willingness on the part of the stakeholders to provide information systems and resources as Linked Data in accordance with W3C standards.
STRENGTHS minimizing RISKS	STRATEGIES avoiding RISKS
• LOD technologies prevent potential additional effort and expense in the data integration of heterogeneous information resources; • A web portal in connection with the knowledge graph and terminology service prevents missing visibility and enables findability; • A legal helpdesk minimizes the risk of individual (and possible divergent) legal solutions; • Involving ethic experts assures compliance regarding data ethical questions and can have external impact on other consortia	• The legal helpdesk actively connects itself with the corresponding institutions of other NFDI consortia and the international professional community; • Prompt search and approach of appropriate members to the ethical advisory panel; • Assured resource allocation and clear task assignment for the implementation of the NFDI4Culture portal; • Assured resource allocation for the implementation, management, and maintenance of the LOD platform

Measures

TA5|M1: Cross-area and cross-cutting infrastructural tasks

In this measure, **cross-cutting technical concerns** which are not restricted to one individual task area and which require technical implementation as well as coordination are gathered. These include the creation of a uniform authentication and authorization infrastructure for NFDI4Culture research data as well as the development, coordination and implementation of common technical data quality criteria in close cooperation with **TA2**. In future, cross-cutting technical tasks which are yet to arise will have to be allocated sufficient means and mechanisms.

Overview of tasks in TA5|M1

TA5|M1|T1 Joint authentication and authorization infrastructure across the NFDI: Conception and implementation of an AAI procedure across the consortium with the potential option of single-sign-on access in coordination with all other NFDI consortiums.

TA5|M1|T2 Determination, implementation, monitoring and securing of general quality criteria across the NFDI at large to determine and monitor technical data quality in agreement with all NFDI consortiums (in close collaboration with TA2).

TA5|M1|T3 Future infrastructural cross-cutting topics in the NFDI which are still unknown at the time of writing must be addressed, implemented and permanently solved using appropriate means (cf. Berlin Declaration).

Key Value Propositions (KVP): Users receive simple and uniform access to NFDI4Culture information resources **(KVP1)**. Users obtain information and reassurance of the technical data quality of NFDI4Culture information resources **(KVP2)**. Users receive support in the event of future technical issues which may arise within the consortium's framework **(KVP3)**.

User stories getting added value: 8, 13, 18, 21, 22, 27.

TA5|M2: NFDI4Culture information portal

The measure comprises planning, implementation, management, and sustainable operation of a research portal for users to access all of NFDI4Culture's services and resources. This also applies to a uniform web-based access of the consortium's data repositories and information systems. Coordination of the platform occurs centrally through **TA5**. Content is contributed and maintained by the task areas. The measure is also aimed at making NFDI4Culture's content and offerings available in European infrastructures.

Overview of tasks in TA5|M2

TA5|M2|T1 Conception, implementation and operation of the NFDI4Culture information portal: Information architecture, UX, software architecture, content management-based implementation, development and implementation of an editorial and operational concept in cooperation with all TAs.

TA5|M2|T2 Web-based access to NFDI4Culture information resources: Conception and planning, technical implementation, operation and maintenance.

TA5|M2|T3 Integration of portal content in European research infrastructures: Creation of CERIF-mapping/application of CERIF ontology to the portal content, implementation of a CERIF-compatible HTTP-interface, harvesting of portal content by European and international providers.

Key Value Propositions (KVP) and relevance for research community: Users gain access to all information resources and services related to NFDI4Culture through a web

portal **(KVP1)**. Users can obtain information about NFDI4Culture and the information resources and services it administers **(KVP2)**. Users obtain additional information through the portals of European research infrastructures **(KVP3)**.

User stories getting added value: [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [21], [22], [23], [24], [26], [27], [28].

TA5|M3: Linked Open Data 4Culture

The measure comprises the conception and implementation of a collaboratively curatable Linked Open Data service for the consortium including an overarching knowledge graph and terminology service. The central strategy is to make this service collaboratively curatable by the users of the consortium. This will be implemented by the use of a Wikibase instance as core of the LOD platform. NFDI4Culture's close cooperation with Wikimedia e.V. will help with the coordination and alignment of data integration activities, authority files (GND) and hubs (Wikidata) together with TA2. Related to this, Wikimedia will give support with Wikibase and LOD onboarding workshops for the users during which selected data sets will be lifted to structured data for integration into the collaborative knowledge graph. The knowledge graph itself will observe W3C standard technologies in order to enable overarching search access on the distributed research data collections internally, across the consortium, and beyond.

Overview of tasks in TA5|M3

TA5|M3|T1 LOD working group: Coordination with other TAs and invitation of external experts and stakeholders to the working group in close cooperation with Wikimedia.

TA5|M3|T2 Ontologies for NFDI4Culture: Survey of relevant metadata standards and corresponding ontologies, design, implementation, and evaluation of additionally required ontologies in coordination with TA2 (vocabularies), TA3 (registry).

TA5|M3|T3 Survey of Linked Data resources and generation of Linked (Open) Data for existing NFDI4Culture data and metadata based on ontologies in coordination with TA2; availability, publication and documentation in coordination with TA4 and the institutions responsible for the information resources.

TA5|M3|T4 Implementation of the NFDI4Culture knowledge graph: Examination of centralized and distributed solutions for the implementation; set up and permanent operation of a Wikibase instance; Determination of an implementation strategy (distributed vs. centralized index) in coordination with TA2 and TA4 and implementation; connection to external NFDI knowledge graphs by linking to authority data (GND) and hubs (Wikidata).

TA5|M3|T5 Integration of the NFDI4Culture knowledge graph into the NFDI4Culture portal: design and implementation of a web-based user interface; definition of use cases and demonstrations of the knowledge graph; integration into the information platform.

Key Value Propositions (KVP) and relevance for research community: Users can participate in the curation of the NFDI4Culture knowledge graph (**KVP1**). Users receive

access to detailed information resources, metadata and associated background knowledge in a machine-readable form **(KVP2)**. Users can effectively implement advanced information systems (in particular retrieval and recommender systems) on the basis of the NFDI4Culture knowledge graph **(KVP3)**. Users can link and effectively integrate their own information resources with other information resources via a uniform interface **(KVP4)**. NFDI4Culture information resources become part of the Web of Data, thus, effectively implementing all FAIR principles (**KVP5**).

User stories getting added value: 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 18, 21, 22, 23, 24, 26, 27, 28.

TA5|M4: Legal Helpdesk

By setting up the legal helpdesk, a central point of contact for researchers as well as for institutions which provide and use data is established. Researchers and institutions need to be supported in keeping research data as open and easily accessible as possible and compliant with the FAIR principles by clarifying legal issues and legal uncertainty. To this end, researchers and institutions receive legal assistance and information. First and foremost, the helpdesk refers them to the relevant contact persons (e.g. rights holders, other organizations) and draws their attention to relevant information and regulations. For this purpose, legal information and regulations which already exist are jointly identified and reviewed from an expert legal point of view. At the same time, the helpdesk collects legal queries and arranges them into legal problem areas allowing for legal problems in individual research data sets to be prioritized. This will enable legal expertise to be developed based on actual requirements and for a future legal Culture Competence Centre to be established.

Overview of tasks in TA5|M4

TA5|M4|T1 Setting up, coordinating, and organizing the legal helpdesk: Coordination with the other TAs (especially TA2) with regard to the implementation of the FAIR principles; legal support for queries by researchers and from the consortium.

TA5|M4|T2 Identification and activation of a legal network: Identification of relevant legally active stakeholders in and outside science; establishment of communication channels for the purpose of clarifying rights.

TA5|M4|T3 Documentation of enquiries according to legal categories (clustering); prioritization of legal problems.

TA5|M4|T4 Specialist legal support (legal analysis) and assessment: Evaluation of legal information which has evolved due to individual scientific stakeholders; review and evaluation of legal information and procedures from a scientific perspective.

TA5|M4|T5 Legal basis for training offers: training offers with legal content, in strong coordination with the Cultural Research Data Academy TA6.

TA5|M4|T6 <u>Coordination with other NFDI consortiums</u> to establish legal standards for the NFDI.

Key Value Propositions (KVP): With the legal helpdesk, researchers as well as institutions that provide and use data receive a legally competent point of contact **(KVP1)**. Researchers and scientific institutions are referred to reliable legal information and contact partners **(KVP2)**. Legal expertise, which is occasionally initiated by scientific institutions, is mutually examined and made available in a bundled form (**KVP3**). NFDI4Culture continually contributes legal expertise to the NFDI as a whole, above all in the fields of copyright, intellectual property (IP) and personal rights. It also actively participates in handling legal issues in the NFDI as a whole **(KVP4).** In educational and professional development legal content is included **(KVP5)**.

User stories getting added value: <u>3</u>, <u>7</u>, <u>8</u>, <u>9</u>, <u>10</u>, <u>15</u>, <u>17</u>, <u>19</u>, <u>20</u>, <u>27</u>.

TA5|M5: Stakeholder process to improve legal certainty in the research context

In light of the manifold interests, ethical aspects (cf. M6) and only rudimentarily existing legal requirements for research data, this measure aims to establish a forum with a variety of stakeholders. These include collecting societies (VG Bild-Kunst, GEMA/VG Musikedition, Cultural Commons Collecting Society, GVL, and others), monument preservation authorities, churches, museum and archive associations, and the stakeholders and associations of the cultural and creative industries, such as publishing houses. At regular meetings and/or conferences, the needs of science in the digital age should be made apparent. This procedure aims to jointly develop codified foundations to regulate rights of use (rights management) which enable legally compliant procedures for technical administration, public accessibility, and an exchange of culture-related research data which is as unrestricted as possible. Furthermore, NFDI4Culture puts forward criteria and proposals for a legal framework to ministries, governments and parliaments, for example in view of the foreseeable amendment of the scientific copyright law in 2023, or the still outstanding implementation recommendation of the EU Copyright Directive.

Overview of tasks in TA5|M5

TA5|M5|T1 <u>Stocktaking:</u> Review and collection of existing agreements and procedures.

TA5|M5|T2 <u>Conception of a stakeholder process:</u> Evaluation of experiences at scientific institutions; identification of potential dialogue partners, setting objectives; clarification of the legal framework; internal and external legal obligations; mandating and legal capacity to act as a consortium and beyond.

TA5|M5|T3 <u>Coordination and organization:</u> Preparation and organization of dialogue events; point of contact for rights holders, collecting societies, and other stakeholders.

TA5|M5|T4 <u>Legal development:</u> Identification of relevant legal developments (national, EU, and international legislation); drafting lines of argumentation and model cases in order to

work towards legal improvements; comments and participation in legislative processes; cooperation or exchange with competent supervisory authorities.

TA5|M5|T5 Cooperation with stakeholders: Cooperation with relevant stakeholders such as the German National Library (DNB), Deutsche Initiative für Netzwerkinformation (DINI), Creative Commons, Research Data Alliance, Data Ethics Commission as well as initiatives such as the European Network for Copyright in Support of Education and Science (ENCES), focus initiative "Digital Information" of the Alliance of German Science Organizations, Action Alliance "Copyright for Education and Science", Copyright for Creativity (C4C); representation of the NFDI4Culture Community in committees.

Key Value Propositions (KVP): The previous commitment of scientific institutions to the clarification and management of rights is identified and integrated (**KVP1**). Dialogues on the usage requirements of research should be initiated between institutions providing and those using data or with their associations and should lead to agreements in which interests are fairly balanced (**KVP2**). NFDI4Culture is committed to improving the legal framework and works with politicians and legislators to do so (**KVP3**). NFDI4Culture represents its community in relevant national and international committees (**KVP4**).

User stories getting added value: 3, 7, 10, 18, 19, 25, 27.

TA5|M6: Advisory Panel for Data Ethics

The measure is aimed at providing researchers and their institutes with a data ethical framework of action. Relevant aspects should be identified in science ethics, the history of science, and the newly emerging digital ethics, for instance regarding works with historically and culturally sensitive contexts and their digital representations. At the same time, the possibility of adaptation and further development of existing ethical standards from the disciplines involved in NFDI4Culture must be considered. To do so, an advisory panel is created which comprises experts from the consortium as well as third-party experts. Its function is to ensure that ethical questions are treated in sufficient depth within the consortium and to provide input regarding further issues. Panel membership is to be adjusted over the course of the project.

Overview of tasks in TA5|M6

TA5|M6|T1 Initiation of the panel of experts: Appointment of suitable members followed by organizational support in terms of meetings and follow-up. Preparation and organization of external outreach and dialogue.

TA5|M6|T2 Legal framework of ethical recommendations for action: Examination of the legal framework for ethical options for action, in particular, for international research contexts.

TA5|M6|T3 Drafting of ethical recommendations for action: Moderation of the process of drafting ethical recommendations for action;sSearch for suitable options for publication and organization of workshops to disseminate the results.

Key Value Propositions (KVP): A data ethical framework of action is to be developed by a panel of experts within NFDI4Culture (**KVP1**). Ethical recommendations are drafted within the applicable legal framework, in particular for international research contexts (**KVP2**).

User stories getting added value: <u>15</u>, <u>17</u>, <u>19</u>, <u>22</u>, <u>27</u>, <u>28</u>.

Key performance indicators

Aim 1: Guarantee a permanent, cross-cutting technical horizon for the entire consortium which includes legal and data-ethical aspects

- **KPI 1.1**: Legal helpdesk and technical coordination office (cf. TA7|M2) is set up after 6 months and leads to a proven network among the consortium partners as well as between the consortium and external parties.

Aim 2: Research data on tangible and intangible cultural assets and associated services can be found and accessed in the World Wide Web via a uniform interface

- **KPI 2.1**: A first version of the NFDI4Culture portal is implemented after 6 months. Annual reviews of user interaction/conversions on the portal show a growing acceptance and awareness of the offer within the community.
- **KPI 2.2**: The NFDI4Culture portal continually integrates new offers and resources from the consortium. Continual integration of these resources into European infrastructure (EOSC, OpenAIRE)
- **KPI 2.3**: (Standardized) ontologies have been developed as a basis for the NFDI4Culture knowledge graph and are established after 24 months. The number of available information systems/data collections which have implemented the provision of information resources employing Linked Data principles and the degree of this implementation increases annually
- **KPI 2.4**: Verifiable increase in cross-networking of the NFDI4Culture knowledge graph and terminology service with other consortia as well as with resources from outside NFDI

Aim 3: Consensually achieved consortium agreements, memoranda of understanding enable a reliable and potentially automated management of rights

- **KPI 3.1**: Initial agreements are concluded after 2-3 years

Aim 4: Implementation of a Legal Helpdesk (copyright, property and personal rights, treatment of culturally sensitive objects)

- **KPI 4.1**: Legal helpdesk for researchers and institutions that provide and use data is set up and is operational after 6 months
- **KPI 4.2**: Legal helpdesk has dealt with all incoming matters

Aim 5: Development of a legally reliable and ethically trustworthy framework for researchers as well as for service providers

- **KPI 5.1**: Ethical Advisory Panel is established after 10 months
- **KPI 5.2**: Guidelines for dealing with data ethics in the context NFDI4Culture have been drafted

Task Area 6: Cultural Research Data Academy (CRDA) – Professionalization, qualification and training

The **overall objective** is to develop subject-specific **"cultural" Data and Code literacy** as well as **computational thinking** for the humanities and cultural studies in all target groups and with regard to the FAIR4S-framework of the EOSC. To promote our endeavor, NFDI4Culture will establish an interdisciplinary, decentralized **Cultural Research Data Academy (CRDA).** This includes addressing **generic aspects** as well as **subject-specific** Data and Code Literacy. The latter might pertain to a framework and specific criteria for training, teaching modules, workshops, labs and schools appropriate to the handling of tangible and intangible cultural assets. In detail this regards fundamental expertise in the application of the **CIDOC CRM** in art history, of **MEI** in musicology, the **development of interdisciplinary connected data standards for performing arts**, as well as the stronger participation of film/media studies in standardization initiatives within the framework of GND and Wikidata. Also, **the ability to assess the performative aspects** of the creation, annotation, analysis of cultural research data and research software is to be addressed. The CRDA faces several challenges:

Participants in NFDI4Culture and others institutions already provide a (growing) amount of high quality offers for increasing Data and Code Literacy, e.g. workshops on cooperative metadata handling (SPK), data carpentries and hackathons as well as MOOCs (TIB), (international) summer schools on digital methods (AWLM) or musicology (Edirom, UPB), master studies on Cultural Data Studies (UMR) etc. A first evaluation shows that other initiatives have e.g. set up a registry for courses on Digital Humanities, e.g. https://registries.clarin-dariah.eu/courses or e.g. set up a network (http://www.parthenos-project.eu/portal/trainingsuite). However, training options for GLAM employees, for multipliers and for the citizen science community are lacking, as well as acknowledged quality criteria or frameworks. Therefore, it is **necessary to thoroughly evaluate and identify existing training offers compliant with the FAIR4S-framework and to then bundle, network, and publish them** in a centralized and needs-oriented manner.

- In ongoing discussions with the community and the target groups it has become apparent that in some cases **new, innovative offers to fill identified gaps must be developed**.
- The CRDA will also tackle the lack of an established teaching canon: A **clear and viable conceptual framework** for cultural Data and Code Literacy, **clear recommendations**, as well as **quality assurance measures** for teaching and training content in close reference to the FAIR4S-framework are to be created. Framework, recommendations and measures need to be continuously developed to

a further extent by NFDI4Culture in a highly responsive process with users planning to implement these materials in their teaching practice (multipliers).

- Accordingly, appropriate **opportunities to articulate the need** for skills development **in a central and easily accessible place** is so far missing, as is the possibility of receiving accessible and personalized **support**: for example, regarding suitable offers, or current developments concerning management of a specific type of data.
- Also, a lack of support concerning **institutional surroundings** is to be encountered: Researchers and lecturers also must deal with institutional frameworks that delay and/or prevent the integration of the necessary Data and Code Literacy content into training concepts and curricula.

Against this background, TA6 **aims**

1. to enable target groups and in particular researchers to **actively contribute their needs and suggestions** to the CRDA;
2. at active and participatory development of **quality criteria, framework, and recommendations** for cultural data and code literacy;
3. at CRDA to **provide researchers, lecturers and GLAM employees with easy access** to **certified, FAIR4S-compatible and up-to-date training offers**, including self-managed and network-based skills development;
4. at the target groups **actively *using*** framework, recommendations and access, i.e. to access **practice-based concepts, teaching material**, and **recommendations** as well as **well-directed modular course offers for on-site academic teaching, in-house training, summer schools, pre-conference workshops**, etc., and to be supported in a precise and personalized manner.

Risk analysis and risk management

SWOT analysis for TA6 is summarized in Table 7.

Cooperation with other task areas and cross-area dependencies

A close exchange on content is planned with all other task areas **(TA1–5)** in order to establish the current state of data capturing and enrichment **(TA1)**, to incorporate emerging standards, vocabularies, FAIR principles, data quality, and data curation **(TA2)** with regard to current tools **(TA3)**, to include data publications and availability **(TA4)** as well as integrating specific consulting on legal and ethical issues in trainings and other CRDA offerings **(TA5)**. Moreover, close cooperation with other NFDI consortiums concerned with Data and Code Literacy is planned, especially in the humanities, but also with NFDI4Ing. Contributions to cross-cutting developments provided by the NFDI directorate regarding Data and Code Literacy is also planned (**TA6|M1|T1**).

Table 7.

SWOT analysis for TA6.

STRENGTHS creating OPPORTUNITIES	WEAKNESSES preventing OPPORTUNITIES
• By bundling experience and existing offers in the CRDA findability and accessibility of taining offers are enhanced; the institutional level can be included; joint framework and recommendations can be (further) developed; • A sound practical relevance of offers for the development of Data Literacy skills can be achieved; • When combining forces, the CRDA can quickly react to specific needs and invest on new offers fitting those needs	• To date, insufficient sustainability in community building prevents a framework; • Didactically sustainable design for the development of skills is only partly existing. Moreover, the focus lies primarily on the classic institutional offers; • The link between development of skills and practice is too marginal; • The development of skills is not up to date. This, e.g., applies to the need to stay in sync with current versions of standards, data quality etc. • Offers for skill development are too spread out and a central point of access is missing; • Institutions hardly react to initiatives for the inclusion of content on Data and Code Literacy in curriculum and education formats
STRENGTHS minimizing RISKS	STRATEGIES avoiding RISKS
• Access to specialized offers is facilitated by knowledge of the various offers bundled in the consortium (one point of contact/information); • Close networking within the consortium allows for practical relevance to be established.	• Community building within the measures links specific offers to each other as well as to the needs of the users; • The highly interactive approach in the forum helps to avoid the risk of offers not being integrated, accepted, and supported within the community.

Measures

TA6|M1: Forum on Cultural Data and Code Literacy

M1 provides the basis for the coordination and the participatory framework of the task area. The aim of the measure is to create a communication platform operating as the **Cultural Research Data Academy (CRDA)** which both provides and facilitates access to knowledge and the development of skills for the community. As users are still often institutionally embedded the forum also aims at providing further impetus for organizational adaptation to foster the integration of digital competencies. By networking with similar offers provided by other consortiums, the measure also aims at making innovative impulses as well as cross-disciplinary and specific knowledge available and accessible to the communities. After five years, the forum is established and active processes for joint

work on framework, quality criteria, and recommendations (TA6|M2) are available (cf. KPIs 1.1, 1.2).

Overview of tasks in TA6|M1

TA6|M1|T1 <u>Launch, activation and moderation of a forum:</u> Gathering potential participants, with particular consideration to the professional bodies; operating the forum through at least one annual meeting and optional additional event formats; moderation and design of the forum, possible thematic sub-forums.

TA6|M1|T2 <u>Coordination and networking:</u> NFDI and DCL networks and projects: Point of contact for other NFDI4Culture TAs with concern to training; topic-related coordination and synergy effects with other consortia and with the respective cross-cutting offers of the directorate; active networking and coordination with existing and future networks and projects on data and code literacy and computational thinking.

Key Value Propositions (KVP) and relevance for research community: Exchange on specific questions in relation to skill development is facilitated through the participatory approach **(KVP1)**. In addition, researchers, lecturers and GLAM employees are given the opportunity for interact beyond the boundaries set by their own disciplines **(KVP2)**.

User stories getting added value: [2](#), [5](#), [6](#), [15](#), [16](#), [26](#).

<u>**TA6|M2: Needs, CRDA quality criteria, framework and recommendations**</u>

The measure aims, on the one hand, to identify, collect, and continually monitor the topics, competencies and skills which are needed for dealing with research data and software in the field of material and immaterial cultural assets. Together with the communities, needs are identified and systematized. On the other hand, M2 also serves the joint development of FAIR4Scompatible quality criteria, as well as a conceptual framework for the teaching of data and code literacy in the curricula of participating disciplines. Also, hands-on recommendations on content and didactic and methodological concepts will be produced and sustained. **After five years**, productive methods for identifying and monitoring the needs of the communities have been developed. Also, viable quality criteria for educational offers in the field of data and code competence have been developed and accepted by the community. In addition, a conceptual framework for data and code literacy is developed and adapted to the requirements of the disciplines. Together with hands-on-recommendations, these processes can be sustained digitally (cf. KPI 1.2, 2.1, 2.2, 2.3).

Overview of tasks in TA6|M2

TA6|M2|T1 <u>Identifying, monitoring and systematizing cultural data and code literacy needs:</u> Activating the community, identifying and monitoring needs, classification and transfer to taxonomies, publication.

TA6|M2|T2 <u>Developing FAIR4S-compatible quality criteria based on needs</u> (TA6|M2|T1), on existing offerings (cf. TA6|M3), on close cooperation in the forum (cf. TA6|M1), publication and regular assessment.

TA6|M2|T3 <u>Developing a conceptual framework for the teaching of data and code literacy</u> in cooperation with other consortia and initiatives in the field of cultural data literacy education, publication, further adaptation.

TA6|M2|T4 <u>Developing didactic and content-related recommendations for teaching and learning</u> concepts for trainings on cultural data and code literacy.

Key Value Propositions (KVP): Researchers, lecturers, and GLAM employees can articulate their needs (**KVP1**). FAIR4S-compatible quality criteria for training offers allow for clear orientation for users (**KVP2**). Lecturers can use the framework and recommendations to easier conceive their specific courses and to make sure that content is up to date (**KVP3**).

User stories getting added value: 2, 3, 5, 7, 15, 16, 26.

TA6|M3: CRDA portfolio – 4Culture Training Offers

With this measure, **CRDA** aims at **bundling, the promotion, quality assurance, new and further development, and distribution of offers** for data and code literacy education. Together with the participating institutions and other relevant organizations in the field of tangible and intangible cultural assets, offers will be identified, compiled and systematically presented in a **CRDA portfolio**. Existing offers are assessed with respect to quality criteria (M2|T2) and, if desired, certified as part of a standardized process. NFDI4Culture will also make funds available to enable participating institutions to develop new or innovative training offers in the absence of adequate offers for documented needs. Institutions that offer training are also supported in the development and establishment of sustainable models for financing and organizing these offers. As two specific needs have already been identified, specific trainings for GLAM employees and a NDFI4Culture mentoring network will be set up immediately by two co-applicant institutions. **After five years**, the CRDA provides a thorough and quality assured portfolio of training offers on cultural data and code literacy. Information on and access to these offers can easily be kept up to date. Certification of offers is widely accepted in the community (cf. KPI 3.1, 3.2, 3.3, 3.4).

Overview of tasks in TA6|M3

TA6|M3|T1 <u>Identifying, bundling, evaluation and publishing</u> cultural Data and Code Literacy offers.

TA6|M3|T2 <u>Quality management (QM) and certification:</u> Supporting participants in their training courses and teaching offers; certification according to quality criteria developed in M2.

TA6|M3|T3 <u>Development of new and innovative offers:</u> identify new skills needed by the community, facilitate conceptual development by providing resources, integrate new offer to CRDA portfolio and foster use of new offer.

TA6|M3|T4 <u>Specific offer I:</u> CRDA-trainings for GLAM-employees (responsible: SPK) by conceptual development, first training courses, subsequent evaluation and adaptation.

TA6|M3|T5 <u>Specific offer II:</u> CRDA mentoring network (responsible: TIB) by development of structure and of appropriate forms for acquisition of mentors/mentees, by activating and implementing network.

TA6|M3|T6 <u>Organizational and financial modalities for teaching Data and Code Literacy:</u> Development of models for reuse and further development of offers, support for organizational and financial implementation

Key Value Propositions (KVP) and relevance for research community: The research community and the community beyond science can easily overview and access offers that highly contribute to their ultural Data and Code Literacy **(KVP1)**. When the necessity for new skills is visible and articulated e.g., via the forum, the CRDA can quickly facilitate additional training to develop these skills **(KVP2)**.

User stories getting added value: <u>3</u>, <u>5</u>, <u>15</u>, <u>23</u>.

TA6|M4: CRDA Helpdesk – Consulting and support

With M4, the CRDA aims at providing comprehensive consulting and support for participants and users seeking to improve and develop their cultural Data and Code Literacy. The CRDA consulting and support helpdesk serves the scientific community involved in NFDI4Culture. However, it is also open to the citizen science community that is interested in working with data on tangible and intangible cultural assets. Users of all kinds can regularly contact the CRDA consulting and support helpdesk in various matters. For example, regarding appropriate content on data quality or long-term availability when conceiving an own training program. The CRDA helpdesk might be consulted with questions on learning cooperative metadata handling (offered by SPK Berlin), participating in the EDIROM summer school in Paderborn, or a culture data hackathon organized by the TIB in Hannover, or the digital humanities virtual laboratory at the LMU München. Even basic questions on educational materials or concepts for a specific topic can be answered at the helpdesk, that, in itself, will partake in a permanent process of self-evaluation and optimization. The CRDA helpdesk will also provide support and advice for the institutional process of implementing cultural Data and Code Literacy training (modules) in the respective institute, museum, archive or otherwise constituted community (cf. KPI 4.1, 4.2, 4.3).

Overview of tasks in TA6|M4

TA6|M4|T1 <u>One point of information:</u> cultural Data and Code Literacy training offers/options for the NFDI4Culture information portal: Conception in close reference with existing platforms, strategic integration of content in the NFDI4Culture portal (TA5)

TA6|M4|T2 <u>One point of contact:</u> Cultural Data and Code Literacy consulting and support: procuring participants' offers; consulting with regard to the offers and their quality (cf. TA6|M2); advice for content and organizational development and design of training courses, workshops, teaching modules, MOOCs, schools, etc.

TA6|M4|T3 <u>Process assistance in the implementation of offers:</u> collect good practices for implementing digital change; develop and publish precise models for implementing new training offers, consultation on procedural aspects.

Key Value Propositions (KVP): Target groups receive simplified access to core knowledge and information **(KVP1)**. Precise and individually customized offers are conceived **(KVP2)**.

User stories getting added value: <u>1</u>, <u>2</u>, <u>3</u>, <u>5</u>, <u>7</u>, <u>14</u>, <u>15</u>, <u>16</u>, <u>23</u>, <u>26</u>.

Key performance indicators

Aim 1: The target groups and in particular the researchers can actively contribute their needs and suggestions to the CRDA

- **KPI 1.1**: successful establishment of a cross-disciplinary community of experts which is networked and which actively participates in further development of cultural Data and Code Literacy. Criterion: yearly participatory event of the forum, organized by the CRDA.
- **KPI 1.2**: Subjective perception of the opportunity to articulate needs. Criterion: the majority of forum participants indicate that they know how to formulate and contribute their needs.

Aim 2: Active and participatory development of CRDA quality criteria, framework and recommendations for cultural data and code literacy

- **KPI 2.1**: Quality criteria for offers. Criterion: quality criteria are developed in accordance with the FAIR4S framework and are familiar to the majority of forum participants.
- **KPI 2.2**: Subject-specific framework for skills development. Criterion: framework is initiated, mentored and agreed on by the forum.
- **KPI 2.3**: Recommendations. Criterion: at least 3 recommendations have been initiated, drafted by means of a participatory process; use is documented with at least 30 cases of application.

Aim 3: CRDA provides researchers, lecturers, and GLAM employees with easy access to certified, FAIR4S-compatible and up-to-date training offers, including self-managed and network-based skills development

- **KPI 3.1**: CRDA-portfolio compatible with FAIR4S-framework. Criterion: the majority of forum participants know about the CRDA-portfolio.
- **KPI 3.2**: Training offers for GLAM-Employees. Criterion: a series of training offers for GLAM employees has been developed and executed at least three times.
- **KPI 3.3**: Active and effective network of mentors. Criterion: a mentor network is formed by members of the forum. It has mentored at least 25 mentees.

- **KPI 3.4**: New and innovative training offers. Criterion: At least three new and innovative training offers have been identified and set into practice by NFDI4Culture participants.

Aim 4: The target groups use CRDA framework, recommendations and access, i.e. they are supported as precisely and personalized as possible

- **KPI 4.1**: Reception of framework and recommendations. Criterion: the majority of the forum participants are familiar with the subject-specific framework and recommendations.
- **KPI 4.2**: Use of CRDA-portfolio compatible with FAIR4S-framework. Criterion: at least 50 users have been referred to one of the training offers in the CRDA-portfolio via NFDI4Culture.
- **KPI 4.3**: Consulting for multipliers. Criterion: at least 25 lecturers and other multipliers have been consulted on the further development of their teaching and learning offers.

Task Area 7: Governance and administration

TA7 brings together all administrative and coordinative activities and bundles financial controlling, contract management, reporting, governance operations, incentives for participation and inward-outward cooperation in RDM, dissemination, community engagement and outreach into adequate measures. The general aim of TA7 is to ensure the efficient operation of the consortium with smooth governance processes and a sound cross-area technical management that secures the convergence of infrastructural approaches on the consortium level and the NFDI level (**Aim 1**). A second aim is to enable knowledge pooling and exchange during the work on NFDI-wide cross-cutting topics with other NFDI consortia (**Aim 2**). NFDI4Culture will only find acceptance in its community of interest if it stays open for the integration of new and innovative contributions by its users (natural persons and institutions alike). Therefore, incentives and defined admission processes need to be set up that enable the future participation in the consortium on the basis of innovative ideas for identified but not yet covered demands (**Aim 3**). A professional innovation management will be the basis for a continuous increase in NFDI4Culture's user community and a warranty for the long-term acceptance and sustainability for the services developed by the consortium (**Aim 4**).

Risk analysis and risk management

SWOT analysis for TA7 is summarized in Table 8.

Cooperation with other task areas and cross-area dependencies

TA7 is overarching and bundles coordination and administration activities of all other task areas.

Measures

TA7|M1: Administrative Coordination Office (ACO)

As described in detail in the chapter "Organizational structure", the ACO is responsible for the administration of the consortium (financial flows, controlling, reporting, contracts) including the event management for all task areas and the preparation and proper execution of all governance related tasks. The ACO is set up as a centralized-decentralized structure with the lead staff (spokesperson, scientific coordinator and financial coordinator) at the AWLM and administrative staff for each task area at each of the co-applicant institutions. Digital communication channels (messengers, web conference calls etc.) and collaborative documentation platforms will be used by the ACO to provide everybody with a transparent and comprehensible reporting system. The centralized-decentralized structure and digital workflows have already been tested very successfully during the application phase.

Table 8.

SWOT analysis for TA7.

STRENGTHS creating OPPORTUNITIES	WEAKNESSES preventing OPPORTUNITIES
• Structure and workflows for the coordination offices as well as governance events have already been successfully tested during the application phase; • Concrete requests and project ideas for the Culture RDM Kickstarter already exist and can immediately be put into practice; • Concrete working groups on several cross-cutting topics have already been agreed with cooperating consortia.	• Delays in administrative routines or governance procedures; • Lack of transparency during decision making processes; • Failure to ensure proper and timely reporting; • Lack of participation in the governance bodies prevents the ability to come to decisions; • Conflicts arise and are badly managed or unresolved.
STRENGTHS minimizing RISKS	**STRATEGIES avoiding RISKS**
• Academic societies are fully integrated into the governance and dissemination channels ensuring a permanent connection between users and providers; • Long-standing expertise in fund administration ensures a smooth administration of the consortium in compliance with the funding policies; • Successful workflows between the applicant and co-applicants are already in operation for more than a year.	• Minimize financial risks with regular (external) audits; • Evaluation and optimization of governance bodies and procedures every two years (2022, 2024); • Ensure full transparency on all decisions through open documentation; • Mandate stand-in persons to always ensure decision-making ability; • Establish a sound conflict management.

Tasks in TA7|M1

TA7|M1|T1 <u>General coordinative action</u> in each task area (e.g. organization of forums etc.).

TA7|M1|T2 <u>Coordination and controlling</u> of the financial flows.

TA7|M1|T3 <u>Coordination and operation</u> of the reporting.

TA7|M1|T4 <u>Contract management</u>.

TA7|M1|T5 <u>Coordination of all governance events and processes</u> (Community Plenary, Steering Board, Advisory Council etc.).

TA7|M2: Technical Coordination Office (TCO)

Similar to the centralized-decentralized structure of the ACO the TCO has the responsibility to provide the overall technical management and coordination of the consortium. FIZ will have the TCO lead and each of the co-applicants will also have expert technical staff that together as TCO team monitor and support all technical developments across the different task areas on the consortium level and the general infrastructural approaches on the NFDI level. The TCO must ensure the overall technical sustainability and future viability of implementations and will play a major role in certification processes for software and services of the consortium. The TCO will also coordinate the working groups on technical cross-cutting topics and provide the consortium with the technical infrastructure for digital communication, collaborative documentation, software development and service monitoring.

Tasks in TA7|M2

TA7|M2|T1 <u>Coordination of technical and infrastructural components</u> between the task areas.

TA7|M2|T2 <u>Coordination of cross-area technical documentation</u> for all infrastructure components and services of the consortium.

TA7|M2|T3 <u>Coordination of and participation in working groups</u> on technical cross-cutting topics.

TA7|M2|T4 <u>Networking on technical and infrastructural components</u> in the NFDI and beyond.

TA7|M3: Governance operations

This measure bundles all events and operations for the governance of NFDI4Culture as described in the chapter "Organizational Structure and Viability". It contains the costs for the yearly Culture Community Plenary and the Advisory Council, the regular Culture Steering Board and Culture Spokesperson Committee meetings and actions like consultation (legally, financially) on adequate long-term operating models for the consortium in accordance with the developments on the NFDI level. In this respect, an

important task in this measure for the governance bodies of NFDI4Culture will be to identify services from the consortium that need specific operating models for their sustainability and to introduce pilot models for operation that will be evaluated with a feedback process and adapted to ensure permanent operation for the users of these services.

Tasks in TA7|M3

TA7|M3|T1 Governance events and processes (Community Plenary, Steering Board etc.).

TA7|M3|T2 Demand-oriented identification services from NFDI4Culture to be sustained.

TA7|M3|T3 Analysis of existing operating models (public/private partnership, consortium models, etc.) in accordance with developments on the NFDI level.

TA7|M3|T4 Pilot introduction of adequate operating models on identified services to be sustained.

TA7|M3|T5 Evaluation and feedback processes for introduced pilot operating model.

TA7 / M4: Coordination, cooperation, knowledge pooling for cross-cutting topics

The aim of this measure is to bring together and coordinate activities in the area of cross-cutting topics throughout the consortium. As laid down in the *Berlin Declaration*, cooperation with other consortia will be organized via working groups dedicated to subject-specific cross-cutting topics such as, in the case of NFDI4Culture, the standardization of imaging techniques, the treatment of complex digital object representations, or generic topics such as the establishment of an AAI. In this context several concrete working groups have already been agreed with other consortia.

Tasks in TA7|M4

TA7|M4|T1 Cross-area working groups (with other NFDI consortia on cross-cutting topics; subject-specific and transdisciplinary).

TA7|M4|T2 Knowledge pooling between task areas and cooperating consortia.

TA7|M4|T3 Identification, documentation and operationalization of new cross-cutting topics.

TA7|M5: Culture RDM Kickstarter

The Culture RDM Kickstarter is a governance measure to promote innovation and development of the NFDI4Culture service portfolio. Through a consolidated evaluation process, members and participants of the consortium can make proposals for the implementation of subprojects that strengthen the consortium's research data management and expand the range of services offered. All proposals must be based on a precise analysis of the needs of the community. The "checks and balances" process on which M4 is based means that a coordinated admission of new participants can also take place in later years of the consortium. The Kickstarter promotes the research data management of

NFDI4Culture in two directions: internally by financing case studies (integration of new services into the consortium) and externally by financially supporting the initiation of third-party funded projects between members and participants with external sponsors (e.g. DFG, BMBF etc.). The measure is already in great demand and has received strong support during the evaluation of the work program by the community at the Community Workshop in September 2019.

Tasks in TA7|M5

TA7|M5|T1 <u>Acquisition of Kickstarter projects and case studies</u> from the NFDI4Culture community.

TA7|M5|T2 <u>Definition and operationalization</u> of an admission procedure for new participants to the consortium.

TA7|M5|T3 <u>Evaluation and generalization of results</u> from finished Kickstarter projects case studies for the consortium and the NFDI as a whole.

TA7|M6: Dissemination, outreach and community enlargement

This measure bundles all strategies and tasks of the consortium in the field of dissemination of work results and continuous community expansion with new users. This includes the presentation of the consortium and its services at national and international conferences (during subject-specific conferences of the participating disciplines and during conferences in the field of (inter-)national research data management). For the regular transfer of the results of the consortium's work and its services into the community of interest, a close strategic cooperation with the academic societies participating in the consortium is pursued. In annual workshops, the academic societies will bring NFDI4Culture closer to their respective communities (musicology, art history, theatre and dance studies, architecture, media studies, digital humanities). The dissemination strategy of NFDI4Culture also includes the production of marketing materials and the facilitation of Open Access publications on specific topics developed in the task areas (e.g. rights or data ethics in the field of tangible and intangible cultural assets).

Tasks in TA7|M6

TA7|M6|T1 <u>Presence of ACO/TCO (lead) and selected delegates of NFDI4Culture on national and international conferences</u>, workshops etc.

TA7|M6|T2 <u>Dissemination workshops on NFDI4Culture</u> and its research data management by the participating academic societies in their respective research communities

TA7|M6|T3 <u>Production of NFDI4Culture marketing materials</u> and OA publications

Key performance indicators

Aim 1: Efficient governance operations, administration and technical coordination

- **KPI 1.1**: Efficient administration of the consortium. Criterion: ACO has been set up at all co-applicant institutions and all necessary contracts and administration processes are established after 3 months. ACO has established regular conference calls and stand ups. Yearly financial audits and regular compliance checks have been established. All governance operations (CCP, CSB, CSC) take place without delays, efficient reporting is in place.
- **KPI 1.2**: Overarching technical convergence and sustainability. Criterion: TCO has been set up at all co-applicant institutions and regular conference calls have been established after 3 months. TCO has established cross-area technical documentation for the infrastructure and services of the consortium after 1,5 year. TCO has built a knowledge exchange network with other NFDI consortia (through working groups) for overarching infrastructural topics on the NFDI level (cf. Berlin declaration)
- **KPI 1.3**: Continuous optimization of NFDI4Culture's governance towards its community of interest and the NFDI as a whole. Criterion: All governance bodies and processes have been evaluated and optimized in year 2 (2022) and year 4 (2024).
- **KPI 1.4**: Development of sustainable operating models in accordance with the NFDI as a whole. Criterion: Completed analysis of existing operating models (public/private partnership, consortium models, etc.) after 3 years, piloting of adequate long-term operating models for selected services after 4 years, performance evaluation of introduced operating models for selected services after 5 years.

Aim 2: Knowledge pooling and NFDI-wide cooperation on cross-cutting topics

- **KPI 2.1**: Close cooperation, knowledge exchange and joint action with other NFDI consortia on defined cross-cutting topics. Criterion: Concrete working groups and collaboration formats are in place after 1 year. Cooperation is evaluated each year and new working groups are set up or joined depending on developments on the NFDI level.

Aim 3: Integration, innovation and openness for new contributions to the consortium

- **KPI 3.1**: NFDI4Culture and its work program is open to respond to newly arising demands from the users and has defined processes for the integration of new participants to the consortium. Criterion: At least 3 research projects (external funding) or integration case studies between members and (new) participants of the consortium have been kickstarted and supported by the consortium each year.

Aim 4: Increased acceptance and outreach for the consortiums' services

- **KPI 4.1**: NFDI4Culture's services find increased acceptance in the community of interest and also become broadly known beyond the consortium's core research domains. Criterion: Yearly dissemination workshops and/or (international) conferences have been carried out by the participating academic societies. Delegates of NFDI4Culture present the consortium and its services on yearly national and international conferences.

Abbreviations

Table 9

Table 9.
List of abbreviations.

AAF	Advanced Authoring Format
AAI	Authentication and authorisation infrastructure
AAT	Art & Architecture Thesaurus
ACE	Architects Council of Europe
ACO	Administrative Coordination Office
ADHO	Alliance of Digital Humanities Organisations
Affecav	Association française des enseignants et chercheurs en cinéma et audiovisuel
AGATE	European Science Academies Gateway for the Humanities and Social Sciences
AIM	Associação de Investigadores da Imagem em Movimento
ALLEA	European Federation of Academies of Sciences and Humanities
API	Application Programming Interface
AR/VR	Augmented Reality / Virtual Reality
BAFTSS	British Association of Film, Television and Screen Studies
BMBF	Bundesministerium für Bildung und Forschung
CAAD	Computer-aided architectural design
CAC	Culture Advisory Council
CAD	computer-aided design
CCP	Culture Community Plenary
CIDOC	Comité International pour la Documentation
CIHA	Comité International d'Histoire de l'Art
CI/CD	Continous Integration / Continous Delivery
CLARIN	Common Language Resources and Technology Infrastructure
CRM	Conceptual Reference Model
CRDA	Culture Research Data Academy

CSB	Culture Steering Board
CSC	Culture Spokesperson Committee
DARIAH	Digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften
DOI	Digital Object Identifier
EADH	European Association for Digital Humanities
EASTAP	European Association of the Study of Theatre and Performance
EOSC	European Open Science Cloud
Exif	Exchangable Image File Format
FID	Fachinformationsdienst
FOKUS	Forschungsdatenkurse für Studierende und Graduierte
FRBR	Functional Requirements for Bibliographic Records
FTE	Full Time Equivalent
GLAM	Galleries, Libraries, Archives, Museums
GND	Gemeinsame Normdatei
ICOM	International Council of Museums
IFLA LRM	International Federation of Library Associations and Institutions Library Reference Model
IFTR/FIRT	International Federation for Theatre Research
IGR	Immaterialgüterrechte in verteilten Informationsinfrastrukturen
IIIF	International Image Interoperability Framework
IMS	International Musicological Society
I-SAW	ICONCLASS seeAlso Widget
L3S	Research center, Hannover
LIDO	Lightweight Information Describing Objects
LoC	Letter of commitment (by NFDI4Culture participants)
LOD	Linked Open Data
LoS	Letter of support (by NFDI4Culture supporters)
LTA	Long Term Archiving
M	Measure
MEI	Music Encoding Initiative
MeCCSA	Media, Communication and Cultural Studies Association (GB)
METS	Metadata Encoding & Transmission Standard
MODS	Metadata Object Description Schema
NCCO	NFDI4Culture Coordination Office
OMR	Optical Music Recognition
OpenAIRE	Open Access Infrastructure for Research in Europe
ORCID	Open Researcher Contributor Identification Initiative
PID	Persistent Identifier

PREMIS	Preservation Metadata: Implementation Strategies
RADAR	Research Data Repository, FIZ Karlsruhe
RDA	Resource Data Alliance
SCMS	Society for Cinema and Media Studies (USA)
SIBMAS	International Society of Libraries, Archives and Documentation Centres of the Performing Arts
TA	Task Area
TCO	Technical Coordination Office
TEI	Text Encoding Initiative
TGN	Getty Thesaurus of Geographic Names
TOPORAZ	Topographie in Raum und Zeit
TRUST	Training zum Umgang mit sensiblen Forschungsdaten
ULAN	Union List of Artist Names
W3C	World Wide Web Consortium
XMP	Extensible Metadata Platform

Funding program

Hosting institution

Academy of Sciences and Literature | Mainz

References

- Alge B (2019) Forschungsdatenmanagement in der Musikethnologie. Universitätsverlag Hildesheim, Hildesheim, 110 pp. [In German]. [ISBN 978-3-487-15835-8] https://doi.org/10.18442/031
- Aliverti C, Fabian C, Kailus A (2015) RDA und Kultureinrichtungen. Zeitschrift für Bibliothekswesen und Bibliographie 62 (6): 329-338. https://doi.org/10.3196/186429501562643
- Arendt T, Kranz S, Mantz F, Regnat N, Taentzer G (2011) Towards syntactical model quality assurance in industrial software development: process definition and tool support. In: Reussner R, Grund M, Oberweis A, Tichy W (Eds) Reussner, R., Grund, M., Oberweis, A. & Tichy, W. (Hrsg.), Software Engineering 2011 – Fachtagung des GI-Fachbereichs Softwaretechnik. Bonn: Gesellschaft für Informatik e.V.. (S. 63-74). [In English]. URL: https://dl.gi.de/handle/20.500.12116/19847 [ISBN 978-3-88579-277-2].
- Arendt T, Taentzer G (2013) A tool environment for quality assurance based on the Eclipse Modeling Framework. Automated Software Engineering 20 (2): 141-184. https://doi.org/10.1007/s10515-012-0114-7

- Association of Research Libraries / Wikimedia Foundation (2019) ARL White Paper on Wikidata. In: Puente M (Ed.) Opportunities and Recommendations. URL: https://www.arl.org/publications-resources/4751-arl-white-paper-on-wikidata-opportunities-and-recommendations
- Bicher K (2018) FRBR-Ebenen und Normdaten zu Werken der Musik für Musikwissenschaft und Musikbibliotheken. In: Fischer B (Ed.) Sessionspace bei der GNDCon 2018, 4.12.2018. Frankfurt a. Main. URL: https://wiki.dnb.de/display/GNDCON2018/Dokumentation+der+Sessions+der+GNDCon+2018
- Bicher K, Wiermann B (2018) Normdaten zu „Werken der Musik" und ihr Potenzial für die digitaleMusikwissenschaft. Bibliothek Forschung und Praxis 42 (2): 222-235. https://doi.org/10.1515/bfp-2018-0043
- Bove J, Schmahl K (2015) Fotografische Nachlässe. Sammlungsund Aktivierungsstrategien am Beispiel des Archivs der Fotografen in der Deutschen Fothek. Zeithistorische Forschungen, Heft 2/2015, Göttingen 334-336. [In German].
- Bove J (2018) Fotografie aktivieren. Kooperative Strategien des „Archivs der Fotografen" in der Deutschen Fotothek. Kooperative Informationsinfrastrukturen als Chance und Herausforderung 384-397. https://doi.org/10.1515/9783110587524-040
- Bracht C (2016) Warum eine sinnvolle Dokumentation von Kunstwerken ohne die Bibliotheken nicht mehr länger denkbar ist. Neue Strategien am Bildarchiv Foto Marburg, Beitrag zur VDA-Fortbildung für Fachreferent*innen der Kunstwissenschaft. Heidelberg 2016. URL: https://www.vdb-online.org/veranstaltung/718/bracht_foto_marburg.pdf
- Brett A, Croucher M, et al. (2017) Research Software Engineers: State of the Nation Report 2017. Zenodo. https://doi.org/10.5281/zenodo.495360
- Bundesministerium für Bildung und Forschung (2019) Data Literacy – Mehr Datenkompetenz für Studierende. URL: https://www.bildung-forschung.digital/de/data-literacy---mehr- datenkompetenz-fuer-studierende-2355.html
- Busch H, Rettinghaus K, Schrade T, Schulte S (2018) Aktuelle Situation der RSEs: Karriere – welche Karriere? Verband Digital Humanities im deutschsprachigen Raum. Blog der AG DH-RSE. URL: https://dh- rse.github.io/workshop/dhd2018/karriere/anerkennung/arbeitsbedingungen/2018/08/20/aktue lle-situation-der-rses-karriere-welche-karriere.html
- Coburn E, Light R, McKenna G, Stein R, Vitzthum A (2010) LIDO – Lightweight Information Describing Objects. Version 1.0.. ICOM URL: http://www.lido-schema.org/schema/v1.0/lido-v1.0-specification.pdf
- Crofts N, M. Doerr M, Gill T, Stead S, Stiff M (Eds) (2011) Definition of the CIDOC Conceptual Reference Model. Version 5.0.4. URL: http://www.cidoc- crm.org/sites/default/files2/cidoc_crm_version_5.0.4.pdf
- Czmiel A, Druskat S, Schrade T (2018) Research Software Engineering und Digital Humanities. Reflexion, Kartierung, Organisation. Kritik der digitalen Vernunft, Köln Konferenzabstracts der DHd 2018: 56-56. URL: https://dh-rse.github.io/dhd-workshop-2018-presentation
- Deutscher Museumsbund (Ed.) (2018) Leitfaden zum Umgang mit Sammlungsgut aus kolonialen Kontexten. Berlin. URL: https://www.museumsbund.de/wp- content/uploads/2018/05/dmb-leitfaden-kolonialismus.pdf

- Div. Authors (n. d.) (2020) Wikidata: WikiProject Cultural heritage.
  https://www.wikidata.org/wiki/Wikidata:WikiProject_Cultural_heritage. Accessed on:
  2020-7-16.
- EC EGFD - European Commission Expert Group on FAIR Data (2018) Turning FAIR
  into reality. Final Report and Action Plan from the European Commission Expert Group
  on FAIR Data,. Brüssel. https://doi.org/10.2777/54599
- Effinger M (2018) Wissen verbreiten – im Open Access publizieren. Infrastrukturen für
  die Digitale Kunstgeschichte. In: Kuroczynski P, Bell P, Dieckmann L (Eds) Computing
  Art Reader. Einführung in die digitale Kunstgeschichte. Heidelberg. Heidelberg, 269-285
  pp. https://doi.org/10.11588/arthistoricum.413
- Effinger M, Krabbes F, Withanage D (2018) Crossmediales Publizieren bei Heidelberg
  University Publishing (heiUP). Bit-Online 21 (5): 393-403. URL: https://www.b-i-t-
  online.de/heft/2018-05-fachbeitrag-effinger.pdf
- Effinger M, Maylein L, Šimek J (2019) Von der elektronischen Bibliothek zur innovativen
  Forschungsinfrastruktur. Bibliothek Forschung und Praxis 43 (2): 311-323.
  https://doi.org/10.1515/bfp-2019-2067
- Gesellschaft für Informatik e.V. (2019) Data Literacy – Kompetenzen in der Hochschule.
  https://gi.de/dataliteracy/. Accessed on: 2020-7-16.
- Glöckner F, et al. (Ed.) (2019) Berlin Declaration on NFDI Cross-Cutting Topics (Version
  1). Zenodo. https://doi.org/10.5281/zenodo.3457213
- GND-Kooperative, Kett J (2017) Initiative für Normdaten und Vernetzung:
  GNDEntwicklungsprogramm 2017–2021, Stand 06/2017. URL: https://wiki.dnb.de/
  download/attachments/125418325/GND_Entwicklungsprogramm17- 21_2017-06.pdf
- Hadjakos A, Iffland J, Keil R, Oberhoff A, Veit J (2017) Challenges for Annotation
  Concepts in Music. International Journal of Humanities and Arts Computing 11 (2):
  255-275. https://doi.org/10.3366/ijhac.2017.0195
- Hartmann T (2013) Zur urheberrechtlichen Schutzfähigkeit der Forschungsdaten. J.
  Taeger (ed.): Law as a Service (LaaS) – Recht im Internetund Cloud-Zeitalter. vol. 1,
  Edewecht 2013, pp. 505–515. URL: http://hdl.handle.net/11858/00-
  001M0000-0014-1208-E
- Hartmann T (2014) Urheberrecht in der Bildungspraxis. Leitfaden für Lehrende und
  Bildungseinrichtungen. Bielefeld.
- Hartmann T (2017) 1f. Open Access rechtlich absichern – warum es ein Opt-in braucht.
  Praxishandbuch Open Access 45-52. https://doi.org/10.1515/9783110494068-006
- Hartmann T (2018) Urheberrecht abgelaufen, trotzdem abgemahnt? Wikimedia kämpft
  vor Gericht für Gemeinfreiheit. Netzpolitik.org. URL: https://netzpolitik.org/2018/
  urheberrecht-abgelaufen-trotzdem-abgemahnt-wikimedia- kaempft-vor-gericht-fuer-
  gemeinfreiheit/
- Heidrich J, Bauer P, Krupka D (2018a) Future Skills: Ansätze zur Vermittlung von Data
  Literacy in der Hochschulbildung. Hochschulforum Digitalisierung beim Stifterverband
  für die Deutsche Wissenschaft e.V. Berlin. Arbeitspapier Nr. 37, ed. by Geschäftsstelle.
  URL: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/
  HFD_AP_Nr37_DALI_Stud ie.pdf
- Heidrich J, Bauer P, Krupka D (2018b) Strukturen und Kollaborationsformen zur
  Vermittlung von Data-Literacy-Kompetenzen. Hochschulforum Digitalisierung beim
  Stifterverband für die Deutsche Wissenschaft e.V. Berlin. Arbeitspapier 32, ed. by

Geschäftsstelle. URL: https://gi.de/fileadmin/GI/Hauptseite/Aktuelles/Meldungen/2018/HFD_AP_Nr32_Data_Literac y_Kompetenzen_Literatur.pdf

- Kailus A (2017) Iconclass als Baustein des Semantic Web? Eine Positionsbestimmung, 23. Berliner Veranstaltung der internationalen EVA-Serie Electronic Media and Visual Arts (= EVA Berlin 23), Heidelberg 2017 (2016), pp. 47–53. DOI: https://doi.org/10.11588/arthistoricum.256.338 . In: Bienert A, et al. (Ed.) Konferenzband EVA Berlin 2016. 23. Berliner Veranstaltung der internationalen EVA-Serie Electronic Media and Visual Arts (= EVA Berlin 23). Elektronische Medien & Kunst, Kultur und Historie, Heidelberg, 47–53 pp. URL: https://doi.org/10.11588/arthistoricum.256.338

- Kailus A (2018) Spartenübergreifende Öffnung und Weiterentwicklung der Gemeinsamen Normdatei: Das Projekt GND für Kulturdaten (GND4C). In: Bienert A, et al. (Ed.) Konferenzband EVA Berlin 2018. Elektronische Medien & Kunst, Kultur und Historie: 25. Berliner Veranstaltung der internationalen EVA-Serie Electronic Media and Visual Arts (= EVA Berlin 25), Heidelberg, 241–245 pp. https://doi.org/10.11588/arthistoricum.442

- Kailus A, Stein R (2018) Besser vernetzt: Über den Mehrwert von Standards und Normdaten zur Bilderschließung. Informationszugang, semantische Interoperabilität, Linked Open Data, Normdaten, Koreferenzierung. In: Kuroczynski P, Bell P, Dieckmann L (Eds) Computing Art Reader, 2018. Einführung in die digitale Kunstgeschichte, Heidelberg 2018, 119-139 pp. https://doi.org/10.11588/arthistoricum.413

- Klimpel P (2015) Eigentum an Metadaten? Urheberrechtliche Aspekte von Bestandsinformationen und ihre Freigabe. Handbuch Kulturportale 57-65. https://doi.org/10.1515/9783110405774-006

- Klimpel P, König E- (2015) Urheberrechtliche Aspekte beim Umgang mit audiovisuellen Materialien in Forschung und Lehre. Gutachten für die Gesellschaft für Medienwissenschaft und den Verband der Historiker und Historikerinnen Deutschlands. Berlin.

- Klimpel P, Rack F, Weitzmann JH (2017) Neue rechtliche Rahmenbedingungen für Digitalisierungsprojekte von Gedächtnisinstitutionen. Handreichung. 4th Completely Rev. Edition. Berlin. https://doi.org/10.12752/2.0.002.3

- Knaus G, Stein R, Kailus A (2019) LIDO-Handbuch für die Erfassung und Publikation von Metadaten zu kulturellen Objekten. Vol. 1. Graphik, Heidelberg. https://doi.org/10.11588/arthistoricum.382.544

- Manovich L (2011) Cultural Software. 2011. URL: http://manovich.net/content/04-projects/070- cultural-software/67-article-2011.pdf

- Martin RC (2018) Clean Architecture: A Craftsman's Guide to Software Structure and Design, Boston 2018. Rat für Informationsinfrastrukturen (2018): Zusammenarbeit als Chance. Zweiter Diskussionsimpuls zur Ausgestaltung einer Nationalen Forschungsdateninfrastruktur (NFDI) für die Wissenschaft in Deutschland. Göttingen. URL: www.rfii.de/?wpdmdl=2529

- Münzmay A (2018) Lesen und Schreiben im digitalen Dickicht. Bibliothek Forschung und Praxis 42 (2): 236-246. https://doi.org/10.1515/bfp-2018-0031

- NFDI (2019) **Memorandum of Understanding** by NFDI initiatives from the humanities and cultural studies. Zenodo. https://doi.org/10.5281/zenodo.3265763

- NFDI4Culture Working Paper (2019) Fokusthemen und Aufgabenbereiche für eine Forschungsdateninfrastruktur zu materiellen und immateriellen Kulturgütern. Living

Document der NFDI-Initiative NFDI4Culture. Zenodo. (Version 1.0). https://doi.org/10.5281/zenodo.2763576

- Niggemann E, Goebel R, Kretzschmar R, Bracht C, Ahlers J (2017) GND Für Kulturdaten (GND4C). Beschreibung des Vorhabens – Projektanträge im Bereich „Wissenschaftliche Literaturversorgungsund Informationssysteme" (LIS). URL: https://wiki.dnb.de/download/attachments/134055796/B_eRT_GND4C_public.pdf?version=1 &modificationDate=1518198808000&api=v2

- Patton GE (2010) Funktionale Anforderungen an Normdaten: Ein konzeptionelles Modell. De Gruyter Saur, Berlin. [In German]. https://doi.org/10.1515/9783110232516

- Poulter M (2017) Wikidata: the new hub for cultural heritage. https://blog.wikimedia.org.uk/2017/01/wikidata-the-new-hub-for-cultural-heritage/. Accessed on: 2020-7-16.

- Ridsdale C, et al. (2015) Strategies and best practices for data literacy education. Knowledge synthesis report. URL: http://hdl.handle.net/10222/64578

- Röwenstrunk D (2018) Langzeitverfügbarkeit von wissenschaftlicher Software im Bereichhistorisch-kritischer Musikedition. Bibliothek Forschung und Praxis 42 (2): 302-308. https://doi.org/10.1515/bfp-2018-0029

- Rula A, Zaveri A (2014) Methodology for Assessment of Linked Data Quality. 1st. Workshop on Linked Data Quality. URL: http://ceur-ws.org/Vol-1215/paper-04.pdf

- Schelbert G (2017) ... warum nicht gleich Wikidata?! M. Stolz (ed.): Konferenzabstracts DHd 2017 Bern. Digitale Nachhaltigkeit (13.–18. Februar 2017), pp. 287–288. URL: https://wikis.hu-berlin.de/mediathek/...Warum_nicht_ gleich_Wikidata%3F

- Schrade T (2017) Nachhaltige Softwareentwicklung in den Digital Humanities. Konzepte und Methoden. In: Stolz M (Ed.) Konferenzabstracts der DHd2017. Bern, 168–171 pp. URL: http://www.dhd2017.ch/wp-content/uploads/2017/02/Abstractband_ergaenzt.pdf

- Simon H (2007) „prometheus" und Justitia – Bildarchive der Kunst- und Kulturwissenschaften im Spannungsfeld des medialen Umbruchs hin zu einer digitalen Informationsgesellschaft. Forschung und Lehre im Informationszeitalter - zwischen Zugangsfreiheit und Privatisierungsanreiz 65-86. https://doi.org/10.1515/9783110893038.65

- Stein R, Balandi O (2019) Using LIDO for Evolving Object Documentation into CIDOC CRM. Heritage 2 (1): 1023-1031. https://doi.org/10.3390/heritage2010066

- Universitätsbibliothek Mainz (Ed.) (2019) Handout: Rechtliche Fragen bei der Bereitstellung von Forschungsdaten. Mainz. URL: https://www.digitale-bibliotheksdienste.uni- mainz.de/files/2019/02/Handreichung-Forschungsdatenbereitstellung.pdf

- van der Graaf M, Waaijers L (2014) Authority files: Breaking out of the library silo to become signposts for research information. URL: http://repository.jisc.ac.uk/6224/1/Authority_files_-_Breaking_out_of_the_library_silo.pdf

- Veit J, Richts K (2018) Stand und Perspektiven der Nutzung von MEI in der Musikwissenschaft und inBibliotheken. Bibliothek Forschung und Praxis 42 (2): 292-301. https://doi.org/10.1515/bfp-2018-0026

- Wiermann B (2018) Bibliothekarische Normdaten und digitale Musikwissenschaft. Die Musikforschung 71/4. pp. 338–357.

- Wilkinson M, Sansone S, Schultes E, Doorn P, Bonino da Silva Santos LO, Dumontier M (2018) A design framework and exemplar metrics for FAIRness. Scientific Data 5 (1). https://doi.org/10.1038/sdata.2018.118

Endnotes

[*1] TA2 will collaborate with other consortia such as MaRDI, Text+ and NFDI4Neuroscience in the area of modelling and standardisation of complex data types. NFDI4Ing will cooperate with NFDI4Culture in the area of standardisation and curation of 3D data types (e.g., CAAD models and other forms of 3D representations).

[*2] TA5 is dedicated to work on cross-area technical and legal topics (such as a aoint authentication and authorisation infrastructure AAI and a collaborative terminology service) that are relevant to many other consortia. The measures of this task area will be continuously expanded to address cross-cutting topics in the NFDI which were still unknown at the time of application (cf. the Berlin Declaration).

[*3] NFDI4Ing and NFDI4Culture have identified Data Literacy, Code Literacy, and the provision of open educational resources as cross-cutting topics for close collaboration in TA6.

[*4] TA7|M4 aims to coordinate activities in the area of cross-cutting topics throughout the consortium. As laid down in the Berlin Declaration, cooperation with other consortia will be organised via working groups dedicated to subject-specific cross-cutting topics.