UNION
DER DEUTSCHEN AKADEMIEN
DER WISSENSCHAFTEN

AGATE: Concept for a European Academies Internet Gateway for the Humanities and Social Sciences

Ulrike Wuttke, Carolin Ott, Dominik Adrian and Simon Worthington

SPONSORED BY THE

Federal Ministry
of Education
and Research

AGATE: Concept for a European Academies Internet Gateway for the Humanities and Social Sciences

Final Report of the Project “Aufbau eines europäischen Akademienportals”

Ulrike Wuttke, Carolin Ott, Dominik Adrian and Simon Worthington

Union of the German Academies of Sciences and Humanities

June 2017

SPONSORED BY THE

Federal Ministry of Education and Research

AGATE: Concept for a European Academies Internet Gateway for the Humanities and Social Sciences

Final Report of the Project “Aufbau eines europäischen Akademienportals”

Ulrike Wuttke, Carolin Ott, Dominik Adrian and Simon Worthington

June 2017

Union of the German Academies of Sciences and Humanities

URL: https://doi.org/10.5281/zenodo.806179

Contact: agate@akademienunion-berlin.de

Citation: Ulrike Wuttke, Carolin Ott, Dominik Adrian and Simon Worthington, ‘AGATE: Concept for a European Academies Internet Gateway for the Humanities and Social Sciences‘, Final Report of the Project “Aufbau eines europäischen Akademienportals”, Union of the German Academies of Sciences and Humanities (2017).

This work is licensed to the public under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.
https://creativecommons.org/licenses/by/4.0/legalcode

Union der deutschen Akademien der Wissenschaften
Geschwister-Scholl-Straße 2
55131 Mainz
Germany
http://www.akademienunion.de/

The views expressed in this publication are exclusively the opinions of the authors and not those of the German Federal Ministry of Education and Research.

Funded by the German Federal Ministry of Education and Research, reference number 01UG1503.

Thanks and appreciation are due to the German Federal Ministry of Education and Research (BMBF) for funding the project and to the German Aerospace Center (DLR) for their cooperation as project management agency.

Contents

AGATE Scientific Advisory Board

Dr. Elena González-Blanco García (Director of the Digital Humanities Lab (LINHD) at the National Distance Education University (UNED))

Dr. Natalie Harrower (Director of the Digital Repository of Ireland, Royal Irish Academy)

Prof. Dr. Gerhard Lauer (Professor of German, University of Göttingen, Member of the Göttingen Academy of Sciences and Humanities)

Prof. Dr. István Monok DSs (Director General of the Library and Information Centre of the Hungarian Academy of Sciences (MTAK))

Prof. Dr. Claudine Moulin (Professor of German, Trier University, Director of the Trier Center for Digital Humanities)

Prof. Dr. Laurent Romary (INRIA (French National Institute for computer science and applied mathematics), DARIAH-EU, Berlin-Brandenburg Academy of Sciences and Humanities (BBAW))

List of Tables and Figures

Tables

Figures

List of Abbreviations and Acronyms

acatech	Deutsche Akademie der Technikwissenschaften (German Academy of Science and Engineering)
ACDH	Austrian Center for Digital Humanities
AGATE	European Academies Internet Gateway
AIP	Leibniz Institute for Astrophysics Potsdam
Akademienunion	Union der deutschen Akademien der Wissenschaften (Union of the German Academies of Sciences and Humanities)
ALLEA	All European Academies
API	Application Programming Interface
ARIADNE	Archaeological research data infrastructures
BAdW	Bayerische Akademie der Wissenschaften (Bavarian Academy of Sciences and Humanities)
BARTOC	Basel Register of Thesauri, Ontologies & Classifications
BBAW	Berlin-Brandenburgische Akademie der Wissenschaften (Berlin-Brandenburg Academy of Sciences and Humanities)
BMBF	Bundesministerium für Bildung und Forschung (German Federal Ministry of Education and Research)
CC	Creative Commons
CCeH	Cologne Center for eHumanities
CERIF	Common European Research Information Format
CERN	Conseil européen pour la recherche nucléaire
CESSDA	Consortium of European Social Science Data Archives
CLARIAH	Common Lab Research Infrastructure for the Arts and Humanities (NL)
CLARIN	Common Language Resources and Technology Infrastructure
DARIAH	Digital Research Infrastructure for the Arts and Humanities
DASISH	Digital Services Infrastructure for Social Sciences and Humanities
DCC	Digital Curation Center
DESIR	DARIAH-ERIC Sustainability Refined
DFG	Deutsche Forschungsgemeinschaft (German Research Council)
DH	Digital Humanities
DiMPO	Digital Methods and Practices Observatory Working Group
DINI	Deutsche Initiative für Netzwerkinformation e. V. (German Initiative for Network Information
DiRT Directory	Digital Research Tools Directory
DMP	Data Management Plan
DNB	Deutsche Nationalbibliothek (German National Library)
DOI	Digital Object Identifier
DRP	Digital Research Practices
DRT	Digital Research Tools
EASSH	European Alliance for Social Sciences and Humanities
EOSC	European Open Science Cloud
ERA	European Research Area
ERIC	European Research Infrastructure Consortium

ESF	European Science Foundation
ESFRI	European Strategy Forum on Infrastructures
ESR	Early-Stage Researchers
ESS	European Social Survey
euroCRIS	European current research information systems
EUROHORCS	European Heads of Research Councils
FAIR Principles	Guiding principles for findable, accessible, interoperable and re-usable data
FHP	Fachhochschule Potsdam (University of Applied Sciences)
FOSTER	Facilitate Open Science Training for European Research
FP	Framework Programme (European Commission)
GESIS	Leibniz-Institut für Sozialwissenschaften (Leibniz Institute for the Social Sciences)
GIS	Geographic Information Systems
GLAM	Galleries, Libraries, Archives, Museums
H2020	Horizon 2020 (8th FP of the European Commission)
HERA	Humanities in the European Research Area
IANUS	Forschungsdatenzentrum Archäologie & Altertumswissenschaften (Research Data Centre for Archaeology and Ancient Studies)
ICT	Information and Communication Technology
ISSP	International Social Survey Programme
KNAW	Koninklijke Nederlandse Akademie van Wetenschappen (Royal Netherlands Academy of Arts and Sciences)
LIS	Library and Information Systems
MPG	Max-Planck-Gesellschaft (Max Planck Society)
MQ	SASSH Main Questionnaire
MTA	Magyar Tudományos Akadémia (Hungarian Academy of Sciences)
NWO	Netherlands Organisation for Scientific Research
OA	Open Access
ÖAW	Österreichische Akademie der Wissenschaften (Austrian Academy of Sciences)
OKF	Open Knowledge Foundation
OpenAIRE	Open Access Infrastructure for Research in Europe
PAN	Polska Akademia Nauk (Polish Academy of Sciences)
PARTHENOS	Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies
PI	Principal Investigator
PID	Permanent Identifier
RI	Research Infrastructure
SAGW	Schweizerische Akademie der Geistes- und Sozialwissenschaften (Swiss Academy of Humanities and Social Sciences)
SASSH	Survey and Analysis of Basic Research in the Social Sciences and Humanities at the Science Academies and Related Research Organisations of Europe
SHARE	Survey of Health, Ageing and Retirement in Europe

SQ	SASSH Short Questionnaire
SSH	Social Sciences and Humanities
STEM	Science, Technology, Engineering and Math
TaDiRAH	Taxonomy of Digital Research Activities in the Humanities
TeLeMaCo	Teaching and Learning Materials Collection
UAI	Union Académique Internationale
UC3	University of California Curation Center
W3C	World Wide Web Consortium
WP	Work Package

Executive Summary/Kurzfassung

Executive Summary

The European academies of sciences and humanities' research significantly contributes to the study and preservation of Europe's cultural heritage. Still, the great potentials of an increasing digitalization to better access and exploit this wealth of knowledge are not fully being seized. The project of the Union of the German Academies of Sciences and Humanities funded by the German Federal Ministry of Education and Research (BMBF) from October 2015 until June 2017 aimed to conceptualize the framework of a pan-European digital infrastructure that improves the visibility and findability of the academies' SSH research and promotes international cooperation. Therefore, the basic requirements of a future European Academies Internet Gateway (AGATE) for the Social Sciences and Humanities were mapped and possibilities examined for integration, reuse, and scaling of existing services and solutions.

Two stakeholder workshops, numerous bilateral talks with representatives of European science academies and research infrastructure experts, as well as interviews with future users provided the necessary background information and feedback to conceptualize digital access to academies' research. Furthermore, a potential core consortium of partners was formed, that is a working group comprising of academies and research infrastructures that declared a dedicated interest in actively contributing to a future implementation of AGATE. The project results led to a conceptual exposé for the projected portal that serves as a well-founded starting point for the actual implementation and is comprised of recommendations for technical and organizational aspects, user and outreach concepts, as well as for sustainability and the planning of estimated resources.

Based on the project results, the conceptual exposé recommends a modular structure, building up on a central web portal for user entry. As an integral part of the portal, the recommendation envisages the development of a database that comprises detailed information about the research projects at European Academies and points to available digital resources, the "AGATE Knowledge Map". At the same time, the web portal would provide guidelines that enable the users to enter or rather integrate the project information and digital resources into the "Knowledge Map" and offer focused information and resources on topics of special interest for the academies' researchers such as Open Access and Open Data.

In addition to online offers like FAQs and webinars it is recommended to develop, preferably in cooperation with strategic partners, individual workshops and trainings that meet the specific user needs, and could be promoted via the portal and channels of the partners. The enhanced integration of social media could strengthen and

facilitate knowledge exchange and cooperation in the field of sustainable digital research and publication practices among the researchers of the European academies and beyond, as well as with pan-European infrastructures.

To guarantee the sustainability of the AGATE project, several strategies were pursued. First, the project substantially participated for the Union of the German Academies in the Horizon2020 proposal “Designing an Open Innovation Research Infrastructure demonstrated on the example of European Scientific Academies (AGATE)”, that was submitted under the leadership of the Austrian Academy of Sciences on 29 March 2017. The H2020 project aims at the design of an Open Innovation Research Infrastructure with a focus on the academies’ research. Second, the project provided important input for the national project database, that was resolved by the Executive Committee of the Union in March 2017 and will be implemented under the auspices of the Academy of Sciences and Literature, Mainz.

Kurzfassung

Die europäischen Wissenschaftsakademien tragen mit ihren Forschungsvorhaben maßgeblich zur Erforschung und Bewahrung des kulturellen Erbes Europas bei. Die großen Potenziale, welche die zunehmende Digitalisierung für die Verfügbarkeit und Erschließung dieses Wissensschatzes birgt, werden jedoch noch nicht voll ausgeschöpft. Das vom Bundesministerium für Bildung und Forschung (BMBF) von Oktober 2015 bis Juni 2017 geförderte Projekt der Union der deutschen Akademien der Wissenschaften „Aufbau eines europäischen Akademienportals“ verfolgte das Ziel, den Rahmen für eine paneuropäische digitale Infrastruktur zu sondieren, welche die Sichtbarkeit und Findbarkeit der geistes- und sozialwissenschaftlichen Forschung an den Akademien verbessert und internationale Kooperationen befördert. Dazu wurden die Grundanforderungen für ein zukünftiges europäisches Akademienportal für die Geistes- und Sozialwissenschaften (European Academies Internet Gateway, kurz: AGATE) ermittelt und Möglichkeiten geprüft, bestehende Angebote und Lösungskonzepte zu integrieren, weiterzuverwenden und anzupassen.

Zwei Workshops, bilaterale Gespräche mit Vertreterinnen und Vertretern europäischer Wissenschaftsakademien und Infrastrukturexpertinnen und -experten sowie stichprobenartige Nutzerinterviews lieferten die erforderlichen Vorgaben und Rückmeldungen bei der Konzeptionierung eines digitalen Zugangs zur Akademienforschung. Außerdem wurde ein potentielles Kernkonsortium gebildet, das heißt eine Arbeitsgruppe aus Partnerakademien und Forschungsinfrastrukturen, die daran interessiert sind, bei einer späteren Umsetzung von AGATE eine aktive Rolle zu übernehmen. Die Ergebnisse des Projekts flossen in ein ausführliches Konzeptexposé

für das geplante Portal ein, das einen fundierten Ausgangspunkt für eine spätere Umsetzung bietet und Empfehlungen für technische und organisatorische Aspekte, Nutzer- und Outreachkonzepte sowie zur Nachhaltigkeit und zur Ressourcenplanung enthält.

In dem Konzeptexposé wird eine modulare Struktur empfohlen, die auf einem zentralen Webportal als Nutzereinstieg aufsetzt. Als zentrales Modul soll eine Datenbank aufgebaut werden, die detaillierte Informationen über die Forschungsprojekte an den europäischen Wissenschaftsakademien enthält und verfügbare digitale Ressourcen aufzeigt („AGATE Knowledge Map“). Zugleich sollen auf dem Webportal Richtlinien bereitgehalten werden, welche die Eingabe bzw. Einbindung der Projektinformationen und digitalen Ressourcen in die „Knowledge Map“ erleichtern. Darüber hinaus sollen Informationsangebote zu Themen wie Open Access und Open Data für die Wissenschaftlerinnen und Wissenschaftler der Akademien zur Verfügung gestellt werden. Neben reinen Online-Angeboten wie FAQs und Webinaren sollten zusätzlich möglichst in Kooperation mit strategischen Partnern individuelle Workshops und Trainings entwickelt werden, die spezifischen Nutzerbedürfnissen begegnen, und diese über das Portal sowie Kanäle der strategischen Partner beworben werden. Durch die verstärkte Einbindung sozialer Medien und digitaler Kommunikationskanäle könnte der Wissensaustausch und die Zusammenarbeit im Bereich nachhaltiger digitaler Forschungs- und Publikationspraktiken unter den Wissenschaftlerinnen und Wissenschaftlern der europäischen Akademien und jenseits davon sowie mit paneuropäischen Infrastrukturen vorangetrieben werden.

Um die Nachhaltigkeit des AGATE-Projekts sicherzustellen, wurden mehrere Strategien verfolgt: Zum einen hat sich das Projekt für die Akademienunion maßgeblich am Horizon 2020-Antrag „Designing an Open Innovation Research Infrastructure demonstrated on the example of European Scientific Academies (AGATE)“ beteiligt, der unter Leitung der Österreichischen Akademie der Wissenschaften am 29. März 2017 eingereicht wurde. Das H2020-Projekt hat das Design einer europäischen Open Innovation Forschungsinfrastruktur mit einem Schwerpunkt auf der Akademienforschung zum Ziel. Zugleich lieferte das Unionsprojekt wichtige Impulse für die vom Präsidium der Akademienunion im März 2017 beschlossene nationale Projektedatenbank, die unter der Federführung der Mainzer Akademie der Wissenschaften und der Literatur implementiert wird.

1 Background and Research Programme

1.1 Background

One of the key endeavours of the European science academies is to document, research, and preserve the rich and diverse European (and global) cultural heritage. By doing this they contribute significantly to the European research landscape and the formation of an European consciousness and identity. This has been revealed by the SASSH Survey (Survey and Analysis of Basic Social Science and Humanities Research at the Science Academies and Related Research Organisations of Europe) that was conducted from 2013 to 2015 by the Union of the German Academies of Sciences and Humanities (Akademienunion) and the European Federation of Academies of Sciences and Humanities (ALLEA) which provided the first inventory of the SSH (Social Sciences and Humanities)[1] research activities of the European science academies, a specific SSH research sector whose peculiarities had not been studied before.[2] The survey gathered data of over six hundred SSH projects run at or by science academies[3] and resulted in a report titled "Survey and Analysis of Basic Social Science and Humanities Research at the Science Academies and Related Research Organisations of Europe".[4]

The diversity of the forms, functions, and research priorities of the European science academies mirrors the diversity of Europe itself. While some of them primarily act as learned societies, others are national centres of research, in the sciences as well as in the social sciences and humanities.[5] The majority are rooted in a centuries old tradition of an academic elite, which is one of the reasons why they generally have a significant role for their respective national research communities and their respective areas of research.

---

[1] Social Sciences and Humanities (SSH) – the usual phrasing in many contexts and common English abbreviation, for example within the Horizon 2020 framework, though for example the DFG Classification of subject areas turns the word order around (Humanities and Social Sciences). See 'DFG | Subject Areas Structure', accessed 30.04.2017. The authors keep to SSH in the document without wishing to indicate any preference of one of the research areas over another.

[2] Leathem and Adrian 2015, p. 12: "Beyond the potential awareness that there are many of them Europe-wide, that they are traditional and elite with high scholarly renown, there has been no integrative record of the work and working methods of the academies until now".

[3] As can be derived from the name, the SASSH Survey collected data from European science academies and related research organisations. It was out of scope of the current project to recalculate the results to arrive at results only for the academies. Given the fact that 85% of the projects that responded to the SASSH Survey are conducted at or associated with European science academies, it seems justifiable to consider the SASSH results as representative, respectively at least giving a good indication. Additional desk research and bilateral exchanges with academies' representatives and researchers during the current project manifested the SASSH results in general.

[4] The SASSH Survey resulted in the publication Leathem and Adrian 2015.

[5] Around fifty-seven academies in more than forty countries from the Council of Europe region are brought together by ALLEA, the European Federation of Academies of Sciences and Humanities, see 'ALLEA | Homepage', accessed 30.04.2017.

The European science academies form a separate non-university sector because of their distinctive organisational structure. Science academies are self-governed communities of scientists and scholars. Generally speaking, two types of science academies can be distinguished today: 1) the research academy type (academies that are involved in research activities, which often take place in more or less independent research institutes or research centers); and 2) the "learned society academy type" (academies without any direct involvement in research activities).

Historically, science academies are learned societies with elected members, i.e. distinguished individuals, who may be elected by other members.[6] Excellence is the main selection criterion for academy membership. Only the most excellent scholars of their fields are elected as (lifetime) members (sometimes the word academician is used to denote an elected member of an academy). The academies' members form a link between the academies and the universities, and other research performing organisations. Given their prominent role the members are destined to act as ambassadors for scientific communities. Nowadays, many academies are also important regional or national non-university research organisations, often organised as an academy with organisational and institutional functions, with more or less independent research units or institutes. Some national academies act as the most important non-university research organisations of their respective country (for example the Hungarian Academy of Sciences (MTA))[7] and also the scholars directly involved in the respective academies' research are important ambassadors to and members of wider scientific communities. Though the SASSH Survey and the scope of AGATE focus solely on the academies' SSH research, one needs to be aware of the fact that the European science academies have different research scopes. While some academies are solely active in the SSH,[8] others cover with their research activities a broad variety of scientific fields,[9] and some are even specialised in non-SSH research.[10] It is also important to note that in general the learned society level of a science academy (e.g. the level of the academy's members) is organisationally separated from its research activities. That means, being an academy member does not postulate any involvement in the home academy's research activities[11] and that being responsible for a research project carried out under the umbrella of an academy or working for such a project is not equivalent with being a member of an academy.

[6] See for example the procedure described for the Royal Society of Edinburgh, 'RSE | Becoming a Fellow', accessed 30.04.2017.
[7] See 'MTA | Homepage', accessed 30.04.2017.
[8] For example the Swiss Academy of Humanities and Social Sciences (SAGW), see 'SAGW | Homepage', accessed 30.04.2017, and most of the member academies of the Akademienunion.
[9] Many national academies, e.g. the Hungarian Academy of Sciences (MTA).
[10] For example acatech (German National Academy of Science and Engineering), see 'acatech | Homepage', accessed 30.04.2017.
[11] Even if an academy is only active in SSH research, often the academicians represent a broad variety of scientific fields (e.g. Mathematics and Natural Sciences, Humanities and Social Sciences). Often they are organised in classes and there is a difference between full and corresponding members.

Both academy types that have been discussed above serve as fora for scholarly exchange for their members and beyond. The academies are dedicated to the promotion of excellence in research in general, which is demonstrated by awarding prizes and speaking on behalf of the scientists as a united voice, often in an advisory function for policy makers and/or as national think tanks (for example the U.K. based Royal Society).[12] Therefore, they often play an important role in science diplomacy efforts, i.e. in efforts of "bridging the world" via science,[13] and a leading role in their respective national research landscapes by providing role models for example through statements on science relevant topics.[14]

Academies' SSH research as a community of practice

The SASSH Survey revealed many similarities between the SSH research carried out at the European academies that belong to the research academy type.[15] Medium-term to long-term projects (6-15 years) and long-term projects (over 15 years) in the humanities prevail, although especially in Eastern and Central Europe and in the social sciences also short-term projects (up to and including 5 years) occur.[16] The majority of the research projects can be classified as long-term basic SSH research[17] that is concerned with collections, dictionaries, and editions[18] mainly in the historical humanities.[19] Social science projects — mainly with a historical sociological focus — play a minor role, but again especially at some of the academies in Eastern and Central Europe contemporary social science projects proliferate next to the humanities.[20] For example to the Hungarian Academy of Sciences belongs the Research Centre for Social

[12] See 'Royal Society | Homepage', accessed 30.04.2017.
[13] For a recent campaign in this context see 'Royal Society | European Academies' Statement: Science Is Global', accessed 30.04.2017.
[14] See for example 'Leopoldina | Policy Advice', accessed 13.05.2017.
[15] Given the above explanation, generally speaking, no research activities are carried out at learned society type academies that fell under the scope of the SASSH Survey. But as with all rules there are exceptions, this is the case here. For example the British Academy (for the Social Sciences and Humanities) is a learned society, but at the same time the UK's national body for the humanities and social sciences, which includes not only fellowships and short term grants, but also the (partial) funding of long-term projects, the so called Academy Research Projects (which are then carried out at independent institutions), see project list 'British Academy | Academy Research Projects', accessed 10.05.2017.
[16] See Leathem and Adrian 2015, pp. 42-45.
[17] Leathem and Adrian 2015, p. 2: "Firstly, the academies are bodies of expertise in basic (as opposed to applied) research in the humanities, and particularly in the historical humanities. Research interest in history, language, religion, literature and geography (including anthropology and ethnography) is particularly widespread." See also Leathem and Adrian 2015, pp. 27-28.
[18] See Leathem and Adrian 2015, p. 2.
[19] See Leathem and Adrian 2015, pp. 29-41.
[20] See Leathem and Adrian 2015, p. 2.

Sciences[21] with its own Research Documentation Centre,[22] to the Polish Academy of Sciences the Institute of Philosophy and Sociology,[23] and to the Bulgarian Academy of Sciences the Institute for the Study of Societies and Knowledge[24] which is involved for example in SHARE (Survey of Health, Ageing and Retirement in Europe).[25] The SASSH Survey thus manifested that the European science academies are major generators and storehouses of knowledge and contribute with their ample long-term research projects fundamentally to European basic SSH research on cultural heritage and identity, not least because excellence is a characteristic of the academies' research activities.[26]

The similarities between the research projects carried out under the umbrella of the European science academies and their special organisational structure and position in the research landscape strongly suggests to consider the SSH researchers at the European academies as a discrete community of practice.[27] Because academies provide an outstanding climate for research activities that need tenacity, they proliferate especially in collaborative[28] long-term historical SSH basic research projects that need to be compiled and maintained over decades. Therefore, these projects share common research methods and practises with a special emphasis on research outputs, such as collections, dictionaries, and editions, that are based on extensive amounts of material. The output of these long-term basic research projects lays the scientific foundation for various further research activities and is often valid for a long time.

The wealth of similarities and overlaps between research fields, topics, and methods would lend itself ideally for research cooperation. Although the SASSH Survey revealed that researchers at the European academies are generally interested in increasing cooperation amongst each other on a European scale, these potentials are often not being exploited for various reasons, amongst them the lack of information about ongoing research, but also the lack of appropriate funding schemes for European long-term basic SSH research.[29] By stronger promoting and interlinking their SSH research with each other and beyond, the European academies would at the one hand greatly

[21] See 'MTA | Centre for Social Sciences', accessed 8.05.2017.
[22] See 'MTA | Research Documentation Centre', accessed 8.05.2017.
[23] See 'PAN | Institute of Philosophy and Sociology', accessed 8.05.2017.
[24] See 'BAS | Institute for the Study of Societies and Knowledge', accessed 8.05.2017.
[25] See 'SHARE | Homepage', accessed 08.05.2017.
[26] Irrespective of their concrete organisational form, regular evaluation of the projects to ensure their scientific excellence, often in the form of external evaluation, is the norm. See Leathem and Adrian 2015, pp. 123-128.
[27] Academies are for example considered as a specific stakeholder group in the stakeholder analysis of IANUS (Heinrich et al. 2015, p. 7), besides non-university research organisations, monument protection services, research cooperations, cultural resource management services (archaeological excavation companies), museums, and universities.
[28] The majority of the research projects are conducted in teams (with 3-6 members). See Leathem and Adrian 2015, pp. 55-57.
[29] See Leathem and Adrian 2015, p. 2, pp. 113-119.

benefit from the increased potential for collaboration, knowledge exchange, scientific innovation, and on the other hand they could contribute together to "combat the linguistic and scientific fragmentation that exists at the pan-European level".[30]

1.2 Status quo of digital SSH research methods and publication practices at the European science academies

Given the objective of the SASSH Survey to identify potentials for the improvement of the connectivity amongst the European academies SSH projects, the survey paid special attention to digital practices, such as digital tools, Open Access, archiving practices, the role of data standards, and institutional support and training resources,[31] because of the premise that "connecting geographically dispersed researchers and projects requires digital resources and tools accessible to and useable by all".[32]

The survey indicated for the authors particular room for improvement in the following three domains

1. general visibility;
2. findability, access, reusability, and sustainability; and
3. knowledge exchange about standards, good practices, and infrastructure partners.

The status quo is outlined below.

1. Low general visibility of SSH research of the academies

In general, the visibility of the research activities of the academies SSH projects' (and their digital resources) leaves great room for improvement. Many projects use mainly exclusive scientific or limited dissemination channels (e.g. scientific publications, conferences, press releases, email mailing lists) and do not fully employ the possibilities of the net and digital communications as dissemination channels. This applies to information about the project, for example on a dedicated project landing

[30] See ESF 2011, pp. 6-7, citation p. 7.
[31] See Leathem and Adrian 2015, p. 13.
[32] See Leathem and Adrian 2015, p. 13. In 2015 the DARIAH Digital Methods and Practices Observatory WG (DiMPO) conducted the first European survey on scholarly practices and digital needs in the arts and humanities. (National) highlight reports were published from the 2016 onwards (see 'DARIAH | Digital Methods and Practices Observatory Working Group (DiMPO) - Zenodo Community', accessed 02.05.2017). The final report is announced to be published in the middle of 2017. Because the DiMPO survey focusses explicitly on "digital humanists" and their individual research methods and practices and a much broader research community the results are only remotely comparable with the SASSH Survey.

page or on an institutional homepage as well as the publication of research data and results (also or solely) in digital form (and the acceptance of Open Access for electronic publishing varies geographically, see also the following point (2)).[33] Also multiple respondents of the SASSH Survey called for "improved public visibility of research data, sources, and findings"[34].

In the following, the term "digital resources" will be used to denote various forms of digital research data and research results.[35] The term digital resources can refer to statistical data, textual corpora, image databases, audio-visual material, and digital editions, but also for example to source code for digital software tools, and last but not least to more classical forms of electronic publishing, especially monographs (in PDF format) — which are (despite the growing role of other forms of digital research data and research output) especially in printed form still considered as the gold standard for scholarly research in the humanities, along with scientific articles.

The low visibility and lack of information on the SSH research projects undertaken at respective academies makes their research virtually invisible and thus unknown to non-specialists. If the research does not use identifiers, such as DOIs etc., and complete the metadata thoroughly then the work will not be visible in RIs. The equivalent is not giving a book an ISBN or completing its metadata and expecting the books to be found in libraries or distributed to bookstores. Despite the very important role of the European science academies as political adviser on national and European scale in general,[36] their significant role for basic SSH research in particular, and the relevance of their research for a wider non-specialist public, their SSH research activities, results, and special needs are barely present in the eyes of the wider public and in discussions of research policy at European level.[37]

---

[33] See Leathem and Adrian 2015, p. 3, pp. 107-109, pp. 111-112.

[34] Leathem and Adrian 2015, p. 80.

[35] Especially in the humanities the definition of research data is more complex than in those research fields that work with measured or observed data (and most humanities researchers would not think not of their digital sources or results as such) (see Sahle 2015). Humanities data are often highly complex and strongly connected with the technical surrounding that was used to produce, process, and publish them. This has far-reaching consequences for data curation (denoting "the active and ongoing management of datas throughout its entire lifecycle of interest and usefulness to scholarship" Cragin et al. 2007), which exceeds straight forward archiving of the data.

[36] Recently funded in this context was the H2020 project SAPEA (Science Advice to Policy by European Academies). Its consortium consists of our project partner ALLEA together with four other European Academy Networks. See 'ALLEA | SAPEA: Science Advice for Policy by European Academies', accessed 13.05.2017. ALLEA has just recently released the revised edition of the European Code of Conduct for Research Integrity (see ALLEA 2017), which is used by the European commission as a reference document in the H2020 Model Grant Agreement (see 'ALLEA | News entry 24.03.2017', accessed 30.04.2017).

[37] Arnold, Barker and Slipersæter 2010, p. 1: "Research institutes, variously defined, account for almost half of Europe's public expenditure on R&D, yet they are in many respects almost invisible. There are no systematic statistics about them. What they do is to a large extent undocumented. The institutes have been consistently ignored until very recently in ERA development and discussions, despite their key nodal role in the Framework Programmes. Very little reform has taken place in the institute sector,

2. Barriers to the utilisation of digital resources of SSH research at the academies (findability, access, reusability, sustainability)

Findability

The low general visibility of the SSH research project in general affects the findability of the digital resources. Especially digital resources beyond articles or monographs are difficult to find as they are still more or less out of the scope of classical library catalogs. In these cases more specialized search entrances have to be used that are often only known by a limited group of experts (such as subject specific databases, OpenAIRE[38] for research publications and data, DataCite[39] for research data, GitHub[40] for source code, GESIS[41] for social science data, re3data[42] for research data repositories), if the data are stored digitally in dedicated and interoperable (data) repositories at all and do not remain on internal institutional servers.[43] In the specialised search entrances mentioned above the academies' projects and their digital resources may "get lost in the shuffle", the search results only represent subsets of the overall output, and various forms of digital resources belonging to the same project may not be examined at the same time, which may lead to the undesirable result that relevant resources may be overlooked because they are located in different digital "silos" which are (not yet) well interlinked with each other or searchable by popular search engines such as Google. This problem manifests itself also at a more abstract level, that is concerning searches that exceed the scope of a project, but are aimed at identifying thematic, chronological, or methodological related materials.

Access

Identifying a digital resource does not equal having immediate access to it as many resources are not available in Open Access (OA), but locked behind paywalls.[44] Although the majority of the European science academies has signed and issued an

---

except for changes to bring former Soviet-style academies into line with EU practice. Unlike the universities, the institutes are barely present in discussions of research policy, especially at the European level."

[38] See 'OpenAIRE | Homepage', accessed 13.05.2017.

[39] See 'DataCite | Homepage', accessed 13.05.2017.

[40] See 'GitHub | Homepage', accessed 13.05.2017.

[41] See 'GESIS | Data Archiving', accessed 13.05.2017.

[42] See 're3data.org | Homepage' , accessed 13.05.2017.

[43] Leathem and Adrian 2015, p. 96: "At present, archiving is chiefly an internal process within the institution (e.g. the institution's server). The use of data repositories or the archives of large-scale online library initiatives like GESIS remain an exception."

[44] It is clear that not all digital resources can be OA. Especially in the social sciences data are often sensitive and require limited access. In the case of GESIS they are findable via a metadata catalogue, but subject to various access level controls and sometimes with costs attached. See 'GESIS | Access', accessed 13.05.2017.

Open Science statement[45] and an Open Access statement[46] under the umbrella of ALLEA and several academies have individually signed the important Budapest and/or Berlin declarations on Open Access[47] and/or also have their own institutional Open Access policies,[48] Open Access seems to be "an ideal professed but not practiced".[49] The results of the SASSH Survey indicate that only half of the academies' projects that publish their results in electronic form (about 60% of the total), do so as Open Access, which amounts to around just one third of the responses.[50] The low percentage is particularly noteworthy as the researchers themselves have to bear the consequences and name limited access as a major hindrance in their research activities.[51] The reasons for the low acceptance or implementation of Open Access were out of the scope of the survey, but it can be assumed that they are manifold. First of all, copyright restrictions are not specific to the academies' research outputs or SSH alone[52] as must be noted that the humanities are in general lagging behind because of a different publication culture.[53] However, it can be assumed from the project's staff discussions with researchers that a common and fundamental obstacle is the long tradition of the academies' projects, which leads to the fact that still many of the academies respective projects have long-standing contracts with publishing houses that include copyright restrictions (non OA).[54] Furthermore by the SASSH Survey the lack of awareness and knowledge on the side of the researchers[55] — the latter not

[45] See ALLEA 2012.
[46] See ALLEA 2013 and ALLEA 2015.
[47] The Budapest Open Access Initiative first defined OA in 2002 (see 'Budapest Open Access Initiative | Homepage', accessed 13.05.2017). For the actual list of signatures see 'Budapest Open Access Initiative | View Signatures', accessed 13.05.2017). Another milestone of the Open Access movement was the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities in 2003 (see 'MPG | Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities', accessed 13.05.2017). For the actual list of signatures see 'MPG | Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities: Signatories', accessed 13.05.2017).
[48] See for example the open access policies of the Royal Society ('Royal Society | Open Access Publishing', accessed 13.05.2017) or the SAGW ('SAGW | Open Access Strategie der SAGW', accessed 13.05.2017).
[49] Quote taken from the title of Andreoli-Versbach and Mueller-Langer 2013.
[50] See Leathem and Adrian 2015, p. 98.
[51] See Leathem and Adrian 2015, p. 68.
[52] Harrower (ed.) 2015, p. 12: "Shared data is still not standard in scholarly nor in scientific research practices."
[53] See 'Science.ORF.at | Open Access: Warum die Geisteswissenschaftler zögern (Interview with Walter Scheidel), 14.08.2013', accessed 13.05.2017.
[54] Harrower (ed.) 2015, pp. 12-13: "Access to cultural heritage materials is restricted by copyright laws that recall the traditions of the 19th century thinking. Projects run by Academies long before the digital age began are bound to contracts which do not allow data sharing."
[55] Leathem and Adrian 2015, p. 91: "A major issue in fostering online publication, particularly concerning Open Access, is a lack of clarity regarding copyright. This is reflected in responses by ICT and library staff to the question in the short questionnaire of the most frequently asked questions by researchers concerning digital research practices: 21% state that these concern copyright issues. Although online publications are protected by copyright laws, these vary according to country, and are often shrouded by ambiguity. In the face of this, copyright on electronic publications is often disregarded. The

surprising given the complex legal situation concerning copyright, licences, and IPR on national and European levels — and the lack of positive institutional reinforcement (including clear guidelines and support)[56] are recurrently identified as such barriers to OA at the European academies.

A holistic approach to data management planning (DMP), based on the FAIR principles,[57] making use of discipline specific data management guidelines and DMP-tools is required to unleash data's full potential (see also below, point 3):

> **"**Publishing/depositing in OA is a major step, but doing it properly is equally important. Your best bet is to use repositories/journals that use persistent IDs (e.g., CrossRef, DataCite, CNRI handles, ORCID, FundRef) and they expose data in an interoperable form."[58]

Reusability

Academies produce a lot of highly relevant qualitative SSH data,[59] especially so called small data, but also big data.[60] Both could be of use to a vast research community if progress was made in the field of Open Science — especially in the areas of Open Access and (Linked) Open Data, and Open Source — as well in the promotion of common standards and sharing of best practices to improve interoperability. Still many of the academies' digital resources are not accessible in a form that allows reuse for all research needs (e.g. text and data mining).

Although it is difficult to estimate on the basis of the SASSH Survey the overall amount of electronic publications which are only in the form of monographs or articles in PDF format[61] — a format which is quite unsuitable for reuse even if accessible in OA[62] —

---

erroneous assumption that everything freely available online can also be freely edited and re-used is wide-spread."

[56] Leathem and Adrian 2015, p. 91: "34% of respondents to the short questionnaire state that their institution has a policy on electronic publishing in place (including open access/open data), 42% state that it does not, and 24% are unsure. Upon the request for further details on these policies, 40% of responses explicitly refer to international standards, e.g. the Berlin Declaration."

[57] See 'FORCE11 | The FAIR Data Principles', accessed 13.05.2017.

[58] 'OpenAIRE | For Researchers', accessed 13.05.2017.

[59] Harrower (ed.) 2015, p. 13: "The output of [computational] methods relies crucially on the quality of the data input, and here projects run by the Academies are relevant. The resources that are created in Academy projects lend themselves particularly well to this kind of computer-based study as they are painstakingly precise in their coding of data. [...] Given the leading role that Academies have played in providing such [correct, authoritative and well-structured, UW] cultural data in digital formats, they should also take the opportunity to lead the way in the use of these resources, not only as advanced books, but by focusing on the genuine new possibilities that they open up."

[60] Academy of Finland 2015, p. 4: "In recent years, there has been growing recognition in the field of digital humanities of the need not only for "big data", but also "small" (deep, rich) data, underlining the importance of expertise and knowledge of content that facilitates insightful analysis of data in each specific field. The complementarity of "big" and "small" data opens up many new avenues for research."

[61] See Leathem and Adrian 2015, pp. 92-93.

the survey results point (besides the low implementation of Open Science) to general obstacles in the area of interoperability (common shared metadata schemes and standards) that cause restrictions to the reuse of the academies´ digital resources.[63]

Sustainability

Especially the aspects "access" and "reuse" of digital resources are strongly connected with the sustainability of the digital resources, because if the digital resources are not preserved, this naturally means an end to all access and reuse. This problem begins with the well-known phenomenon of a broken link if a website is no longer maintained and no more of use to verify information.

The common denominator "digital sustainability" (also long-term preservation or digital curation) describes a span of activities that more or less encompass the whole research (data) lifecycle and exceeds the narrow sense of archiving in general linguistic usage. The term archiving means to archives, museums, and libraries more than permanent storage on a medium, it encompasses the notion of ensuring long-term access and therefore includes the need to preserve modes of reuse and retaining the interpretability of the digital resources.[64] This is a collective task, that includes many stakeholders, from researchers to digital preservation specialists.

The European academies are without doubt big storehouses of data, but their digital content is endangered. Especially the task to maintain complex, enhanced publications, one of the academies' main digital SSH output exceeds many academies' individual resources and the landscape of trustable preservation partners in this field is still under way.[65] Data, once vanished into oblivion when the web application cannot be maintained anymore are hard to revive.[66] Not only individual solutions for the actual research process with little regard to compatible formats and standards prevail,[67] but also long-term preservation solutions that exceed (predominantly closed) archiving beyond an institutional server are not yet common. This concerns the digital sustainability of all digital resources, research results and research data alike, as the latter are often stored (only temporarily) using free storage space from commercial services such as Google that are neither a long-term preservation solution nor adequate for sensitive data.[68] Although the SASSH Survey did not differentiate between research data and research results and did not go into detail concerning

[62] The PDF format is also critical concerning long-term preservation. Only the PDF/A format (ISO 19005: PDF/A) fulfills digital preservation requirements, for details see for example 'IANUS | IT-Empfehlungen: PDF-Dokumente', accessed 13.05.2017.
[63] See Leathem and Adrian 2015, pp. 75-101.
[64] See Neuroth et al. (nestor Handbuch) 2010, Kap. (title) 1:3.
[65] See Buddenbohm, Engelhardt and Wuttke 2016.
[66] See for example the LAZARUS project ('CCeH | LAZARUS', accessed 13.05.2017).
[67] See Leathem and Adrian 2015, pp. 3-4.
[68] See Leathem and Adrian 2015, p. 77.

current archiving practices (where in which form digital resources are preserved), it nevertheless gives some indications. Printed forms still play an enormous role as 'archival form of research output', but in the majority of cases also the electronic equivalents are being archived, but apparently often on non-interoperable/accessible, institutional servers: "At present, archiving is chiefly an internal process within the institution (e.g. the institution's server). The use of data repositories or the archives of large-scale online library initiatives like GESIS remain an exception."[69]

As already mentioned above, digital sustainability encompasses the whole research (data) lifecycle, requires an encompassing approach to DMP and thus includes the sustainability of digital research tools (DRT). Also in this area, the SASSH Survey revealed room for improvement and concrete actions, as concerns about the sustainability of the DRT were formulated explicitly on the researcher's side: "A further wish across the spectrum of responses pertains to the continuity or sustainability of DRT. This is a particularly important consideration for long-term research projects, which, without sustainable DRT, may find themselves having to change the tools and methods on which the project is based along the way."[70] The rather bleak picture regarding the digital sustainability of the academies' SSH research may be partially explained by institutional inadequacies as well as with lacking resources, but the SASSH Survey also indicates that either fitting sustainable digital tools and services do not exist or are unknown to the researchers. One striking example from the survey is the reliance on Google services for many research areas, especially free storage which involves risks because of Google's privacy policy and business policies.[71]

3. Lack of resources, awareness, information and guidelines

This point is strongly connected with the previous point, because the issues mentioned above can only be tackled in the course of long-term processes, for which the prerequisites often yet have to be created.

[69] Leathem and Adrian 2015, p. 95.
[70] Leathem and Adrian 2015, p. 78.
[71] Leathem and Adrian 2015, p. 77: "And Google is not the only IT enterprise that procures access to user data via cost-free programmes which is then sold on to third parties. These practices offer no form of data protection and are particularly inadmissible for the preservation of sensitive, personal sociological data. It is with good reason that elaborate processes are in place to anonymize data before it is used or made public."

Lack of resources

The SASSH Survey revealed an uneven development at the European academies concerning digital research methods and publication practices.[72] Developing and maintaining digital humanities and digitisation projects (especially in the sense of sustainable research methods and publication practices) requires specialized support which cannot be offered by all academies sufficiently on their own. While some academies have specialised information services departments, or cooperate with external information service departments (such as partner universities, libraries, or computing centers) or European SSH-infrastructures, some academies lack sufficient in-house infrastructure support or access to national or international infrastructures.[73]

Amongst the European infrastructures mentioned in the SASSH Survey as either known or actively used by academies' researchers, and/or as active cooperation partner of an academy (academy as data provider or partner institution) were: Europeana, DARIAH, CLARIN, ESS, CESSDA, SHARE, CENDARI, DASISH, and ISSP.[74] Although the SASSH Survey thus revealed a sound general awareness of the most prominent European digital SSH-initiatives amongst the academies researchers - with Europeana taking the lead before CLARIN, DARIAH, or CESSDA, it seems to be necessary to raise awareness for the fact that these infrastructures have more to offer than search tools and can be considered as competent research partners who provide research tools, and research supportive services (including long-term preservation) and that the European academies can actively participate in shaping them either as content contributors or strategic partners.[75]

Lack of awareness, information and guidelines

Fostering an ideal environment for Open Science and ensuring digital sustainability are collaborative tasks with many stakeholders involved. Only bundled efforts, especially

[72] Further information on the role of national academies for the Digital Humanities in Europe may be derived from the essays published as part of the Editorial: The Status Quo of Digital Humanities in Europe 2014 (see 'H-Soz-Kult | Editorial: The Status Quo of Digital Humanities in Europe', accessed 13.04.2017).

[73] The uneven development may not only be explained by different research traditions in specific disciplines or by general regional infrastructural weaknesses, but also by the fact that - with the exception of the German Academies' programme - most academies' projects are not supported by stable and reliable funds, as was also revealed by the SASSH Survey. See Leathem and Adrian, p. 2.

[74] See 'Europeana | Homepage', accessed 13.05.2017, 'DARIAH | Homepage', accessed 13.05.2017, 'CLARIN | Homepage', accessed 13.05.2017, 'ESS | Homepage', accessed 13.05.2017, 'CESSDA | Homepage', accessed 13.05.2017, 'SHARE | Homepage', accessed 13.05.2017, 'CENDARI | Homepage', accessed 13.05.2017, 'DASISH | Homepage', accessed 13.05.2017, 'ISSP | Homepage', accessed 13.05.2017. COST, which was also mentioned as an digital infrastructure is excluded here, because it is obviously not an infrastructure, but a funding programme. More digital infrastructures, especially in the SSH can be found via European Strategy Forum on Research Infrastructures (ESFRI) 2016 and Wuttke, Buddenbohm and Engelhardt 2014.

[75] See Leathem and Adrian 2015, pp. 87-90.

in the areas of awareness raising, guidance and support, may win over the scepticism and insecurities on the researchers' side that emerged from the SASSH Survey.

Although some academies already offer internal institutional trainings and informative events or support the participation in external events to improve knowledge about digital research tools or topics such as copyright and Open Access,[76] the SASSH Survey revealed insecurities based on missing information or clear guidelines, and a general scepticism concerning the sustainability of digital resources and tools.[77]

The need for stringent data management based on guidelines and using DMP as a tool is for example a rather new topic, especially in the humanities and social sciences. This is compounded by the fact that many national and international initiatives for the development of guidelines and good practices are still in its infancy. The SASSH Survey revealed that only 35% of the academies' ICT and digital library staff can clearly state that their institution has a data management policy and/or institutional guideline on data standards (35% no, 30% unsure). From the further details provided on these policies in the follow-up question it is difficult to estimate in how far these policies and guidelines are in accordance with international practices and standards. The varied answers indicate at least a need for harmonisation.[78] The lack of standards, sustainable digital tools and insecurities concerning long-term preservation and legal issues are not solely a problem of the academies' SSH-research projects. But due to the long-term perspective of their research activities and the long relevance of their digital resources; common standards, sustainable data management practices, data curation solutions, the continuity and sustainability of tools and digital infrastructures are especially urgent.[79]

The fact that 21% of the respondents of the SASSH Short Questionnaire (mainly ICT and library staff) mentioned copyright issues concerns on the side of the researchers as a major issue in fostering Open Access, which is not surprising given the complex legal landscape on the European scale, more information, clear guidelines and positive encouragement may make a huge difference. Under this area fall concrete measures, such as information about positive effects of Open Access in general, as well as more

[76] See Leathem and Adrian 2015, p. 87: "The short questionnaire asks library and ICT staff whether their institutions offer further training or informative events for digital research tools: 42% of respondents answer in the affirmative, and for 58% this is not the case."
[77] Leathem and Adrian 2015, p. 80: "Responses reflect the concern that the preservation of data is turning away from traditional archiving methods established over centuries and towards comparatively young methods of digital preservation that have not yet proven their permanence and stability. It would seem that the efforts of libraries in particular to secure sustainable and enduring archives for digital data are as yet unfamiliar to most researchers, who do not mention any such initiatives."
[78] Leathem and Adrian 2015, pp. 97-98.
[79] Leathem and Adrian, p. 78: "A further wish across the spectrum of responses pertains to the continuity or sustainability of DRT. This is a particularly important consideration for long-term research projects which, without sustainable DRT, may find themselves having to change the tools and methods on which the project is based along the way. Sustainability also concerns electronic archives, whose long-term sustainability has yet to stand the test."

practical advice concerning copyright issues and the possibilities of Creative Common licences in the form of trainings or tutorials for the researchers.[80]

In general, the SASSH Survey as well as exchanges during the present project phase point to a high potential for training and knowledge exchange activities aimed at (if not exclusively) academies' SSH researchers and governing bodies to raise more awareness for innovative digital research and publication practices and methods that could be employed either into running projects (digital transformation) or right from the start concerning new projects.

1.3 A digital infrastructure for the European academies

Based on the results of the SASSH Survey it was recommended to lend greater visibility to the European academies' SSH research activities by providing a central comprehensive scope of information about the academies' research activities, especially the ongoing research projects. This would create networking and cooperation possibilities between the European academies SSH projects and the broader scientific community and enhance the general impact of their research activities beyond their specific field. It was recommended to incorporate the project information in an accompanying digital infrastructure that "could and should lead to the pooling of corresponding digital resources"[81] (sources, data and publications) and that would promote "common, compatible systems for accessing, collecting, generating, sharing, analysing, storing and disseminating data and information" and "digital interoperability".[82]

Research infrastructures are well established in the humanities and social sciences. Research infrastructures such as archives, libraries, academies, museums and galleries are central to many strands of this kind of research by identifying, ordering, preserving, and making accessible of sources.[83] The European academies have a venerable tradition as research infrastructures with a strong focus on long-term basic SSH research and a key endeavour to document, research, and preserve the rich European cultural heritage, but unlike similar research bodies they lack a common digital infrastructure. While it is necessary for the SSH to speak with one voice to strengthen their position in the competition about the allocation of funding,[84] it doesn't mean that the existing SSH-infrastructures already meet the needs of all

[80] See Leathem and Adrian 2015, pp. 91-92, see also pp. 96-98.
[81] Leathem and Adrian 2015, p. 3.
[82] Leathem and Adrian 2015, p. 13.
[83] See ESF 2011, p. 2.
[84] In 2015 the European Alliance for Social Sciences and Humanities (EASSH) was founded to promote research on social sciences and humanities as a resource for Europe and the World by bringing together scientific networks, associations, disciplines and universities. See 'EASSH | Homepage', accessed 13.05.2017.

communities.[85] At the moment, the academies' digital resources are spread over various storage locations (and only to a small degree are Open Access) and not accessible via a centralised search entrance (and therefore only visible as subsets) and the information relevant to the community of researchers involved in academies' SSH activities is spread over various infrastructures, and there are no common networking tools.

In the SASSH Survey five priorities that need to be addressed by the proposed digital infrastructure were formulated

1. access to resources;
2. alignment of research data management policies;
3. alignment of Open Access policies;
4. information about user specific aspects of digital research and publication practices; and
5. cooperation with SSH infrastructures.[86]

Therefore, the development of the proposed infrastructure in the current project is based on the following basic tools, services, and requirements: a project database, a specialised search engine over the academies digital SSH resources, information about existing tools, standards, and relevant initiatives to increase their integration into the academies' projects research practices and the researchers' everyday research work, and the creation of state-of-the-art storage locations for academies' research data and research results where needed.

The elaboration of the basic framework into the present conceptual exposé for the proposed digital academies' infrastructure for SSH-research was guided by two overarching goals:

- Firstly, that the future infrastructure should bring measurable improvement in the problematic domains identified in the SASSH Survey: the low general visibility of SSH research of the academies, existing barriers to the utilisation and sustainability of digital resources of SSH research at the academies, and the lack of resources, awareness, information, and guidelines.

[85] Doorn 2014, p. 1: "Why are one or two research infrastructures not enough for the SSH? The silliness of this question becomes apparent as soon as we turn it around: why not have one research infrastructure for the natural sciences? You cannot look at the stars with a nuclear icebreaker and you cannot break ice with a telescope. The SSH are just as heterogeneous as the natural and life sciences, and therefore one tool or virtual lab does not fit all demands."

[86] See Leathem and Adrian 2015, p. 144. These areas correlate with the priority tasks for academies in the Digital Humanities identified by Immenhauser as 1) Creating of data; 2) Preserving of data; 3) Ensuring access to data; and 4) Creating conditions for the reuse of data and encompass the whole research data lifecycle. See Immenhauser 2015, p. 38. For a typical research data lifecycle in the SSH, see 'Data Archive | Research Data Lifecycle', accessed 13.05.2017.

- Secondly, in order to offer a functional fully-fledged research infrastructure that is of a supranational and supra-institutional relevance.[87] It will need to provide access to resources, services and tools that will foster innovation, collaboration, and knowledge exchange. Enabling a broad range of scientific users to conduct excellent research, as well as supporting cultural citizenship.

Potentials of the proposed Infrastructure (AGATE)

The proposed common digital infrastructure would on the one hand strengthen the European academies SSH research and on the other hand have a Europe wide impact. This applies especially in the fields of improving (open) access to, and the dissemination of digital and digitized cultural items, as well as scientific resources to research communities — to a wide range of users (educators, museums, exhibition curators, and the public), and stimulating internal and external interdisciplinarity.[88] Due to the potential European scale of the proposed infrastructure and the geographical scope of the core consortium (see Appendix 1 and 2) it will also act as a research enabler and capacity building incubator for areas of Europe that are less represented in current ESFRI landmark projects[89] and has a huge potential for collaboration and partnership beyond Europe.

Expected impact:

- Enhanced research efficiency: central access to information about academy SSH research in Europe (project information and digital resources).
- Facilitation of international and interdisciplinary cooperation, development of new joint research projects and new methods especially pertaining to European cultural heritage and identity.

[87] See European Strategy Forum on Research Infrastructures (ESFRI) 2016, p. 181: "As soon as they satisfy the coverage (with respect to geography as well as languages) requirements, multilingual and transcultural RIs such as the ESFRI landmarks CLARIN and DARIAH will bring European citizens into a position to become familiar with their common cultural heritage as well as with its richness and diversity. This is key to future oriented and sustainable development of our societies. In fact, "Cultural citizenship" is a key dimension for building and strengthening European citizenship and identity; studying, preserving and making available cultural items through the most advanced technologies is a highly relevant economic asset for European economy. Open access to historical material and heritage as well as the critical means to assess it can be considered as crucial to the development of any inclusive and reflective society". See also BMBF 2013, p. 2: "Forschungsinfrastrukturen im Sinne dieser Roadmap sind umfangreiche Instrumente, Ressourcen oder Serviceeinrichtungen für die Forschung in allen Wissenschaftsgebieten, die sich eine mindest nationale Bedeutung für das jeweilige Wissenschaftsgebiet auszeichnen sowie über eine lange Lebensdauer (in der Regel über 10 Jahre)."
[88] Named as science drivers and key priorities for SSH infrastructures in ESFRI 2016, pp. 171-175.
[89] See European Strategy Forum on Research Infrastructures (ESFRI) 2016, pp. 182-183.

- Strengthening of cross-links with disciplinary and/or national communities of practice,[90] especially by reaching academy researchers and members — who have a special potential to act as multiplicators or ambassadors to subcommunities — centrally via the AGATE portal.
- Stimulation of research through broader dissemination of a linguistic, regional, and disciplinary wide and varied range of scientifically excellent research data and research results to various scientific communities across Europe.
- Enabling of cultural citizenship through engagement of the wider public with these datasets and results (especially via new forms of scientific communication and by opening up the academies' datasets to the wider public in order to experiment and/or create alternative/unforeseen types of research).[91]
- Promotion of Open Science (especially Open Access and Linked Open Data) by providing a strong incentive to European academies — who often are amongst the respective national leading research institutions to use state-of-the-art technologies and open licences to make their digital resources widely accessible and fit for the European Open Science Cloud (EOSC).[92]
- Provide extra services (support) and success stories of early adopters to convince sceptical data providers and offer help to those who lack resources such as (programming of APIs, guidelines, translation, etc.).
- Act as research enabler and capacity builder for the move to digital research methods in academies' humanities and social science research through the brokering/provision of online information material and means of communications as well as (on-line) trainings to researchers from specific fields or regions and the pooling of the academies' resources to tackle common tasks and challenges.
- Stimulation of engagement, cooperation, and knowledge exchange with pan-European infrastructures[93] for example concerning issues of standardisation[94] and sustainability.
- Identification of needs for new research community specific services.
    - The proposed academies' infrastructure forms an excellent departing point for the development and implementation of tools and services for specific user

[90] This interdisciplinary approach can be compared to DARIAH, as described in Anderson, Blanke and Dunn 2010, p. 3790: "DARIAH, however, includes the claim that one can build an infrastructure based on cross-disciplinary scholarly activities, not just within discipline boundaries."
[91] See ESF 2011, p. 9.
[92] On the EOSC see 'European Commission | European Open Science Cloud', accessed 13.05.2017.
[93] The low visibility caused by the concentration on exclusive dissemination channels and a lack of engagement in actual discussions, especially with the pan-European infrastructures, has a negative impact on the contribution the European academies could "make to debates regarding sustained digital infrastructures and project-funded artefacts, long-term durable digital preservation, and societal responsibility for the preservation of our cultural heritage" See Harrower (ed.) 2015, p. 6.
[94] The role and importance of standards and European infrastructure's roles in promoting and supporting them as an added value was also underlined during the infrastructure conference organized by the SASSH project, see esp. Romary 2014.

groups/fields of research which are currently not addressed by existing infrastructures with national and international partners.[95]

- The development of ideas and areas for these additional new services and tools will be an important field of action for the core consortium during the further development of the basic concept provided in this document together with prospective users,[96] access to these new services and tools could be offered via the proposed common infrastructure.

[95] ESF 2011, p. 20: "Every researcher in the Humanities in Europe must be assured of finding a service provider for their digital research activities, for short-term accessibility as well as long-term preservation of data and publications. A key dimension to the delivery of this objective is the concept of a Europe-wide Research Infrastructure for the Humanities with strong (virtual) centres, as developed in DARIAH and CLARIN, and designed to support the development, promotion and implementation of shared protocols and standards."

[96] A proposal for AGATE within the framework H2020-INFRADEV_01_2017 (Horizon 2020) has been submitted which aims at designing an Open Innovation Research Infrastructure applying Open Innovation in Science methods and principles (see Appendix 2). This approach is supposed to detect needs and co-design innovative tools and features.

1.4 AGATE: Academies Projects and Publications Knowledge Map & Hub

Methodology

In the SASSH Survey it was specifically recommended to establish a "digital infrastructure [for and by the European Academies] that enables shared access to as many resources as possible and ensures standardised digital working practices"[97] to better seize the potentials of academies´ SSH research. Based on the findings of the SASSH Survey preliminary features and requirements were prioritised (see section 1.3. A Digital Infrastructure for the European Academies) and developed further into the present concept (AGATE 1.0) by the members of the AGATE project through desk research of the landscape of existing solutions, digital infrastructures and initiatives in Europe focusing on humanities and social sciences and information gained from additional semi-structured interviews with researchers (mainly from research projects run at member academies of the Akademienunion). Valuable input was provided by representatives of European academies and research institutions (researchers, IT experts, digital librarians, etc.) and infrastructure representatives in bilateral talks, during the project's two workshops,[98] and during the AGATE @ H2020 proposal development phase (AGATE 2.0).[99] In this proposal, submitted in response to the H2020-INFRADEV_01_2017 Design Studies call of the European Commission, in March 2017, the core functions of AGATE 1.0 have been integrated into the design of an Open Innovation Research Infrastructure. Especially we would like to thank the members of the AGATE Scientific Advisory Board for their support, contributions and comments on several stages of the concept.

In the following it will be detailed why and how the proposed digital infrastructure of the academies (AGATE) should be developed to have far-reaching positive impact on the three major challenges currently faced by the academies' SSH researchers (see chapter no. 1.2) by:

1. Enhancing general visibility
2. Enhancing findability, access, reusability, and sustainability
3. Enhancing knowledge exchange about standards, good practices, and infrastructure partners

[97] Leathem and Adrian 2015, p. 144.
[98] For the resources of the first and second workshop see Appendix 3: 'AGATE Workshops'.
[99] For further information and a list of the supporters of the AGATE concept 1.0, see Appendix 1, for a list of the consortium of the H2020 proposal, see Appendix 2.

The AGATE concept is based on the overarching aim of offering central access to detailed information about the academies' SSH research (projects) in order to serve as a discovery and matchmaking tool and as an ideal starting point for further knowledge exchange and information. It envisages an expansion of this "Knowledge Map" with additional functionalities that enable a search over the projects' respective digital resources (the AGATE Knowledge Map, see chapter no. 2). The implementation, use and expansion of the AGATE Knowledge Map will be supported by online and offline offers and activities from the AGATE Hub (aka the AGATE website, AGATE staff, and AGATE Consortium and strategic partners). The AGATE Hub addresses additional essential requirements based on user needs such as facilities for knowledge exchange and community building (see chapter no. 3) as well as general dissemination and outreach activities, that are not directly addressed by or related to the AGATE Knowledge Map (see chapter no. 4).

The AGATE concept is presented as individual stages of expansion or modules that can be implemented consecutively or independently from each other. The advantage of this modular approach is that the individual stages and modules will be valuable tools, resources, and activities in of themselves and serve their respective scholarly communities, even if all the stages of expansion of AGATE cannot be realised due to external factors. This modular approach can also be seen as a means of ensuring the infrastructure's sustainability (on sustainability see especially chapter no. 7).

It is neither useful nor possible to plan the entire technical infrastructure in advance ("in the dark"). The organisational and technical refinement, development, and implementation of the individual stages and modules has to be done in a process of piloting and involving members of all relevant stakeholders and user groups according to the given situation and prerequisites (such as project partners of the implementation phase, further developments in the ERA, especially concerning the EOSC, and the SSH ERICs). The objective of this AGATE concept report therefore is to map the basic technical and organisational functionalities and requirements, and match them with existing solutions in order to give recommendations for the next phase.

The recommendations for the future AGATE are guided by the conviction that a maximum of user friendliness, system compatibility and interoperability is needed to fully exploit the potentials of the platform, guaranteeing its far-reaching bottom up uptake by the envisaged user communities.

User groups of AGATE

The potential user groups of AGATE are various and diverse and depend on the respective component or rather module, as not every module addresses all user groups and different communities might use AGATE with different objectives (see

potentials, chapter no. 1.3). A first overview is provided in the following table (for further details see the use cases in chapter no. 2 and no. 3 as well as chapter no. 4).

User group	**Knowledge Map**	**AGATE Hub**
SSH researchers at European academies (PhD students, early-stage researchers)	adopt AGATE as first point of contact for academies' SSH research in Europe, feel an incentive to add information about/data of their projects to database, use AGATE as an information and research tool	use AGATE as an information, training and communication/ networking instrument
SSH researchers outside the European academies	use AGATE as one-stop-information source to get a general overview over academies' research projects according to a specific research question and get in touch with relevant persons	use AGATE as an information, training and communication/ networking instrument
Policy makers and funding bodies	use AGATE as platform for strategy/documentation/planning purposes	use AGATE as information platform, plan/participate in networking and strategic activities
Research coordinators, and press and public relations staff at the European academies	use AGATE as platform for evaluation/documentation/ planning purposes, use AGATE as information tool for science communication matters, promote AGATE to their academie's projects as a platform to showcase their research activities and results	use AGATE as a communication and PR channel, use AGATE as information platform
Media and general public	use AGATE as first stop shop to get a general overview over SSH research projects according to a specific research question and get in touch with relevant persons	use AGATE as an information instrument

Table 1: User groups and envisaged use of AGATE

As the central service of AGATE there will be the Knowledge Map, a database over the European academies' SSH projects and research data. At the outset mainly the academies' SSH researchers need to see the project's potentials and benefits, because they should feel incentivised to become data providers. In the long term, however, AGATE wants to reach and involve a broader audience from the academies, such as: ICT experts, digital librarians, press and public relations staff, science coordinators and managers, as well as outside the academies, especially SSH researchers in different stages of their career, but also funding bodies, policy makers, journalists, and the general public.[100]

As regards the AGATE Hub, it obviously has to be attractive for each of the envisaged user communities, too. While some of the contents, such as information about AGATE's background, events, and publications will be of interest for all user groups, other resources, services, and tools — such as AGATE guidelines and manuals — will explicitly address the academies' SSH researchers, to support them as data providers for AGATE and to facilitate their research and networking activities.

Since the targeted users are situated all over Europe, English is suggested as the general language of the AGATE Knowledge Map and the Hub, and for most of its content. However, especially training material for the SSH researchers and other essential information should be offered in more languages to achieve a high awareness, acceptance, and participation in the different countries. Furthermore, it seems advisable to allow contributions of the community members in several languages, for reasons of inclusivity and linguistic diversity. Therefore, it should also be considered to plan dissemination and outreach activities not solely in English, but also in other languages to gain high attention and acceptance especially among the targeted user groups outside the academies.

As for all pan-European projects, the heterogeneity of the envisaged user groups is a major challenge for AGATE. Therefore, starting point for each of the following conceptualizations of the infrastructure's individual modules and stages is a use case analysis from which central requirements and recommendations are formulated. As a broad acceptance, involvement, and active participation of diverse user groups is of utmost importance for AGATE's long-term success (see also chapter no. 7), measures should be taken from the outset to involve them in the development and to encourage a sense of ownership. Therefore, targeting specific user groups for involvement and general dissemination and outreach measures are proposed to guarantee the wide uptake and acceptance of AGATE (including bottom up and top down approaches) (see chapter no. 4).

[100] This classification into two levels of users was supported by the comment of Elena Gonzàlez-Blanco Garcìa, external peer reviewer of the presentation on user involvement at the second AGATE workshop on January 16, 2017. See Gonzàlez-Blanco Garcìa 2017.

2 AGATE Knowledge Map – A Database of Academy Projects and their Digital Resources

AGATE has five types of information related to academies that it will include in its database — a list of academies, research projects, research publications and digital resources, usage statistics, and AGATE user data as a user-layer of data made up of user lists, curated collections, comments etc. AGATE will make this information available via its search interface and a machine-readable API. The API allows for data distribution between AGATE and partner catalogs and repositories. Most importantly the API allows for the research outputs metadata processing as this data will be of an order of several magnitudes greater that all the other data types being gathered.

As a first module, there will be an *inventory of SSH research projects*, with careful attention given to the search interface design and usability to ensure maximum user engagement. Then, as a second and follow on module, the *research publications & digital resources* of the academies will have their metadata collected, indexed and connected to their respective project entries.

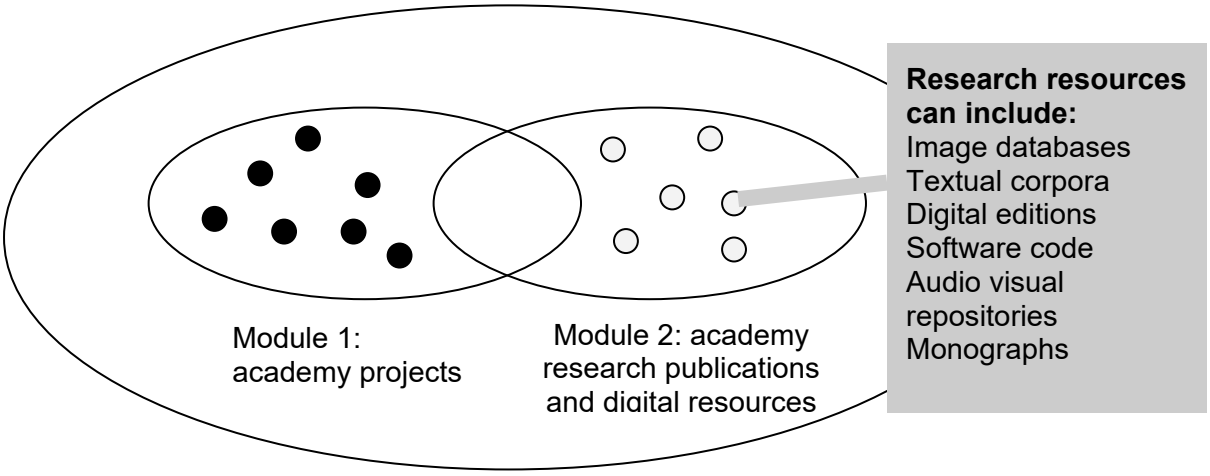

Figure 1: AGATE Knowledge Map

2.1 AGATE 'Knowledge Map': Database of Academy Projects (first module)

In this section, the requirements of the first module, the database of academy projects, are examined and recommendations for the implementation given.[101] The database of the academy projects will act as a foundation and technical test-bed for the later addition of research publications and digital resources by setting up a framework for their cataloguing.

Use cases

The examination of the requirements is derived from the needs of the potential users. A sample set of users have been included in the following table which are representative of a much larger group of users covered in the research (table 2). The table starts with the users' needs and the expected impacts of the AGATE database that have been outlined earlier (see chapter 1.4), followed by the respective source on which these 'expected impacts' are based. Then the use cases are sketched out and the organisational and technical requirements that arise from these scenarios are mapped out. In the final column, the stakeholders related to respective use cases are listed.

[101] For the applicability of use cases as a means of establishing needs assessment, see Cockburn 2006.

Need	Source	Use Case Scenario	Success Scenario	Requirements for the AGATE Database	Stakeholders and Actors
Increase visibility of SSH research projects at European academies	SASSH, Observation, "Facing the Future"[102]	A citizen, a policy maker, a researcher, a journalist, or a student wants to know about academies´ SSH research. At the moment, they have to browse through the websites of the individual academies, often with only scarce information available, and usually only in the national language.	They visit AGATE and have a central access point for information about the research projects of the academies. They can easily get an idea of the richness, variety, quantity, quality, and relevance of SSH-research done at the European academies.	To include a substantial number (best case, all) of European academies' SSH projects To include basic project information To include information about available digital resources in the context of a project To be meaningfully searchable To be multilingual To be intuitively operable To highlight the characteristics of academies´ SSH work, e.g. by clustering academy specific outcomes	Academies as institutions, academies' SSH-projects, general public, policy makers, funding organisations or sponsors

[102] See Duşa, Oellers and Wolff 2014. This book constitutes the proceedings of a conference that took place as part of the SASSH Survey project in November 2013. Among its outcomes it was repeatedly stressed that SSH research faces a lack of visibility, and that creating digital infrastructure would be supportive for its political standing, especially regarding funding.

Facilitate exchange among the researchers at European academies in Europe and beyond	SASSH: wish of researchers, user interviews, desk research[103] Many researchers that took part in SASSH stated that cooperation is hindered by a lack of information on other academies´ projects	A researcher from one academy (or outside the academies) wants to find out about other academies' projects in order to collaborate with these projects.	They can find this information in a structured database	To allow to link projects with their home academies (in case of collaborative projects: to more than one academy) To include information about persons and contacts To be structured in a way that allows identifying shared research interests by searching keywords and browsing faceted lists	Researchers (also beyond European academies)

[103] 'RIA | Digital Humanities Observatory: Digital Research and Projects in Ireland', accessed 10.04.2017.

Facilitate exchange among the SSH ICT community and researchers at the European academies and foster reuse of digital tools and development of common solutions	SASSH, interviews with SSH ICT experts and researchers	An ICT expert is setting up a new (DH) project and looks for inspiration and support.	He/she can search for similar projects on other academies to get information about digital methods and state-of-the-art tools.	To include information about digital tools and digital methods in the project description	ICT experts, researchers, academies' librarians, academies as institutions, SSH-infrastructures, SSH community (beyond academies)
Offer a relatively easy possibility to present a project on the web	SASSH: many projects do not have a web presence because of the lack of resources (financial and human) for activities beyond their actual research	A project coordinator at an academy wants to present his/her project on the web to for outreach, dissemination, and networking goals in an academies' context in an easy way	The database offers the possibility to present his project and update the record whenever necessary.	To be user friendly to enter the project information and to independently update information afterwards. Users can save search results, visit later and add new search result items or edit results	Researchers, Research coordinators

Showcase activity of funder(s)/Transparency	Database showing funding sources by the BAdW,[104] GEPRIS[105]	A citizen, a policy maker, a journalist or a research coordinator want to know which projects are funded by whom.	They find information in the database and can group projects belonging to a common funding scheme	To include information about funding scheme and running time	Funders, academies, researchers, citizens, policy makers
Increase research impact and academies' visibility	BBAW edoc- server (publication repository)[106]	Academies and funder wants maximum dissemination of research publications, outputs and findings.	A comprehensive list of digital resources can be generated with accurate and up-to-date metadata, including links to the sources	Automatic import via API from academy research repositories. Log of imports. Validation of bibliographic records.	Funders, academies
Increase visibility and support academies in demonstrating	SASSH	An academy administrator needs usage statistics of its own digital project,	For administrators easy to read and downloadable statistics on usage in the model	Statics dashboard Downloadable statistics, as data and in report presentation format.	Academies, academy administrators or research leaders.

[104] See 'BAdW | Forschungsvorhaben', accessed 10.04.2017.
[105] See 'DFG | GEPRIS (German Projects Information System)', accessed 10.04.2017.
[106] See 'BBAW | Edoc-Server', accessed 24.04.2017.

impact with usage statistics		web statistics access preferably via a dashboard.	of PiWik[107] Web stats		
Exchange and dissemination	SASSH	A funding application writer from Europe or from outside Europe is looking for EU partners, regional partners.	Faceted search interface to allow user to find partners. Up-to-date contact information for the project.	Search interface and ability to save or download searches. A user could make use of the interface over a long period of time and share search results with colleagues.	Principal Investigators (PIs), project leaders.

Table 2: Use cases for academies project inventory databasing

[107] Open source web traffic analytics platform with user privacy protection. See 'Piwik | Homepage', accessed 06.06.2017.

Synopsis of the use cases

A series of requirements based on the use cases are listed in the column 'Requirements for the AGATE Database' of table 2 (above). These have been grouped into key headers and outlined below:

Simple keywords and browsing faceted lists – the use cases suggest that the database should be organised in a way that it allows identifying shared research interests by searching keywords and browsing faceted lists.

Complex filtering, sorting and relating of information – classification systems and taxonomies will need to be added, as well as the development of a variety of features and interfaces to deal with the need to display and interact with the database to address user requirements. These requirements include; filtering, searching, keyword structuring, faceted search navigation, sorting and the relating of terms. Investment in ongoing user interface design and user interaction research will be important to ensuring success in this area, taking account of the users of AGATE's different interface sections: inputting data, for search, browsing, and data visualisations.

Required fields and relating these to research communities' (digital) standards – the project database will need to capture, at a granular level, information about a large number of projects in order to make search interface deliver insightful search results. Information needed includes the following base sets as examples; project profiles, contact information, project originated publications, and tools and methods (conventional and digital). Controlled vocabularies and a variety of digital cataloguing standards will need to be applied to these base information sets.

Multilingual – since the data providers and potential users derive from various countries and language groups the database should be multilingual.

Summarising the use cases above, the project entries would need to ideally include the following information as individual fields and data sets:

- title of the project/project acronym
- abstract
- research topics (keywords)
- disciplines
- geographic terms
- temporal classification (era of research topic)
- methods, digital methods and digital tools applied
- available digital resources
- digital publication formats (including types of enhanced publications)
- affiliations (to one or more academies, institutes, partners)

- persons involved in the project
- funder information
- running time (this can include information about past projects)
- contact data and further relevant sources of information, e.g. a link to a project's homepage

The definitions of fields, attributes and options to add for information about projects into the database, based on taxonomies and predefined identifiers or otherwise in free text fields, will be the backbone of the database. When determining the attributes, classifications and taxonomies for the AGATE database, the structure of existing relevant repositories and aggregators will have to be considered carefully to guarantee interoperability and easy data exchange.

The reasons for taking the standards and protocols in use in related databasing are multiple. Firstly, the AGATE database, in all its proposed stages of development, will need to exchange data with these other external databases using an API (interoperability). Secondly, in the area of guaranteeing integrity and reliability of information entries, these can be validated on input and when being imported against controlled vocabularies. AGATE will be checking, interacting and exchanging data from API services such as library bibliographic systems, person name repositories such as ORCID, or artifact and textual repositories with metadata schema, such as VRACore.[108] The following sections will touch on a number of factors that demonstrate the importance of the issue of 'standards for interoperability' when working in open data ecologies.

Interoperability and data exchange

Interoperability and data exchange are hot topics in the context of 'current research information systems' (CRIS) that are set up for research management, e.g. by research organisations and research funding organisations (RFOs),[109] as well as in the broader context of the FAIR principles. CRIS is a term that has taken hold over the last decade to describe *digital data systems for research* managed on an institutional level. What is symptomatic of CRIS systems are: *managing change* in technology (hardware, software, standards and methods), having to integrate multiple systems, and supporting organisational change — as habits within institutions themselves adapt and evolve in relationship to new methods and working practices.

[108] 'Library of Congress | VRA CORE - a Data Standard for the Description of Works of Visual Culture', accessed 12.04.2017.
[109] Science Europe 2016.

Though not particularly targeting funding agencies, the AGATE database has obvious overlaps with these CRIS systems. The concern of RFOs in this context is to support the development and use of unique identifiers for relevant important categories (e.g. organisation, person, funding agency). Such identifiers are needed to avoid doubling of entries, e.g. caused by different spellings of the same entity and to prevent confusion, e.g. of an individual person's name. At the same time these organisations have a strong interest in the use of interchangeable systems and formats that allow merging and combining datasets.[110] Furthermore, interchangeability can reduce double efforts through the automated re-use of existing datasets and avoid manually entering the same data over and over in different systems, "enter once, reuse multiple times".[111]

Therefore, the database design should pay special attention to entity mapping with existing systems, especially those that are dedicated to harmonisation and data exchange. In the German context, the Kerndatensatz Forschung (Research Core Dataset)[112] is such an initiative. The Research Core Dataset is advanced and makes use of existing classification to increase its interoperability. Unfortunately, its international applicability is restricted by its national scope.[113]

In the international context, the pan-European non-profit organisation euroCRIS is a dedicated advocate of interoperability and data exchange of research information. EuroCRIS promotes the CERIF data model (Common European Research Information Format). EuroCRIS is in charge of the development, maintenance and further advancement of CERIF and cooperates internationally to harmonise systems using CERIF, e.g. with CASRAI[114] and OpenAIRE. The use of CERIF is recommended to its member states by the European Union.[115] The *CERIF Semantic Vocabulary* defines a large number of so called classes using the Resource Description Framework (RDF).[116] CERIF could play a significant role as a benchmark for AGATE, informing design decisions and implementation as noted in the list below:

1. CERIF is also referred to as a "data exchange model" and in this light can be understood for AGATE as an interface that makes datasets reusable.[117] This is in line with the recommendation to be connectable with other research information systems.

[110] Science Europe 2016, p. 5.
[111] Science Europe 2016, p. 4.
[112] The Research Core Dataset (Kerndatensatz) is recommended by the German Council of Science and Humanities (Wissenschaftsrat) and further developed by several infrastructure related organisations, see 'Wissenschaftsrat | Kerndatensatz Forschung', accessed 13.04.2017. It suggests attributes and classifications for third party funded projects (Drittmittelprojekte).
[113] By now almost all publically available information and the data set itself are in German.
[114] CASRAI is working on commonly applicable and reusable datasets of research information, see 'CASRAI | Homepage', accessed 12.05.2017.
[115] Jörg 2012 et al. (ed.), p. 2.
[116] Sicilia 2013.
[117] See 'euroCRIS | Homepage', accessed 12.05.2017.

2. A research information system can be designed based on a subset of the full CERIF model, e.g. for projects, organisations or publications. The vocabulary is very comprehensive and includes almost all attributes mentioned above. If terms do not yet exist, it is possible to add such terms in coordination with the euroCRIS developers.
3. AGATE is intended to start with a project database and be gradually enhanced by features that allow browsing through the digital resources of projects. Especially the capacity to link certain entities as "projects" with other entities such as "publications" meets the needs of AGATE.[118]
4. Several aggregating services such as OpenAIRE are already compatible with euroCRIS solutions.
5. CERIF includes multilingual features to ensure that key terms (entities) are made interoperable across languages, e.g. names, titles, descriptions, keywords, abstracts, etc.[119]

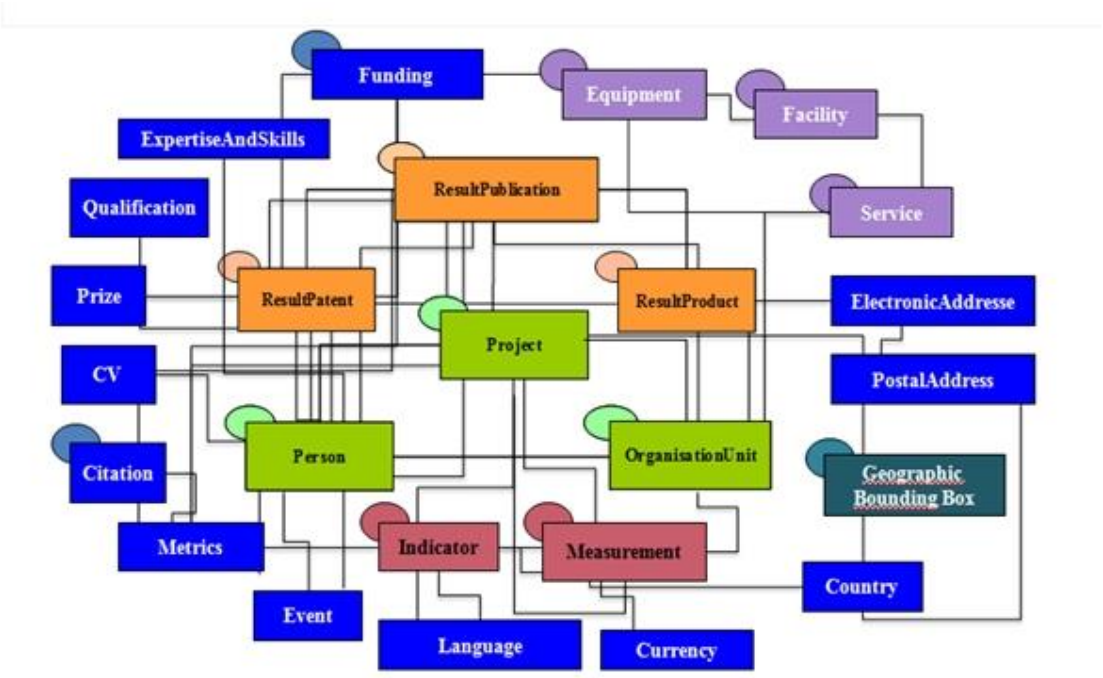

Figure 2: The CERIF data model[120]

[118] See 'euroCRIS | Main Features of CERIF', accessed 12.05.2017.
[119] Multiple language features of CERIF 1.3 Full Data Model, see Jörg et al. (ed.) 2012, p. 34.
[120] See 'euroCRIS | Main Features of CERIF', accessed 18.04.2017.

Relevant persistent identifiers, taxonomies and semantic vocabularies

In order to structure and link the content of the database it is necessary to apply identifiers and taxonomies wherever possible and useful. Identifiers and taxonomies provide clear definitions for the data entered by a user. A semantic vocabulary is necessary for linking the data of different categories or for data exchange with and re-use by different systems.

Examples for relevant identifiers include: ORCID[121] (persons), ISNI[122] (persons as contributors to media), ISBN[123] (publications), DOI[124] (documents), ISRC[125] (audio recordings), ISAN[126] (audio visual works), ISSN[127] (serialised publications), ISTC[128] (15th-century European printing), ISWC[129] (musical works), GND[130] (persons), Crossref Funder Registry[131] (funders).

As a case in point ORCID, a non-profit organisation endorsed by the DFG,[132] is a recent interesting service that provides researchers with a unique identifier (UID). In 2017 the Technische Informationsbibliothek (TIB), Hannover, has taken over the management of coordinating German use of ORCID with a large number of institutions signing up to the consortium.[133] On the ORCID website the researcher can set up a profile page including biographical and contact data as well as educational information and a list of publications that can be updated manually and automatically. This information would be very valuable for the AGATE project database and should be included at least by linking the ORCID IDs of databased project staff (if available).[134]

For the definition of Digital Humanities specific fields such as methods or digital research tools TaDiRAH, the Taxonomy of Digital Research Activities in the Humanities, would be useful:

> "This taxonomy of digital research activities in the humanities has been developed for use by community driven sites and projects that aim to structure

[121] See 'ORCID | Homepage', accessed 12.05.2017.
[122] See 'ISNI (International Standard Name Identifier) | Homepage', accessed 18.04.2017.
[123] See 'International ISBN Agency | Homepage', accessed 12.05.2017.
[124] See 'DOI Foundation | Homepage', accessed 12.05.2017.
[125] See 'ISRC (International Standard Recording Code) | Homepage', accessed 12.05.2017.
[126] See 'ISAN (International Standard Audiovisual Number) | Homepage', accessed 18.04.2017.
[127] See 'ISSN (International Standard Series Number) | Homepage', accessed 12.05.2017.
[128] See 'ISTC (Incunabula Short Title Catalogue) | Homepage', accessed 18.04.2017.
[129] See 'ISMWC (International Standard Musical Work Code) | Homepage', accessed 18.04.2017.
[130] See 'DNB | Gemeinsame Normdatei (GND)', accessed 12.05.2017.
[131] See 'Crossref | Funder Registry', accessed 13.04.2017.
[132] See 'DNB | Projekte DE – ORCID DE – Förderung der Open Researcher and Contributor ID in Deutschland', accessed 13.04.2017.
[133] See 'ORCID DE | Post Heinz Pampel 12.01.2017', accessed 12.01.2017.
[134] See 'ORCID | Homepage', accessed 12.05.2017.

information relevant to digital humanities and make it more easily discoverable. The taxonomy is expected to be particularly useful to endeavors aiming to collect information on digital humanities tools, methods, projects, or readings."[135]

The initiative began with the objective to enhance discoverability of resources in the DiRT directory and the DARIAH-DE bibliography and evolved to be a service for digital humanities research in general and is by now available in several languages.[136] For this purpose TaDiRAH provides definitions and classifications especially in the fields of "Research Activities", "Research Objects" and "Research Techniques".

Furthermore, controlled vocabularies and data models, such as CIDOC CRM,[137] will have to be considered to achieve interoperability. They can be identified for example via the Basel Register of Thesauri, Ontologies & Classifications (BARTOC).[138] Especially regarding semantic vocabularies the Linked Open Vocabularies (LOV)[139] initiative provides a catalog of the semantic vocabularies ecosystem with a dataset that consists of more than five hundred vocabularies.[140]

Given the diverse and dynamic field of persistent identifiers, taxonomies and semantic vocabularies, at this point it is of no use to recommend a single vocabulary that would fit all needs of the organisations and researchers that will participate in AGATE, but important to draw attention to the complexity and relevance of this field for future AGATE development. For example, the 2016 "State of the art report on open access publishing of research data in the humanities"[141] identified for three main areas of scholarly activities in the arts and humanities (archival, electronic scientific text encoding and bibliographic fields) several commonly used controlled vocabularies for each of these fields. The authors concluded with saying: "we want to emphasize the difficulty of predicting what will be the standard that will prevail in the coming five years".[142]

Relevant databases of research projects

Six research databases have been selected after desk research within the project as examples of good practice in the field of research infrastructure. These comparator

[135] See 'DARIAH-DE and DiRT | TaDiRAH (Taxonomy of Digital Research Activities in the Humanities), accessed 15.10.2016.
[136] See 'DARIAH-DE and DiRT | TaDiRAH (Taxonomy of Digital Research Activities in the Humanities)', accessed 18.07.2016.
[137] See 'CIDOC CRM | Homepage', accessed 18.04.2017.
[138] See 'BARTOC | Homepage, accessed 15.10.2016.
[139] See 'OKF | Linked Open Vocabularies (LOV)', accessed 18.04.2017.
[140] See Vandenbussche et al. 2015, p.2
[141] See Buddenbohm et al. 2016, pp. 35–38.
[142] See Buddenbohm et al. 2016, p. 39.

databases serve as a way to benchmark key features of AGATE, such as working across multiple academies and cataloging their projects, comprehensively recording an academy's research output, and ensuring that adequate classifications for SSH and DH research projects are considered. Additionally, these six databases have been compared and analysed in a spreadsheet (Appendix 4: 'Comparison Research Databases') against the database fields finally recommended for AGATE, and any special considerations flagged up in the spreadsheet's "column J: Challenges/Needs for research or coordination" as well as some being mentioned below.

Database comparators:

1. Project database of the Union of the German Academies of Sciences and Humanities (Germany) – http://www.akademienunion.de/forschungsprojekte
2. Project database of the Mainz Academy of Sciences and Literature (Germany) – http://www.adwmainz.de/projekte/geistes-und-sozialwissenschaftliche-projekte.html
3. Wissensspeicher (Digital Knowledge Store), Berlin-Brandenburg Academy of Sciences and Humanities (Germany) – http://wissensspeicher.bbaw.de
4. DRAPIer, Digital Research and Projects in Ireland (Ireland) – http://drapier.dho.ie
5. Digital Humanities Registry – CLARIAH (Netherlands) – http://dh-projectregistry.org
6. OpenAIRE: project entries (EU) – https://www.openaire.eu

1. Project database of the Union of the German Academies of Sciences and Humanities (Akademienunion, Germany)

The Akademienunion has a web based listing of research projects. The projects can be searched, as well as being accessed as listings pages using the following categories: by content type of the project - such as dictionary or edition, or by top-level subject categories e.g. archeology or musicology, etc. Each project entry then has a description, basic project information, links and contact details. The project information is held in a web CMS and is an example of cataloging projects produced by multiple academies. The project information display has limitations, the search is very basic, with no filtering of results and some of displays are links to pre-made static lists. Additionally, the display of information about project is not very prominent, even though the site has comprehensive and high-quality data on projects. If the project entries were also in English, then this would help promote the projects internationally.

2. Project database of the Mainz Academy of Sciences and Literature (Germany)

The Mainz project database holds extensive information of past and current research projects of the academy. The database displays project descriptions and information, as well as listing related publications, news, events and media. The database has further developments in progress that allow for an increase in scale, features and further details on items databased. The well thought out design of the interface, combined with highlighted media, provides a way for the user to access the project database as an information tool for work purposes, or to casually explore the academies research projects because the page layouts are presented in a visually appealing way.

3. Wissensspeicher (Digital Knowledge Store), Berlin-Brandenburg Academy of Sciences and Humanities (Germany)

The Digital Knowledge Store is the first and only academy portal to provide centralised access to all of an academies projects, publications and digital resources, and to "connect them semantically".[143] It is of special significance to AGATE because it collates all the 'digital resources' of an academy and thus provides many insights for AGATE's second module covering the search over research resources.[144] It is possible to browse and search the academy's resources and filter the research results by several classifications, for which the Digital Knowledge Store uses the subject area classification system of the German Research Foundation (DFG),[145] as well as a customised RDF metadata model.

The Digital Knowledge Store has new features being developed in beta for release in 2017. As the Digital Knowledge Store has such a large amount of content, which is so wide-ranging, many of the highlighted system development issues of AGATE database and will no-doubt be addressed and the results contribute to AGATE finding solutions to similar design challenges. These insights will be of use for AGATE's design in its implementation phase. These design challenges include: classifications of SSH disciplines, and categorisation the varied content types of scholarly research publication and digital resources that can be produced[146] (see: Appendix 4: 'Comparison Research Databases', cell J22 and J21).

[143] See 'BBAW | Digital Knowledge Store', accessed 14.05.2017.
[144] For the slides of the project presentation at the first AGATE workshop see Czmiel 2016.
[145] See 'DFG | Subject Areas Classification', 26.04.2017.
[146] See Worthington and Kral 2014.

4. DRAPIer, Digital Research and Projects in Ireland – Royal Irish Academy, DH Observatory

One of the few SSH project databases that shows significant accordance (especially regarding the terms "methods" and "formats") with AGATE is DRAPIer.[147] This comprehensive registry of Irish DH projects was developed by the Digital Humanities Observatory of the Royal Irish Academy. It is searchable by title and keyword but also methods, formats and standards.

Interestingly DRAPIer had collected extensive data but had not inputted this data into specifically design fields. This was not only a problem of DRAPIer but of other databases too. As an example, not having related organisations entered in their own fields means that Semantic Web analysis cannot be applied so easily to determine chronologies of cooperations between organisations, or interrelationships between multiple organisations. The database is no longer actively developed as it came to the end of its funded period, but the database is hosted and kept online by the Royal Irish Academy (see: Appendix 4: 'Comparison Research Databases', cell J12).

5. Digital Humanities Registry – CLARIAH (Netherlands)

The Digital Humanities Registry presents a great number of DH projects conducted in the Netherlands. It has been built and is maintained by CLARIAH, the national branches of CLARIN and DARIAH cooperating in the Netherlands.[148] The database is most notable for its qualitative research in compiling and designing the database, this has been done in consultation with many of the main Netherlands research database maintainer institutions.[149] The categories created by the database and its use of standards, controlled vocabularies and the classification of 'digital methods' is also of importance to AGATE and would be something to make use of and collaborate on with the Digital Humanities Registry in the future (see: Appendix 4: 'Comparison Research Databases', cell J25).

6. OpenAIRE: project entries (EU)

The European Union funded OpenAIRE initiative differentiates in its search interface between three types of entities: publications, datasets, and projects. What is highly relevant for AGATE is that it searches for digital resources over numerous dispersed repositories, giving "access to metadata about projects funded by a selection of

[147] See 'RIA | Digital Humanities Observatory: Digital Research and Projects in Ireland', accessed 14.05.2017.
[148] See 'CLARIAH | DH Project Registry', accessed 14.05.2017.
[149] Information drawn from an exemplary project entry, see 'CLARIAH | Access to and Sharing of Research Data from Public Funding - Global Research Village OECD', accessed 14.05.2017.

international funders".[150] By now there are eight funders that contribute information about their respective projects via APIs (May 2017).

OpenAIRE currently contains more STEM related projects and important SSH and DH properties like methods or temporal terms are missing. But as OpenAIRE offers a very valuable open model of working for AGATE to follow, and in order to be compatible with OpenAIRE, the selection of categories for AGATE should be mapped onto OpenAIRE to allow the free exchange of data.

Collaborations

The Mainz Academy has received a mandate to develop the current database of the Akademienunion further based on the findings of the AGATE project and using technology already employed for the Mainz Academy database. This pilot project starting in 2017 also offers an opportunity to consider and test some of the issues related to operating on a pan-European level.[151]

Consultation and surveying of the comparator research databases has opened up the way for potential collaboration between AGATE and these initiatives. Experts from these initiatives have been present at the two AGATE workshops and provided several inputs. The different comparison research databases each encompass specialisations of interest to AGATE. Outlined are a few key examples of areas of cross over with AGATE and where cooperation could be developed. With the Digital Knowledge Store, it is the complexity of database with many varied publication types. Regarding the Digital Humanities Registry, it is on the topic of DH classifications, and with OpenAIRE it is the technical mechanics of data interoperability. In the case of OpenAIRE it is also their welcomed increase in the profiling of SSH/DH research and the inclusion of DH issues in the design of its CRIS system to accommodate this content. This accommodation of SSH/DH issue into their CRIS system is an important step in what is predominantly a STEM influenced design field of CRIS infrastructures.

Social aspects and incentives to involve academy projects as data providers

There are social aspects to be considered which also have an impact on the technical design. For example, AGATE has to create trust and confidence with its users, a foundation of this trust is having a critical mass of academy projects from SSH and ensuring the project data is up-to-date.

[150] See 'OpenAIRE | OpenAIRE API Documentation', accessed 14.05.2017.
[151] For the slides of Torsten Schrade's presentation at the strategy meeting of the Executive Committee of the Union of the German Academies of Sciences on March 1, 2017 see Schrade 2017.

Service design

The methodology AGATE employs for understanding these social aspects and ensuring that they lead the technical design is called *service design*. Service design has its basis in participatory design ideas as outlined by Eric von Hippel and his book *Democratising Innovation*.[152] For a design process to build AGATE, the service design methodology involves identifying the actors involved in the proposed system and then finding a variety of ways to gain insights into their needs. This design method needs to be an ongoing and repeated process, of consultation, prototyping and testing, then refinement, and implementation, and the again back to consultation. This repetitive design process is then put in place via 'Agile'[153] project management method where short terms goals are set to address the gathered requirements and evaluate the design decisions made.

The following points outline the considerations we have in mind for ensuring AGATE reaches a critical mass of projects from academies from SSH and DH sectors.

Defining projects and academy projects

As shown in the introductory chapter the academies in Europe are very diverse in size, mission, and organisational form. Some academies have their own research projects, some run rather as autonomous institutes and some cooperate exclusively with universities. Also, the forms of the projects vary: a digitisation project of an academies´ library working on a collection builds on very different workflows than a research project working on an edition. Therefore, it seems advisable not to be too prescriptive with the term "academy project" and apply a pragmatic approach. It is suggested to set a minimum set of features, for example a clear affiliation to an academy and at least one contact person that is associated with the project. Since the focus of AGATE lies on Social Sciences and Humanities the project should furthermore be linked to these disciplines, be it directly or in the production or preservation of data sources related to these disciplines.

Incentives to involve academy projects as data providers (bottom up and top down approaches)

As shown in the use cases, the database must include a substantial number of the SSH projects that are conducted at the European academies in order to be a useful tool and be representative of the sector. Thus, measures to recruit users who will contribute project data to help ensure we have the required number of research

[152] See Hippel 2006.
[153] See Agile Business Consortium 2017.

projects listed in AGATE should be considered a priority from the very beginning of the implementation process. This includes the application of easy-to-use workflows in the process of data collection that enable the data providers to independently edit and update their own project entry.[154] This would be a valuable aspect to create a feeling of ownership among the data providers who at the same time constitute an important user group. Gathering data on individual project entries can be done in two ways:

1. **Content Management System (CMS) data entry** – The project database is located on a server that is accessible with different access rights (partly restricted), that allow for changing certain entries without jeopardising the master database itself. Essentially, users are allowed to create a *data layers* on the master data set. This calls for a CMS with an elaborate system of user roles and access rights, applicable for thousands of users. In the SASSH report, six hundred projects were identified, from one hundred and fifty research organisations in thirty-one European countries, so the users involved can easily reach high numbers based on this sample project figure.[155]
2. **API data import and interaction** – The project data that is located on external sites, e.g. the project´s or academy´s website is harvested automatically by the AGATE project database via an API. This would have the advantage that the updating procedure is being conducted in the user's own familiar environment and does not require accessing the server of the AGATE project database by external users.

While the aforementioned incentive targets the project staff itself and can therefore be considered 'bottom up', AGATE also needs ways to encourage users from a 'top-down' perspective, for example research managers and funding agencies with incentives such as useful statistical tools. Using the service design approach we would look at other incentives to attract these user groups to add projects. Barbara Ebert of the head office of the German Council for Scientific Information Infrastructures (RfII) recommended in her comment at the second AGATE workshop regarding this topic to "Create multiple ways to access and reuse the data: easy generation of simple lists, analytical tools, or web services".[156] In the case of the AGATE database one could design tools that provide a list of all projects of a single academy or another list presenting all projects of a certain discipline from a selected country. Depending on the attributes that are featured, several scenarios that are interesting from a funder's

[154] This was also confirmed by Barbara Ebert of the German Council for Scientific Information Infrastructures, who commented on a draft of this module as an expert speaker at the second AGATE Workshop. See Ebert 2017.
[155] See Leathem and Adrian 2015, p. 129.
[156] See Ebert 2017.

perspective are conceivable. At a later stage access statistics or even bibliometric features could be included once the project database is connected with *search over publications* functionality in 'module 2 – research publication and digital resource databasing'.

Providing up-to-date data and using Continuous Integration (CI) technology

In order to provide high quality data, the entries must be up-to-date. This necessitates regular data curation that includes regular updates concerning changes in the staff structure of projects, adding new publications or replacing dead links. On the one hand this could be tackled by training, tutorials and guidelines for the data providers (e.g. the researchers in the projects) which can be supplied by the AGATE Hub. On the other hand, a certain degree of manual inputting by the administrators of the AGATE database will be inevitable. The administrators would be responsible for maintaining feedback mechanisms to report errors and fix them and to further develop the system based on these experiences.

It must be understood that the database is not a static component that is completed in a one-time delivery. Not only does the information in it need keeping up-to-date, like any contacts record catalog, but additionally it needs technical maintenance on many levels. For example, interface conventions impacted by new screen use like mobiles and tablets, or additions of new data sources and system design to accommodate this data and resulting new user requirements. If the database is viewed as an ecosystem, with many complex parts needing to be taken into account, then adequate planning can lead to its smooth running.

To ensure that the AGATE database is always up-to-date a new computing process, Continuous Integration (CI)[157] will be used. CI initially rose in popularity in software coding but is now finding applications in other fields such as publishing and databasing. CI is an automated process of continual validation, in some cases real time. Validation takes place against a set of data which is tested and either passes or fails. The CI validation takes place according to a time schedule, and if the data being tested fails, a log is kept of what data has failed and why. This way an operator can make changes and run a manual CI test to pass the data tests. AGATE would make its own custom validation rule set, which would contain controlled vocabularies, standards and required fields, etc. When using CI an administrator or authorised users have a trusted way of knowing that data they use has met certain standards over time and don't need to re-check the data or take up time of administration staff.

[157] See 'ThoughtWorks | Continuous Integration', accessed 03.05.2017.

Recommendations

Based on the examination of; the technical requirements, the requirements for interoperability, the survey of comparable databases and the social aspects, we formulate the following set of field recommendations for the AGATE database (see: Appendix 4: 'Comparison Research Databases').

Key fields for AGATE to include in the database:

project title
partners
project website
digital methods used
acronym
running time
available digital resources
digital tools used
project ID
status
project type
digital formats
parent academy
contact data
discipline
funding
institute
Abstract
temporal terms
early stage researchers involved
country
persons involved
geographic terms
involvement of citizen science

Each of these fields has to be mapped onto existing terms in classification schemas and/or controlled vocabularies. The table in the appendix (see: Appendix 4: 'Comparison Research Databases') suggests equivalent categories in the respective databases that have been examined in the survey of comparable databases. It is suggested to use the two following research classification schemas: 1. the German 'Research Core Dataset' (Kerndatensatz Forschung), and 2. the 'OpenAIRE project entries'.

Database design:

- For the definition of the fields in the AGATE database we recommend to make use of the CERIF semantic vocabulary.[158] TaDiRAH, the 'Taxonomy of Digital Research Activities', could complementarily provide the discipline specific definitions and classifications that CERIF does not feature.

[158] In the course of the AGATE project there has been an important exchange with the board of euroCRIS and the developers of CERIF. They have informed the authors that customised adaptations are possible (and very common) when applying CERIF.

- Regarding the workflows, the operation of the database for end users should be easy-to-use. To meet this goal, the following technical design methods need to be used: Service Design as an overall methodology, User Experience (UX) and User Interface (UI)[159] for interface design, and Agile methods for project management.
- The process of data collection should enable the data providers to independently edit and update their own project entry.
- The database should feature tools for an easy generation, presentation, export and publishing of selections of data from the database. These could be to show off activity in a specific field over times, or as project lists for conference literature. It should include simple lists, visualisations, and statistics to be an attractive and supportive tool for the participating institutions.
- Furthermore, data curation mechanisms and an administration of the platform should be institutionalised, so integrated into day-to-day working, procedures and policies of academies.
- An API will need to be implemented, documented and supported.
- A dashboard admin interface will be needed to cover the following areas; editing, import and export, field mapping, workflow for job monitoring, flagging errors or content status, as well as bug reporting and support requests.

2.2 AGATE 'Knowledge Map' expansion: Search function with added indexing of digital resources (second module)

The second module would expand the project database by including digital resources of the academies projects, with a search interface function over the entire database. As in the first module the examination of the search function starts with a focus on the users. Accordingly, in the table below an overview of concrete requirements is arrived at on the basis of the initially identified general needs and use cases.

[159] 'Quora | What is the Difference between UX and UI Designer and Web Designer?', accessed 25.04.2017.

Use cases

Need	Source	Use case scenario	Success Scenario	Requirements for AGATE Database	Stakeholders and Actors
Enhance general accessibility of academies' digital resources by increasing their findability (and thus stimulating reuse)	SASSH	A researcher compares medieval manuscripts of diverse European origins and is looking for relevant digital resources	She/he finds high quality material stemming from projects of several European academies connected with information about the research projects	To allow browsing through the digital resources. To allow connecting of digital resources with the project entries in the database. To provide access to datasets with APIs, full digital texts. To automate harvesting of collections of OA material provided by academies (e.g. via OpenAire, Europeana, CLARIN VLO, etc.)	Researchers, students (as research tool), general interested public (as information platform)
Create a platform to showcase and find electronic academy specific 'enhanced publications' and research material that do not fit in common library catalogues	SASSH, interviews	A research coordinator would like to make her/his corpus project visible for a broader audience	The data can be found via AGATE, as a platform of the community and the attention of more researchers from relevant disciplines can be drawn	To develop a solution to showcase enhanced publications. To include resources that are work in progress	SSH researchers at academies and the SSH research community in general, interested public

Foster access to digital resources for DH in general, stimulate explorative aspects and serendipity	Observation, wish of researchers, *Facing the Future*	Encourage researchers to browses subject fields outside of their specialisation in the projects and resources covered by AGATE.	She/he learns by chance about projects or research she/he has never heard of before.	To allow intuitive and random browsing, smart suggestion function	SSH researchers, students
Stimulate (re)use	Interviews	A project databased by AGATE (including its digital resources) has ceased to operate.	The project entry and related digital resources can still be found and accessed.	To include content of projects that have been concluded	Researchers, project coordinators, funders
Research dissemination	This is a common need across projects and for individual scholars	A researcher needs to create special showcases of a project for a conference, as a kind of *digital poster*, including a digital downloadable bibliography.	Interface for making a special collection of research publications. Publish these collections into landing pages. Have collections downloadable with bibliographic collection.	Provide landing page creation area to publish web page. Collections authoring admin console. Downloadable bibliographies	SSH researchers, research communities
Increase exchange and visibility, stimulate (re-)use	SASSH	A project releases a major digitisation project and wants the research community to access data for Open Science reuse.	Description and fields for profiling a project or resource properly demonstrate the extent of its resources (including links) and reuse information.	Easy to input web interface. API import functions Field validation with controlled vocabularies Field for licence information	Projects, academies, researchers

Bulk institutional and publisher use. Visibility and project research findings and publications reuse (impact)	Desk Research	An OA publisher, library or museum wants to negotiate bulk digital access to digital publications. E.g. Open Library of Humanities,[160] or Getty Foundation Online Scholarly Catalogue Initiative.[161]	Publications have accurate licensing information. Sources links have 'continuous integration' validation indicators to allow for automatic data access (in the manner of Travis CI)[162]	OpenIPR licence categories available. Licence categories are shown to be compliant with region and funder requirements.	Research publishers, research communities, scholars
Exchange and even tool reuse	SASSH	A museum curator requests digital assets in a specialised metadata format, such as VRACore for use in their digital repositories.	Show the formats that metadata is available in. Invite research communities to contact AGATE to request new features and collaborate.	Ability to convert metadata between formats.	Museums, Libraries and Archives sector, curators, research community
Dissemination, exchange, visibility	SASSH	Crowd-sourcing or Peer-sourcing. A project wants to improve its metadata on its publishing back catalogue for maximum discoverability and	Guidelines for supplying metadata via web input forms or via API automatic provision.	Import of metadata via API. Metadata validation.	Scholars, projects leaders, funders, academies

[160] See 'Open Library of Humanities | Homepage', accessed 12.05.2017.

[161] See 'Getty Foundation | Online Scholarly Catalogue Initiative', accessed 12.05.2017.

[162] See 'Travis CI | Test and Deploy Your Code with Confidence', 03.05.2017.

		invite research community collaborator to contribute to the effort.			

Table 3: Use cases for search function with added indexing of digital resources

Synopsis of the use cases

Databasing the *digital research resources* will greatly increase the amount of data being brought into AGATE and the complexity of handling this data with the increase of data fields to cover. There is a great need to import and export data from a wide variety of repositories, and to check this data against controlled vocabularies, as well carry out other validation checks. Search functions require that there are checks placed on the manual *input* and *automatic* import of data and then on the design of the search interfaces to ensure the information works for our different user groups.

From the summary of the requirements above the AGATE search function is expected to facilitate the following:

- To discover and allow browsing through digital resources.
- To make use of continuous integration technical frameworks which will ensure the quality of the data being searched via validation. That the data has; valid information such as; correct ISBNs, that external resources are available, or that data is in formats required.
- To connect these resources with the information in the project database.
- To harvest information about collections of open access material produced by academies' projects already provided via aggregating initiatives and infrastructures such as OpenAIRE, Europeana, or CLARIN, via APIs.
- To offer itself an API for other aggregators or repositories to use.
- To give a special emphasis to enhanced publications, and works in progress.
- To cover information about digital resources of projects that have been concluded.
- To have intuitive and easy-to-use interfaces.
- To convert metadata between formats.
- To provide functions to create custom digital research resource collections and publish these to web landing-pages.

Technical solutions

The use cases call for a solution that is capable of *harvesting* different resources or respectively information about these resources (metadata) into AGATE from storage systems located in many different countries. The overarching reason for this need to harvest data about digital resources is that in many cases metadata have already been generated elsewhere outside of AGATE (predominantly digital bibliographic and archive catalogue records).

At the moment, many different storage solutions for the digital resources of academies are in place that range from unstructured databases to professionally curated publication repositories and data centers: some data is saved on central academies´ servers,[163] some on repositories located at the library of an academy,[164] some is stored externally in third party repositories (e.g. on Zenodo),[165] and some on a local server as part of a project homepage. The data derives from such disparate disciplines such as archeology, political sociology, or musicology, and the data types vary from: 3D visualisations & scans, image deep-zooming using IIIF, GIS geolocative data, audio and video files to text corpora in TEI, digital editions and monographs. Therefore, it can be expected to constitute a major part of the workload during the implementation to deal with the various forms and qualities of the accompanying metadata. Moreover, access to many of these digital resources is not (yet) open, restricting the service AGATE can offer for the aggregation of the metadata anyway. It is important to note here that closed sources can still be made usable within AGATE by utilising their metadata — descriptions, classification and the location of the source. It is then for the user to resolve access to the source, either via an authorised online login or arranging a visit to an archive. The vast scale of the data sets and the thousands of work hours involved in their creation and maintenance demand the use of automated API access. AGATE's added value is to provide the following: to validate, structure, filter, and present the results to the user in a meaningful and attractive way, including links to the digital resources on their original repository locations.

This leads to the question which type of data should actually be made accessible for a given digital research resource: the metadata of a resource (its catalog record)[166] or the data of the resource itself (e.g. full text). The general user habits shaped by the use of commercial search engines or discovery services may suggest that the display of a search query result would automatically include data of the actual resource, e.g. in the form of teasers, such as text fragments or thumbnails of pictures. This might be applicable and advisable for resources such as PDF, HTML pages, or images, but is more challenging for audio data or 3D visualisations. For these more complex digital objects we would need to go beyond the conventional 'search results' as a textual list. It is worth noting that search results can be much more varied than just text based lists, take these two examples. First, Google Books,[167] and second, the image deep-zoom software 'International Image Interoperability Framework' (IIIF)[168] and its search inside features as implemented by the Wellcome Library.[169] And for searching inside of

[163] E.g. the edoc server of the BBAW: see 'BBAW | Edoc-Server', accessed 24.04.2017.
[164] E.g. the repository of the Library of the Hungarian Academy of Sciences (REAL): see 'MTA | Repository of the Academy's Library', accessed 14.05.2017.
[165] See 'Zenodo | Homepage', accessed 14.05.2017.
[166] For the definition of the terms "resource", "item" and "record", see: Lagoze et al. (ed.) 2002.
[167] See Worthington 2015.
[168] See 'IIIF (International Image Interoperability Framework) | Homepage', accessed 14.05.2017.
[169] See Wellcome Library 2017.

even more complex and extensive textual resources that are typical for academies´ research — like digital editions, databases or text corpora — a full text search, or web index created by a web crawler will not necessarily be the best choice to achieve satisfactory search results.[170]

For reasons of feasibility, it seems advisable to concentrate first on the possibilities that metadata aggregation offers a better way to work with metadata or metadata catalogs provided by repositories. At the same time, we can follow closely the development of other transnational (discipline specific) aggregators and initiatives on the roadmap to the European Open Science Cloud (EOSC)[171] before deciding to aim at enhancing the search functionalities to search "in the data" of the research resources. Internal document searching is often better tackled via discipline specific RIs and Virtual Research Environments (VREs).[172] The efforts of collecting and including the metadata of the academies digital resources into AGATE should be combined with general awareness raising and hands-on advice for achieving more open access rights to the academies' digital resources in general and implementing metadata standards and application programming interfaces (APIs).

APIs

AGATE uses an API to aggregate metadata into AGATE from digital resources stored in repositories, distributed on different servers across the Internet. An API is the technical method used to allow a repository on a server connected to the Internet to communicate with another server anywhere on the Internet. The communication can involve exchanging data but also queries. As an example query using an API, AGATE can ask the Deutsche Nationalbibliothek's API[173] to perform a search on its bibliographic system and send AGATE the search results back with a custom set of specified fields, for example AGATE only wants the titles of a list ISBN publication IDs it provides. A well-known and common API specification in the cultural heritage sector for aggregating data is the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).[174] The advantage of the OAI-PMH is its low threshold for use, because it is based on the *Dublin Core Metadata Element Set* (DCMES), a basic set of (only) fifteen elements to describe a resource.[175] Thus, by now many repositories already provide such an API and if not, the installation of this interface seems not to be a significant obstacle. But this advantage also accounts for some criticism regarding the final search

[170] See Beall 2008.
[171] See Ayris 2016.
[172] See Knowledge Exchange 2017.
[173] See 'DNB | "Bibliographic Services" Data Service', accessed 13.06.2017.
[174] See Open Archives Initiative 2017.
[175] See 'Dublin Core | Dublin Core Metadata Element Set, Version 1.1.', accessed 14.05.2017.

functionality and metadata quality. The original basic DCMES constrains search functionalities to its fifteen elements, and the application of all fifteen elements is not even fully mandatory.[176] Furthermore, not all Dublin Core usage guidelines require the use of controlled vocabularies.[177] This can account for inaccuracies in the metadata aggregated from different sources.

The OAI-PMH is for example used by OpenAIRE, the biggest research data aggregator in Europe. The core mission of OpenAIRE is "pulling together and interconnecting the large-scale collections of research outputs across Europe"[178] by making repositories from various institutions and disciplines accessible. Jochen Schirrwagen, one of the developers of the OpenAIRE platform, recommended on the occasion of the second AGATE workshop on 16th January 2017 "to define mandatory, recommended, optional metadata properties" and because "not all data providers will be able to deliver all metadata, make clear what is required as a minimum".[179] Jochen Schirrwagen also commented that the search function should be thought of as an expert presentation. In terms of creating these expert presentations, focusing on the user interface design and graphic design of the search results is of great importance. As examples, this means making sure key research information if prioritised, or with visual repositories slide show are generated. Several records of academies´ research from different repositories are already captured by OpenAIRE. Therefore, it would be a good starting point to make use of these existing OpenAIRE records via their API and apply the metadata guidelines of OpenAIRE when implementing the AGATE search function.[180] OpenAIRE and AGATE could make mutual use of the respectively aggregated metadata. A closer future cooperation between AGATE and OpenAIRE seems thus recommended. A letter of intent from OpenAIRE stating the general support for AGATE and an active OpenAIRE participation in both AGATE workshops can be seen as initial steps of this direction. At a later stage, when the AGATE has passed a beta phase and works well, an extension of these guidelines e.g. for a more customised metadata enrichment of enhanced publication (but still compatible with OpenAIRE) could be developed on this basis.

[176] See DINI 2010, p. 47.
[177] See Riley 2017.
[178] See 'OpenAIRE | Project Factsheets - General Information', accessed 14.05.2017.
[179] See Schirrwagen 2017.
[180] For more detailed information about these guidelines, see 'OpenAIRE | Guidelines', accessed 12.05.2017.

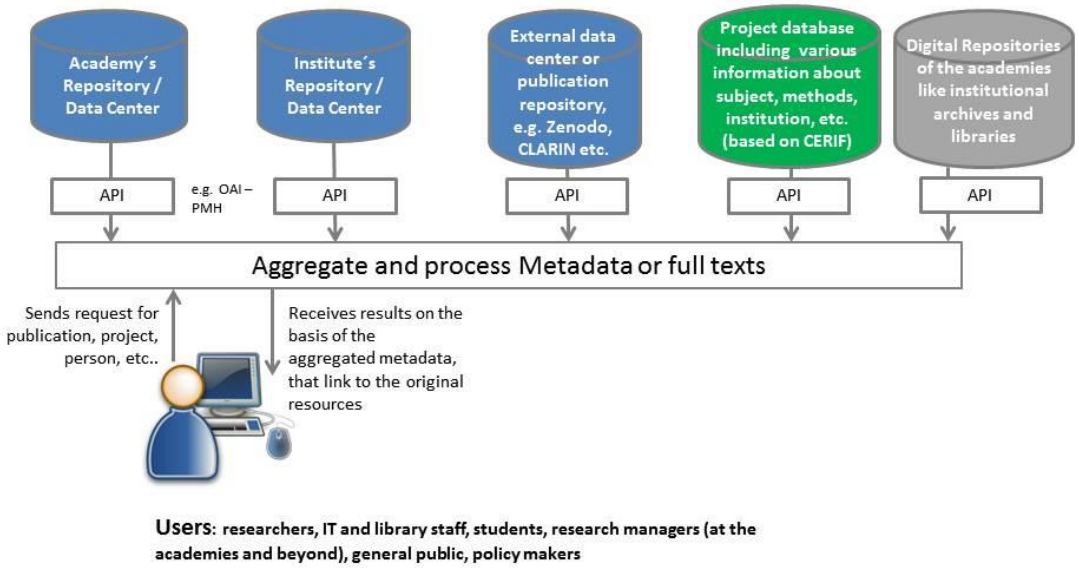

Figure 3: Graphic representation of the AGATE aggregation and search function (simplified)

Processing system

Given that functioning APIs at the relevant repositories are in place and a common metadata format is agreed on, a system is needed that processes the user query by accessing, selecting, filtering and linking (e.g. to the project information) the metadata of the digital resources to present search results.

In order to make use of the metadata accessed via an OAI-PMH interface, the metadata will be aggregated and therefore accessed and copied (harvested) in a metadata repository. In an initial query the metadata of a targeted repository is transferred into a database. This process would be necessary for each repository that is to be covered by the AGATE search function and provides an OAI-PMH data feed. This metadata repository would then regularly and continuously be updated; a process that affects only such metadata records that have been changed, added or deleted since the last updating process.[181] For this process the German National Library has

[181] See Hude and Rupp 2013.

developed an OAI-PMH Harvester Manager as open source software.[182] It should be further examined if this software meets the requirements especially regarding customisation and scalability.

Considering the European dimension of AGATE it seems advisable to build the metadata repository on commonly used and maintained repository systems, as e.g. DuraSpace as consortium of the software DSpace,[183] Fedora[184] and VIVO.[185] These three initiatives are working together and a vital community guarantees permanent development, updating and furthermore interoperability and interconnectivity between the systems.[186] Especially the fact that a DSpace CRIS solution has been developed by the open source community, being compliant with CERIF and combining research information with the digital resources of the research itself[187] is very promising. Linking research project information and publication and digital resources in one system would address a central requirement drawn from the use cases to connect digital resources with information about the projects in the database.

Jochen Schirrwagen recommended during the second AGATE workshop "an aggregation environment (supports but is not limited on OAI-PMH) with normalization features, Solr search software for indexing, and a lightweight portal solution to make the frontend website".[188] In order to be reusable for external users the AGATE database should itself provide an OAI-PMH API and SPARQL Endpoint,[189] an idea that was also highly welcomed by OpenAIRE.

As AGATE will import data that needs associations creating to academies, projects or publications, and a process is needed to carry out this task. By using the CERIF database fields schema, records that were previously unconnected, can be associated by running software processes called 'inference mechanisms', so creating new links. Inference mechanisms have already been implemented by the OpenAIRE system[190] and AGATE can benefit from this previous implementation.

Quality of metadata and search results

The possibility to access actual resources rather than the mere information of the metadata is a primary motive for most users to use the search function. Consequently,

[182] See 'DNB | OAI-PMH Harvester Manager', accessed 28.04.2017.
[183] See 'DuraSpace | Homepage', accessed 30.04.2017.
[184] See 'Fedora Repository | Homepage', accessed 12.05.2017.
[185] See 'VIVO | Homepage', accessed 14.05.2017.
[186] See 'DuraSpace | Technologies', accessed 14.05.2017.
[187] See 'DuraSpace | First Release of DSpace CRIS Module for DSpace 3', accessed 14.05.2017 and 'DuraSpace | DSpace-CRIS 5.5.0', accessed 14.05.2017.
[188] See Schirrwagen 2017.
[189] See 'W3C | Sparql Implementations', accessed 28.04.2017.
[190] See Kobos et al. 2014, p. 93.

the search result should pave the way to do so. According to the needs assessment drawn from the use cases the characteristics of this process should be “intuitive and easy to use”.

Given that metadata is harvested on a regular basis via OAI-PMH APIs from a number of academy related metadata repository, the user can start searching this metadata. The search result should then include a link back to the resource so that it can be accessed in its original environment and context. This could be facilitated via a Uniform Resource Identifier/Locator (URI/URL) that can be part of the Dublin Core set of properties. Since, as mentioned before, the application of Dublin Core elements is in many cases not mandatory, it should be considered to demand a URI as a mandatory element of the metadata as a condition to be harvested by AGATE.

This leads to an important issue, the quality of the search result will mainly depend on the quality of the metadata. To ensure a certain minimum quality data guidelines and minimum standards for the repository should be drawn up. OpenAIRE, but also Europeana[191] or the ARIADNE portal[192] have such guidelines and these should also be taken into account when developing the data model for AGATE.

Even with such guidelines in place, the search function will need continuous adjustments regarding organising, clustering, cleaning, and enriching the aggregated data in a collaborative process with the data provider and in response to the user behavior.[193] It is advisable to support this development in the initial phase with a small group of users, e.g. applying the *lead user* concept[194] introduced by Eric von Hippel from the Massachusetts Institute of Technology (MIT)[195] as done, for example, by the planners of the ARIADNE project, a successfully implemented metadata search engine for archaeological data.[196]

[191] Europeana Data Model (EDM), see: ’Europeana | Provide Data in EDM’, accessed 18.11.2016.
[192] See for example the ARIADNE Dataset Catalogue Model (ACDM) on ‘ARIADNE | Support Portal’, accessed 14.05.2017.
[193] In his aforementioned comment Jochen Schirrwagen also stressed to “not underestimate the effort for metadata curation, healing, normalization, even if guidelines are in place”. See Schirrwagen 2017.
[194] Selhofer and Geser 2015, p. 6.
[195] Hippel 1986, pp. 791-805.
[196] See Selhofer and Geser 2015.

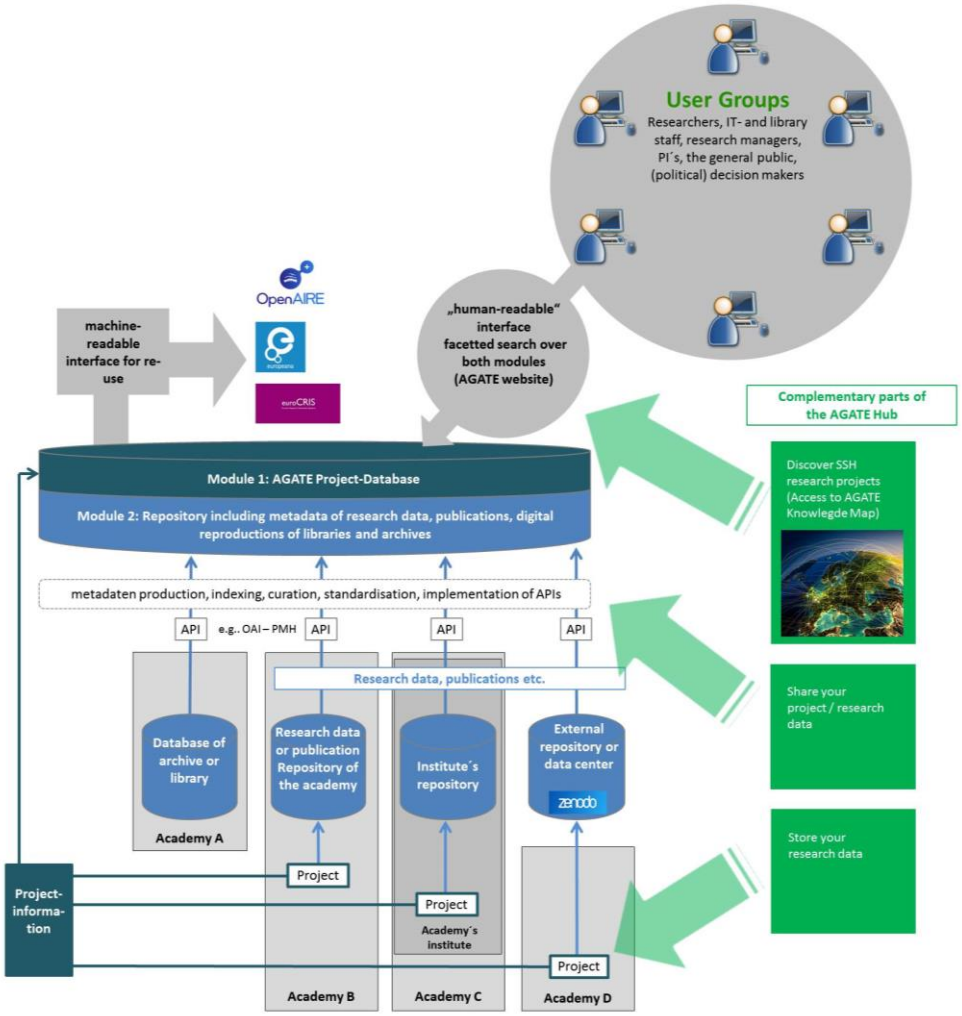

Figure 4: Graphic representation of general functionality of the Knowledge Map (simplified)

Recommendations

Although the actual technical design will be done in a process of piloting and in the agile design cycles, to this end, the following is recommended:

- To focus in the beginning exclusively on metadata harvesting.
- To set up guidelines for the metadata, preferably based on the already very advanced guidelines of OpenAIRE.
- To consider other guidelines of major metadata aggregators as well, especially for the further development once a beta stage is reached.

- To examine in how far the experiences of the Digital Knowledge Store at the Berlin-Brandenburg Academy are supportive for the inclusion and searchability of enhanced publications and “searches in the data”.
- To apply the process of agile design, e.g. using the lead user approach, for the development phase.
- To meet the challenges of a pan-European infrastructure, a metadata repository based on the DuraSpace group of CRIS software should be considered. Especially because of the desired linking of projects and research data, evaluate experiences from related initiatives and prospective partners.
- To start with harvesting the metadata (minimum standard Dublin Core) via OAI-PMH APIs. Several academy resources are already accessible (e.g. via OpenAIRE) and provide these interfaces.
- To evaluate the applicability of the OAI-PMH Harvester Manager software.
- To provide a helpdesk and other forms of guidance (FAQ) via the AGATE Hub to support the application and further installation of such APIs, work very closely with OpenAIRE and other relevant initiatives to exchange experiences and/or (re)use available resources, guidelines, helpdesk solutions, etc.
- To offer support services via the AGATE Hub that enable the academies, their institutes, and projects to create and curate metadata in a way that it can be effectively harvested and processed.

3 AGATE Hub – A Web Portal for Online & Offline Community Engagement

As can be seen from the previous chapter, the implementation, use, and expansion of the AGATE Knowledge Map will have to be supported by a broad spectrum of online and offline offers and activities. These resources and facilities will be provided and communicated via the AGATE Hub that will be developed parallel to the Knowledge Map. On the AGATE Hub users will not only find the central access point to AGATE's core tool, the Knowledge Map, they will also find essential information about AGATE's background and resources that enable them to contribute data to the Knowledge Map and thus become part of the AGATE network.

The specific appearance of the resources for this support function will largely depend on the concrete technical solutions chosen for the Knowledge Map and the resulting requirements for data providers. During the implementation phase, relevant existing guidelines (e.g. for standards) will have to be compiled and made available, and manuals and (video-)tutorials for potential data providers will have to be developed. Moreover, facilities for knowledge exchange and community building among the European Academies' SSH researchers that have been expressed as needs in the SASSH Survey and during interviews and the AGATE workshops, will have to be addressed.

The AGATE Hub also aims to attract a broad public beyond the European academies to win them as users of AGATE as an information tool and as strategic partners. Therefore, the website must be understood as central but not sole instrument of the project's dissemination and outreach activities. While these will be elaborated in chapter no. 4, the present chapter will focus on the analysis of the needs of the main user group of the AGATE Knowledge Map, the SSH researchers at the European academies. The analysis covers both the supportive aspects of the AGATE Knowledge Map and further knowledge exchange and community building aspects. The resulting requirements are matched with already existing solutions. It is examined how these could be integrated or if further development is required to meet the users' needs, before recommendations for the functionalities and basic structure of the AGATE Hub are given.

3.1 Use cases

1. Support of participation in AGATE Knowledge Map				
Use case scenario	**Source**	**Success scenario**	**Requirements for AGATE Hub**	**Benefit**
A researcher finds the AGATE Knowledge Map a meaningful research tool and wants to add his project information to the database	AGATE: Concept, chapter no. 2	Finds detailed instructions how to create a project entry. Finds (contact) information for help. Finds information about trainings.	Create information material for the use of the database (vocabulary, information about standards that have to be met, FAQ, etc.). Establish points of contact/a help desk. Offer training.	Academy projects participate in initial phase of AGATE Knowledge Map — project databasing
A researcher wants to publish his research data and results online, but has no experience in this field and is looking for information and support	SASSH, AGATE: Concept, chapter no. 2	Finds introductory information about e-publishing and Open Access. Finds recommendations where and how to store and publish data.	Promote e-publishing, Open Access and Open Data. Point to appropriate repository solutions, workshops and training resources Establish points of contact/a help desk. Offer trainings.	More projects publish their research data online that can be integrated into the second module of AGATE Knowledge Map — digital resources databasing

		Finds online tutorials as well as information about workshops and training.		
A researcher/project coordinator wants to improve the international visibility and connectivity of his research	SASSH, AGATE: Concept chapter no. 2	Finds a directory of relevant European infrastructures like Europeana, DARIAH, CLARIN, OpenAIRE, etc., with basic information about their objectives and requirements for cooperation. Becomes data provider for/cooperates with relevant European infrastructures.	Directory with Information about Europeana, DARIAH, CLARIN, OpenAIRE and their services and possibilities for cooperation (Open Archive HAL, Zenodo, etc.). As teaser: Presentation of projects that already cooperate with infrastructures/reports/interviews with researchers from such projects.	More academy projects share data via European infrastructures (that can be integrated into the second module of AGATE Knowledge Map — digital resources databasing)
2. Support of knowledge exchange and community building				
A PI plans a digital edition — the ICT staff	Interviews	ICT expert finds academy projects of interest via	Smart linking between AGATE Knowledge	Academy projects reuse existing digital research tools, adopt tools , or develop

involved in the application process/the ICT staff advising the project during implementation want to get an overview of existing solutions and to check, if and how they can reuse them		AGATE Knowledge Map and can inform herself about current activities of this project (workshops, conferences, and publications). ICT expert finds an overview of existing registries of digital research tools. ICT experts give feedback to tool makers and exchanges ideas about further development.	Map and other website categories. Point to existing registries of digital research tools. Provide a forum where tools developed by or relevant for the academies can be presented and discussed if, or how a tool can be reused, or adopted	tools in partnership between academies
A project has developed a new tool and wants to present it to the community	Interviews	ICT experts promote their tools developed in or for academies projects. Can ask for and get feedback.	Provide a forum where ICT experts can present tools and give feedback	Academies' researchers benefit from knowledge exchange and experiences made by peers

A researcher/ICT expert interested in virtual reconstructions is looking for activities on virtual reconstructions at academies throughout Europe (workshops, seminars, and conferences, etc.)	Interviews	Finds academy projects, publications, and activities that deal with virtual reconstructions. This is partly via initial phase of AGATE Knowledge Map - project databasing.	Provide news about current academies' activities (e.g. workshops, conferences, also actively provided by the community members themselves). Smart linking with Knowledge Map and other website categories.	Researchers gain knowledge about digital developments and activities at other European academies and make contact
A PI writes a proposal for a new project and is interested in experiences of similar projects (e.g. regarding estimation of efforts, use of a particular method, adoption of a specific tool)	Interviews	Finds projects with similar topics/methods and can contact them (via Knowledge Map — project database). Finds reports and "lessons learned". Comments on reports and "lessons learned". Starts a discussion/opens a group about a particular issue.	Provide space where reports can be deposited and published/announced. Offer comment function. Offer communication platforms like blogs, discussion groups, mailing lists.	Projects benefit from work experience made by others. European Academies bundle their resources.

A researcher wants to share his experiences ("lessons learned") with reuse/adoption of a digital tool/method	Interviews	Uploads reports and "lessons learned". Finds description of tool/method he has used and leaves a comment.	Provide space to store, publish and prominently announce reports and "lessons learned". Offer comment function. Offer communication platforms like blogs, discussion groups, mailing lists.	Projects benefit from work experience made by others. European Academies bundle their resources.
A researcher wants to communicate and collaborate with his colleagues abroad (virtual meetings, exchange of research data that has not yet been published, collaborate on documents in real time)	Interviews	Finds information on web-based applications that facilitate the daily work-routine like cloud services, video conferences or virtual research environments	Offer information about tools for collaborative work. Liaise with the providers of relevant services and eventually adopt and provide individual solutions (e.g. cloud services, virtual workspace), being aware of limited life cycles of workspace tools.	Projects with partners in different locations work together more efficiently
A researcher wants to document his project and share project-specific knowledge and	Interviews, Desk research, Scientific	Can apply for an account for a restricted area and create a user group	Offer restricted area and project management tools	Projects document and share their work experience more efficiently and safely

information with his team colleagues that work in other cities/countries internationally (e.g. workflows, project specific solutions, milestones)	Advisory Board			

Table 4: Use cases for AGATE Hub

Analysis of the use cases

The table above suggests a number of features that the AGATE Hub should provide. Ideally the AGATE Hub setup should meet the following four main requirements:

1. Provision of information and support for participation in AGATE Knowledge Map

It will be a primary task of the Hub to inform academies' staff and project coordinators about the objectives of AGATE and the features of the Knowledge Map and to provide support for creating a new project entry in the database and for preparing data and repositories in a way, so that their metadata can be harvested by the metadata repository of the Knowledge Map and be compatible with other aggregators.

To meet this requirement both online and offline activities are needed. Besides online resources like guidelines, tutorials and services such as contact forms and ticket systems, also personal contact points should be established. This personal contact can support researchers in the role as a data provider with individual guidance regarding the participation in AGATE, and to connect these researchers or ICT experts with providers of relevant third-party services.

The importance of personal contact for the project success — either via a staffed help desk, hands-on-workshops or onsite events with partners — was underlined by Mareike König, external peer reviewer of the Community presentation at the second AGATE workshop.[197] Her comments were based on her own experiences with the establishment of the German section of Hypotheses, a blog portal for the Humanities and Social Sciences.[198]

2. Provision of services that facilitate communication and strengthen the community

The second important requirement for the AGATE Hub arises from the user need to connect, communicate, and collaborate more easily, efficiently, and safely with colleagues all over Europe. Therefore, the AGATE Hub has to incorporate features that strengthen existing networks and facilitate further networking and knowledge exchange between the researchers at European academies and beyond. In this context services and activities of high relevance are:

- publishing and pointing to new tools, research results, and reports ("lessons learned")

[197] See Ott 2017.
[198] See Adrian, Ott and Wuttke 2017 and 'de.hypotheses | Homepage', accessed 18.05.2017.

- information about relevant events, workshops, and conferences
- exchange about ongoing research activities, new developments, and initiatives (also in a “safe (internal)” environment)
- services that facilitate collaborative work

3. Provision of offline/face-to-face activities and services

Moreover, the use cases show that the online offers have to be flanked by a variety of offline activities and services that will be tailored according to the users’ requirements and promoted via the AGATE Hub and AGATE team, such as; workshops, trainings, or roundtables.

4. Provision of information about user specific aspects of digital research and publication practices

Finally, the Hub will serve as a platform for the transfer of knowledge in the fields of digital research and publication practices that goes beyond the participation in the AGATE Knowledge Map and has a special focus on topics relevant to the SSH research activities at the European academies such as:

- open access publishing and FAIR research data management
- data sharing and cooperation with European infrastructures such as Europeana, DARIAH, CLARIN, OpenAIRE
- policies and guidelines
- discipline specific methods and formats, workflows and best practices
- digital tools and resources, especially developed by European academies, including work in progress
- state-of-the-art science communication
- training material and events, e.g. online courses, videos, workshops
- DH specific events, especially organised by and at the academies — conferences, workshops, trainings

The requirements ask for both static informational material such as informative articles, guidelines, registries or videos and rolling content such as regularly updated news about relevant online and offline-activities. In a next step, the requirements above will have to be matched with the broad range of already existing information platforms and services provided by different institutions and initiatives. It will need to be considered to what extent they already cover the needs of the main AGATE user group and where there are potential gaps.

3.2 Relevant platforms, initiatives, and services

In the last decade, a large number of research institutions have established information and knowledge facilities to deal with the digital turn in academia. A large variety of digital research infrastructures — often supported by national or European funding schemes — have been developed by and with user communities, that partly correspond with the proposed users of AGATE. This means that on the one hand AGATE can benefit from a wealth of experience gained in similar research infrastructures, already existing materials, and services that can be reused or adopted.[199] On the other hand, it has to be analyzed critically if the existing offers fit the needs of AGATE's key user groups and alternative solutions need to be considered if this proves to be the case.

Below is a comprehensive listing of scholarly information portals, either made as support services or for deeper insights into evaluating digital trends. The list has been broken down into a series of grouping based on editorial approaches we identified in the research.

1. Digital Humanities sections of individual academies or institutions with a long experience in collaborating with the academies on national level, such as: the Austrian Center for Digital Humanities at the Austrian Academy of Sciences (http://www.oeaw.ac.at/acdh/); the "Digitale Akademie" of the Academy of Sciences and Literature — Mainz (http://www.digitale-akademie.de/); TELOTA (The Electronic Life of the Academy) at the Berlin-Brandenburg Academy of Sciences and Humanities (http://www.bbaw.de/telota/telota); the Cologne Center for eHumanities of the University of Cologne (http://www.cceh.uni-koeln.de/) that supports the North Rhine-Westphalian Academy for Sciences and Arts; or the Trier Center for Digital Humanities (http://kompetenzzentrum.uni-trier.de/de/) that works together with many projects of the German Academies' Programme.

2. Pan-European digital infrastructures or projects with focus on the Social Sciences and Humanities, such as: DARIAH (Digital Infrastructure for the Arts and Humanities, http://www.dariah.eu/); CLARIN (Common Language Resources and Technology Infrastructure, https://www.clarin.eu/); Europeana (https://www.europeana.eu/portal/de); PARTHENOS (Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies, http://www.parthenos-project.eu/); Humanities and Social Sciences online (https://networks.h-net.org/); or DASISH (Data Service Infrastructure for the Social Sciences and Humanities, http://dasish.eu/).

[199] The need for coordination and sharing of experience to prevent unnecessary duplication is a fundamental point stressed in ESF 2011, p. 2.

3. Digital infrastructures or institutions/projects that offer services regardless of the scientific discipline to facilitate digital research and publication practices in general, such as: EUDAT (https://www.eudat.eu/); Research Data Alliance (https://rd-alliance.org/); OpenAIRE (https://www.openaire.eu/); FOSTER (Facilitate open science training for European research, https://www.fosteropenscience.eu/); or the Göttingen eResearch Alliance (http://www.eresearch.uni-goettingen.de/).

4. Discipline or topic-specific infrastructures and networks in the field of the Humanities and Social Sciences, such as: the European Holocaust Research Infrastructure EHRI (http://www.ehri-project.eu/); the Cracking the Language Barrier Initiative (http://www.cracking-the-language-barrier.eu); or the Collective European Digital Archive Infrastructure for medieval studies and World War I — CENDARI (http://www.cendari.eu/).

5. Supra-national or interdisciplinary professional associations, such as: the Network for Digital Methods in the Arts and Humanities (http://www.nedimah.eu/); European Association for Digital Humanities (http://eadh.org/); EASSH (European Alliance for Social Sciences and Humanities (http://www.eassh.eu); or the Alliance of Digital Humanities Organizations ADHO (http://adho.org/).

6. National or language based DH networks, such as: the national Netherlands branches of CLARIN and DARIAH for example, that joined to form CLARIAH (http://www.clariah.nl/); or national and linguistic branches of the ADHO — like the German association Digital Humanities im deutschsprachigen Raum/DHd [Digital Humanities in the German language territories] (https://dig-hum.de/); or working groups of national professional associations like the AG Digitale Romanistik of the German association of Romanists (http://www.deutscher-romanistenverband.de/der-drv/agdr/).

7. Research Data Centers with specialization in Data from SSH research, such as: IANUS — Forschungsdatenzentrum Archäologie & Altertumswissenschaften (http://www.ianus-fdz.de/); or GESIS — Leibnizinstitute for the Social Sciences (http://www.gesis.org/en/).

8. Digital Registries and Directories of research-relevant services such as: Re3data.org Registry of Research Data Repositories (http://www.re3data.org/); or the DiRT directory of digital research tools (http://dirtdirectory.org/).

9. Specialised Information Services for Researchers located at research libraries and in Germany recently supported by the German Research Foundation (DFG)[200] such as the Fachinformationsdienst Allgemeine und Vergleichende Literaturwissenschaft (http://avldigital.de/). On these platforms libraries develop tailor-made services for

[200] See 'DFG | Press Release no. 63, 2015', accessed 08.05.2017

individual disciplines by making research-relevant materials and resources available that extend beyond basic library resources with emphasis on digital media.

In summary, these potential cooperation partners have a broad spectrum of disciplinary and institutional backgrounds. One added value of the AGATE Hub would be the identification of relevant platforms, initiatives and services from a heterogeneous and dispersed landscape.

3.3 Services that facilitate communication and strengthen the community – mapping existing solutions

The analysis of the use cases suggests that the second pillar of the AGATE Hub should consist of services that facilitate the communication and strengthen the AGATE community (see chapter no. 3). In the following table 5, potential services are presented and examined regarding their desired effects. Then existing solutions are considered and evaluated before the resulting 'Next steps' tasks for AGATE are described.

Service	**Effects**	**Existing solutions**	**Next steps for AGATE**
Possibility to publish and point to new proposals, digital tools and research results	Facilitates knowledge exchange. Allows bundling of resources. Raises visibility of digital research activities at the academies. Facilitates international cooperations.	Hypotheses.org (publication platform for academic blogs). DHd-Blog (only German) http://dhd-blog.org/. None with focus on European academies.	Talk with partners that can provide AGATE with respective infrastructure. Define structure of the blog. Call for contributions, contributions from among the academies' researchers.
Possibility to discuss current academy relevant activities and developments	Increases information and knowledge exchange and pooling of resources.	None	Examine where and how this feature is realized best (news area, blog, mailing list). Develop respective technical environment.

Possibility to inform others and get informed about academy relevant events (workshops, conferences)	Speeds up the flow of information among the members of special interest groups	Some websites of individual academies provide information about events, but there is no central point with information about relevant activities at all European academies	Examine where and how this information is best placed (news area, blog, mailing list). Establish respective technical environment.
Possibility for research teams to document their work and share their knowledge (internally)	Facilitates collaborative work in existing project teams. Facilitates international cooperations.	DiRT Directory, https://dirtdirectory.org/tadirah/collaboration. Cloud Service "B2 Drop" by EUDAT, https://eudat.eu/services/b2drop. Virtual research environments like workspaz by Max Planck Society (https://workspaz.mpdl.mpg.de/) or CENDARI Note Taking Environment, https://docs.cendari.dariah.eu/user/writing_arg.html.	Promote existing services that facilitate collaborative work. In parallel: participate in the development of new solutions relevant for the European academies (cloud services, virtual research environments), where a mandate by the community is given.

Table 5: Services that facilitate communication and strengthen the community – mapping existing solutions

The analysis confirmed the picture that a large number of desired community effects regarding, better communication between the researchers of the European academies and community building, cannot be achieved simply by the provision of static information on a website. Instead it requires more flexible and responsive solutions like the use of blogs and social media channels. This requirement was underlined by the members of the Scientific Advisory Board and the participants of the second AGATE workshop. Again, it was suggested to start with bundling and presenting existing blogs and social media activities of participating projects on the AGATE website before developing newly commissioned material. In most cases it can be learned from experiences already made by other research institutions like for example the Max Planck Society or Max Weber Foundation.

3.4 Information and training resources: re-use and adoption of existing offers and development of new solutions

As the use cases suggested, the AGATE Hub should also provide information and training resources about central topics of *digital SSH research practices* relevant for the European Academies researchers in general, such as:

- open access publishing and FAIR[201] research data management
- data sharing and cooperation with European infrastructures such as: Europeana, DARIAH, CLARIN, OpenAIRE
- policies and guidelines
- discipline specific methods and formats, workflows and best practices
- digital tools and applications, especially developed by European academies, including work in progress
- state-of-the-art scholarly communication
- training resources and events, e.g. online courses, videos, workshops
- DH-specific events, especially organised by and at the academies — conferences, workshops, trainings

For most of the topics, there are already quite a number of initiatives that provide resources suitable for the AGATE community. With regards to Open Access, AGATE could for example refer to FOSTER that provides respective publications and online-trainings and announces events.[202] As for research data management, AGATE could point to existing tools, such as: the data management plans provided by the Digital Curation Center,[203] to DMPTool of the University of California Curation Center for strategic plan sharing,[204] to the *DMP Wizard* provided by CLARIN-D,[205] or to the software developed by the project *Research Data Management Organiser*, located at the University of Applied Sciences Potsdam and Leibniz-Institut für Astrophysik Potsdam.[206] Besides FOSTER, also PARTHENOS, DARIAH, and CLARIN are offering relevant services regarding DH training. PARTHENOS provides training modules and resources in digital humanities and research infrastructures,[207] such as a *standardization survival kit*.[208] The #dariahTeach learning platform is currently developing "open source, community-driven, source, high quality, multilingual

[201] See 'FORCE11 | The FAIR Data Principles', accessed 13.05.2017.
[202] See 'FOSTER | Homepage', accessed 08.05.2017.
[203] See 'DCC | Data Management Plans', accessed 08.05.2017.
[204] See 'UC3 | DMPTool', accessed 12.05.2017.
[205] See 'CLARIN-D | DMP Wizard', accessed 12.05.2017.
[206] See 'AIP and FHP | RDMO (Research Data Management Organiser)', accessed 08.05.2017.
[207] See 'PARTHENOS |Training modules and resources in digital humanities and research infrastructures', accessed 08.05.2017.
[208] For a draft report of the planned online environment see Laurent et al. 2016.

teaching materials for the digital arts and humanities".[209] The collaborative portal TeLeMaCo provided by CLARIN-D offers all kind of online training and teaching materials relevant to linguistics and digital humanities.[210] Lately, CLARIN and DARIAH have relaunched a course registry of face-to-face events in Digital Humanities teaching all over Europe.[211]

The diversity of topics relevant for the AGATE community and multitude of high-quality resources offered by different research infrastructures underline the demand for a first point of contact for a quick overview of the available services. The task for AGATE would therefore be to collect and classify existing resources and present them in a user-friendly way. Nevertheless, existing offers could be accompanied by newly developed resources to better reach the academies' researchers and at the same time enhance the SSH community as a whole.

In regards to digital tools and software applications relevant for the AGATE community, this could mean that the AGATE Hub would not only promote existing software registries, such as the DiRT directory[212] or the DARIAH tools registry,[213] but additionally provide incentives to present digital tools and applications from the academies, as they are only occasionally present in existing registries or on the websites of individual academies.[214] Moreover, the Hub could offer — e.g. via a blog — space to discuss tools, as up until now, there is no such respective service. Also, a low entry level directory of European research infrastructures relevant to SSH long-term research would be very useful for the AGATE community. Such a showcasing could ideally illustrate potentials for collaboration by success stories such as personal profiles and interviews with representatives of academies that already contribute data or cooperate with RIs. For the matter of policies and statements, AGATE could help to make relevant publications (policies, guidelines, statements) by the academies visible and findable. Currently, numerous members and staff of European academies take part in working groups and initiatives that deal with fundamental aspects of digital SSH research activities and publication practices, but the publications cannot be accessed in a central repository.[215]

---

[209] See 'DARIAH | #dariahTeach learning platform', accessed 08.05.2017.
[210] 'CLARIN-D | TeLeMaCo, The Linguistic Teaching Resources Hub', accessed 08.05.2017.
[211] 'CLARIN and DARIAH-EU | DH Course Registry', accessed 08.05.2017.
[212] 'DiRT Directory | Homepage', accessed 08.05.2017.
[213] 'DARIAH | Dienste und Werkzeuge', accessed 08.05.2017.
[214] See for example the Webpage of the Austrian Center for Digital Humanities: 'ACDH-ÖAW | Resources & Tools ', accessed 08.05.2017.
[215] E.g. the report of the working group E-Humanities of ALLEA *Going digital* (Harrower et al. 2015) the white paper by the KNAW eHumanities group *Meaning and Perspectives in the Digital Humanities. A White Paper for the establishment of a Center for Humanities and Technology* (CHAT), the publications by the ICT-staff of the ACDH (see 'ACDH-ÖAW | Publications', accessed 12.05.2017), or the statements, hand-outs and appeals on the core topics: Open Access, Research Data, Virtual Research Environments, National Licensing, worked out by the Priority Initiative "Digital Information", a joint initiative of the Alliance of Science Organisations in Germany, in which the German Leopoldina currently participates

3.5 Recommendation for the implementation of the AGATE Hub

The general structure of the AGATE Hub should be configured to meet the user's needs and their daily work routine.[216] The five main categories that are derived from these work routines are: "Discover research projects", "Start a new project", "Share your project/research data", "Publish/Store your research data", and "Get connected". Further components might be added to AGATE once the basic components have reached a certain maturity, when AGATE has proven its acceptance by the community (sustainability), or if there is a concrete need formulated by the community. However, right from the beginning, AGATE should provide not only static resources but also innovative services that support community building.

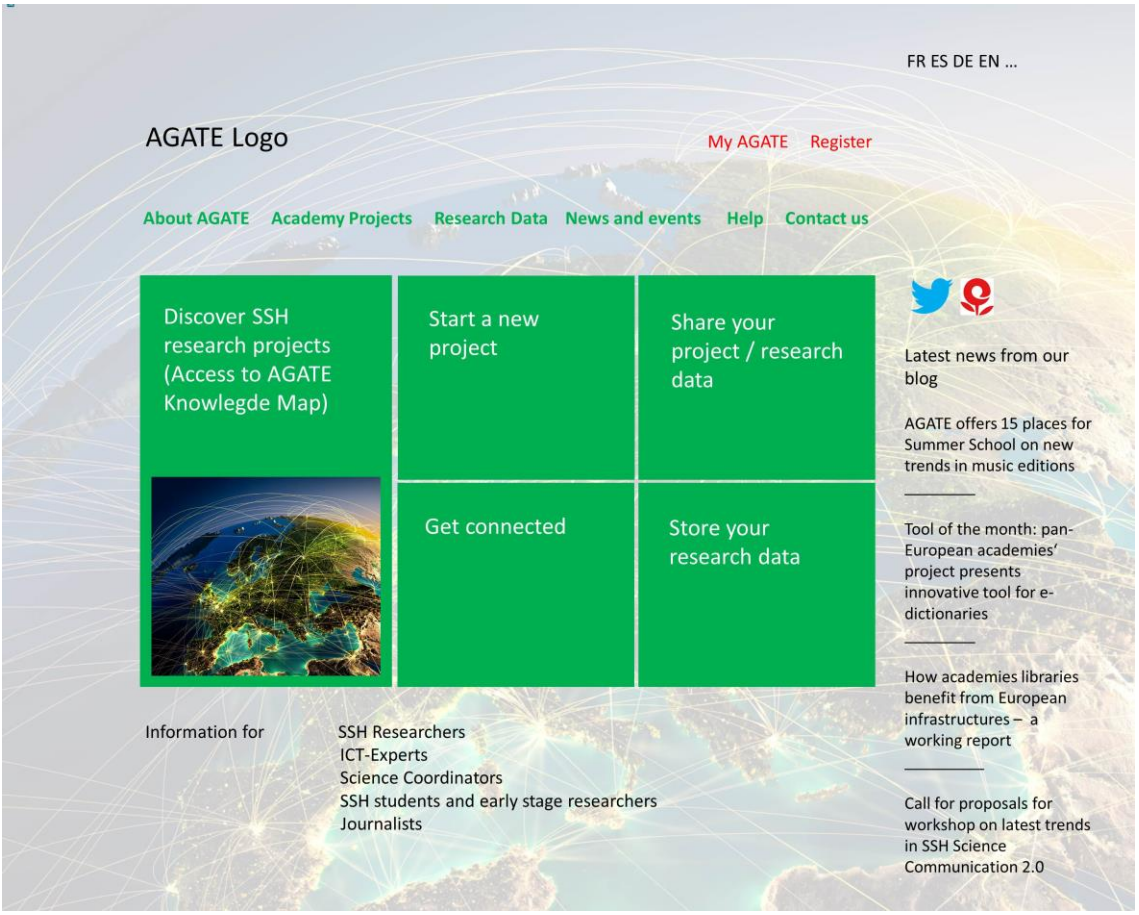

Figure 5: Draft of AGATE start page

(see 'Allianz der deutschen Wissenschaftsorganisationen | Priority Initiative "Digital Information"', accessed 13.05.2017).

[216] This approach was proposed by Mareike König in her expert comment at the second AGATE workshop on January 16th 2017 and was strongly supported by the workshop participants during the subsequent discussion.

Discover SSH research projects

This section will offer the main access point to the AGATE Knowledge Map and provide guidelines for the use of the Knowledge Map and visualisation tools like a map with projects of the same disciplinary affiliation or the use of the same digital methods. These incentives facilitate the use of the Knowledge Map for various user groups with different backgrounds.

Start a new research project

Here the users will find an overview of information essential when starting a new research project. The identified topics range from research infrastructures and their services, to research data management tools, standards, and DH tools as well as to Citizen Science. In most cases, AGATE will point to existing resources offered by their strategic partners, but accompany these with additional, newly developed formats, like concrete case studies from the academies' context.

Share your project or research data via AGATE

Here will be the place for instructions and guidelines how to contribute to the Knowledge Map as a data provider, in the first phase for the 'academy projects' database, later also for 'research resource' data. The resources will include online material like manuals and videos but also information about hands-on-workshops and about the AGATE Helpdesk. Ideally, this material is provided not only in English but in several languages to better address the targeted users.

Store digital project results

This section will provide resources about repositories, Open Access publishing and Open Data. Furthermore, it will point to the services of existing research infrastructures like the CLARIN data centers, DARIAH, or OpenAIRE. Like in section "Start a new research project", additional user specific resources will illustrate the benefits of publishing Open Access and sharing data.

Get connected

This section opens the doors for an individual and deepened exchange of information and experience among the community of European academies´ SSH researchers via a

blog and social media activities of contributing projects, participating academies and cooperation partners. It will be organized according to bottom-up principles. Researchers will be enabled and encouraged to present and discuss relevant topics and share news with the community. This might include expert articles on recent developments in the Digital Humanities, interviews with SSH researchers and ICT experts about their latest projects, presentations of new digital research tools or announcements of events and publications. The provided services facilitate the communication and speed up information flows within the community.

Further essential resources and features

Moreover, the AGATE Hub holds essential information about the AGATE project like mission, organisation structure, partners and team as well as a "library" with project publications such as workshop reports and press releases, e.g. in form of a Zenodo group. Also, a news area, where events and trainings organized by AGATE can be announced, and a help desk would be included. Regarding further requirements for the website, for example to be in English and preferably have a multilingual user interface is recommended. Such language features will help reach high acceptance among researchers and the broader audience from across Europe. Also a registration section enables users to create accounts to join the community. A login area for community members allows them to add a new project to the database or share research data. As the website also wants to promote the Knowledge Map to further user groups at least some of these groups should be addressed explicitly (see chapter no. 4).

4 Outreach, Dissemination, and User Involvement

The dissemination activities will have to concentrate on two main targets: first, the promotion of the AGATE services among the key user group, the SSH researchers at the European academies, to gain a high level of acceptance and active participation, and second, the promotion of AGATE among a broader audience to win them as users and audience. The following chapter outlines the proposed dissemination activities according to the project phases, and includes traditional communication channels such as face-to-face events, project publications and presentations as well as online activities based around the AGATE Hub and through social media channels. The suggested plan is essentially based on the project's core objectives formulated in the previous chapter, and the requirements arising from them, but it also considers reports on experiences from similar projects, and advise expressed by the Scientific Advisory Board and experts attending the two AGATE workshops.

The following table outlines the different dissemination objectives and activities regarding to the project phases:

Project phase	Dissemination objectives	Dissemination activities
Implementation of AGATE services (Knowledge Map and Hub)	Support of the core objectives of the project, win test group for development of the AGATE Knowledge Map and AGATE Hub, consolidate and expand network of partner academies and infrastructures	Promote AGATE project among academy presidents and researchers; win at least one AGATE ambassador at each participating academy (personalised letters, flyer/brochure, meetings). Present AGATE project to a test group of SSH researchers from inside and outside of the academies, involve them with the development of a prototype of the Knowledge Map and with testing/improving the AGATE Hub (workshops, virtual meetings). Develop information resources about the project on the Hub (Mission, organisation, project partners); authoring of guidelines, manuals (why contribute, how contribute, general technical requirements, vocabulary).

Initial offering, evaluation and refinement	Community building and involvement; attract and enable SSH researchers to contribute to the Knowledge Map and participate in the communication services (blog, comment function)	Present beta-version of Knowledge Map and AGATE Hub, advertise the project with face-to face-events to a wider choice of SSH researchers at the European academies; Inform and train AGATE ambassadors (workshops, newsletters). Establish social media presence (e.g. blog, Twitter and Facebook). Use communication channels of project partners (ALLEA, cooperating RIs, etc.).
Establishment phase	Achieve general visibility, extensive use of project deliverables	Full launch of AGATE Hub. Presentation of AGATE at relevant conferences and events at the European academies, partner infrastructures and broader SSH community in Europe, traveling exhibition. Liaison activities with partner infrastructures, academies, existing community networks.
Future use	Extensive use of AGATE Knowledge Map and Hub; provision for sustainability	Refine and expand features of AGATE Knowledge Map and Hub, review and potentially improve participation strategy

Table 6: Dissemination objectives and activities regarding to the project phases

4.1 Recommendation – strategic approach

As in the project's initial phase the focus of the outreach activities lies in the support of the project's main objective, the implementation of the Knowledge Map and the Hub, first, the dissemination activities will concentrate on the academies' SSH researchers and related communities. Since participation requires a substantial contribution in terms of time and workload, it appears most feasible to first invite a selection of researchers from projects from the core consortium, as they already have demonstrated their willingness to cooperate. The group ideally is broadly diversified both geographically and thematically, as the first project entries will demonstrate the broad spectrum of academies' SSH research activities and encourage — once opened to further circles — other projects to create entries. This approach may ideally lead to

a process of “inverted competition”, as was recommended by the Scientific Advisory Board. Moreover, the group should be actively involved in the selection of essential information resources and testing of the AGATE guidelines and manuals. After the implementation of a beta-version of the Knowledge Map and Hub, the dissemination activities will concentrate on promotion of testing and contributing to the services among the community of SSH researchers. Once a significant number of projects have created project entries and the development of the functionalities of the database on the whole is completed, the AGATE Knowledge should go online. At this stage, dissemination strategies concentrate on expanding the number of contributors to the Knowledge Map and reaching a broader audience. Of utmost importance is the cooperation and involvement of the participating academies as well as of partner institutions, infrastructures and community networks.

The target groups can be further segmented and addressed as follows:

Stakeholder group	**Further segmentation**	**Dissemination objectives**
European academies	Academy presidents and academic directors	Win as strategic partners and supporters of AGATE project
	SSH researchers	Win as contributors for AGATE Knowledge Map; motivate to receive training and encourage to share their knowledge with their colleagues
	AGATE “ambassadors”	Win as contact persons on-site; train and support them to act as a multiplier in their academy and intermediary between the central AGATE coordination office and the management level of individual academies (e.g. for organisation of events on-site)
	Research coordinators	Win as multipliers, audience and contributors of AGATE Knowledge Map and Hub

	Staff for press and relations	Win as multipliers, audience and contributors of AGATE Knowledge Map and Hub (announcement of relevant events, publications and activities)
Research Infrastructures and partner organisations	function level	Win as multipliers for AGATE project communication to their communities (on their websites and blogs, mailing campaigns, newsletters). Motivate to promote their activities, services, and resources for the specific user group of SSH researchers at European academies.
SSH researchers without academy affiliation		Win as users of research data and communication/cooperation partners
	(PhD) students and early stage researchers	Win as users of research data and communication/cooperation partners
Funding bodies and policy makers	National and European level	Win as users of AGATE Knowledge Map, enter strategic partnerships, round table discussions about academies' SSH research needs and value
Citizen scientists		Win as users of AGATE Knowledge Map
Media and public		Win as audience and multipliers

Table 7: Stakeholder groups and dissemination activities

4.2 Challenges

There are various challenges for successful community building and user involvement. First of all, the success for the Knowledge Map and the Hub stands and falls with the contents' selection and presentation as well as with its usability. This includes aspects like the technical user friendliness, a clear structure and an appealing visual design.

Already by being scholars, the main user group can be characterized as demanding, independent and critical. Therefore, they will only use AGATE regularly and collaborate actively, if they indeed consider it a convincing and beneficial service. In order to ensure a high level of acceptance and identification, users should have the opportunity to shape the infrastructure in the implementation phase, e.g. by agile software development user consultation methods. Another great challenge lies in AGATE's objective to win users with heterogeneous backgrounds and perspectives as its audience. In order to be interesting and valuable for all of them, diverse entry options and incentives to use and contribute to the different services should be considered right from the beginning.[217] Furthermore, presentations and guidelines focused on the individual needs of the different user groups should support this objective.

Due to the complex and diverse hierarchical structures of science academies in Europe, various communication channels have to be considered for the user involvement. Experience of the SASSH project and the current AGATE project has taught, that especially in the beginning of a new initiative great efforts are needed to get the necessary feedback and collaboration on a broad scale. In our case, the involvement should not be that difficult, as the AGATE project has become well known among the European academies and many of them have expressed their support informally or with letters of intent.[218] Nevertheless, it seems advisable to build up a network of AGATE ambassadors, contact persons at each participating academy that promote the project's services on-site and act as an intermediary between the central coordination office and the management level of the individual academy.[219]

[217] The importance of this need is also reflected in the fact that within the H2020 proposal it was proposed one work package that deals with the engagement and empowerment of the actors in the research infrastructure with Open Innovation methods and practices. For more information about the H2020 proposal see chapter 1.4.

[218] See Appendix 1: 'Partners and supporters of AGATE 1.0'.

[219] See for example the activities of the *Open Access Ambassadors*, who educate and foster Open Access among early career researchers in the Max Planck Society, see 'MPG | Open Access Ambassadors', accessed 13.05.2017.

5 Legal Issues

5.1 General

During the process of forming a consortium for the development of a Pan-European SSH research infrastructure like AGATE and its actual implementation, legal questions in several areas need to be addressed. The following chapter sketches the main prospective challenges and provides preliminary recommendations.

Comprehensive legal recommendations need to be developed by legal specialists in cooperation with the partners of the consortium according to the concrete situation, because they will depend highly on the organisational and financial structure of the consortium, the nature of envisaged tools and services, and on obligations to research funders. The following recommendations are formulated especially with an eye on the H2020 framework as a good practice for European infrastructure projects.

5.2 Legal issues

Profound legal issues arise directly from the objective to develop a transnational research infrastructure with at its heart a knowledge map of the academies SSH research in the form of a database that provides detailed project information and aims at a later stage at metadata harvesting of the academies' digital SSH resources.

They concern especially Intellectual Property Rights (IPR), Privacy Rights, and Copyright. Their complex interplay is caused by differences between national and European laws as been analysed for example in a report provided by the DASISH project.[220] That these laws are constantly changing adds to the difficulty of developing a fitting legal framework.

1. Project internal: IPR, access rights, data management & privacy

Guidelines for H2020 projects are provided by the H2020 Model Grant Agreement[221] and the Guidelines for FAIR Data Management in Horizon 2020.[222] Advice can be obtained from the IPR Helpdesk[223] and national H2020 helpdesks.[224] One basic requirement for an AGATE mailing list, for example for the distribution of a newsletter, will be developing a privacy policy. On EU level the relevant legal framework is the Data Protection Directive (95/46/EC) and the ePrivacy directive (2002/58/EC, as

[220] See Schmidutz et al. 2013.
[221] See European Commission 2017.
[222] See European Commission 2016.
[223] See 'European IPR Helpdesk | Homepage', accessed 13.05.2017.
[224] For Germany see 'EU-Büro des BMBF | Homepage', accessed 13.05.2017.

revised by 2009/136/EC). The privacy policy needs also to be compliant with national laws.

2. Project external: Dissemination and exploitation of results

As AGATE will promote Open Science, it is only natural then, that for its own project results it will adhere to the H2020 principles of Open Access and (Linked) Open Data.[225] This should entail using the Creative Commons license framework (currently CC 4.0), standard open access repositories, and to build on experiences made by Pan-European SSH-infrastructures.[226] Wherever possible, software created during the project should be made available to the community on an open source basis.

The metadata that can be harvested from the AGATE Knowledge Map should be attributed with an open licence (if not in conflict with privacy issues), such as CC 0 (Public Domain) to facilitate its reuse by other aggregators, for research purposes and the by the public.[227]

One of the major challenges for the AGATE Knowledge Map will be the different levels of access to the underlying resources and different rules of how to use and reuse these materials. Though one of the objectives of the AGATE project is to promote Open Access to the academies digital SSH resources in general, the need may also arise to integrate resources with restricted access, as for example in the social sciences restricted access to sensitive data is common. A cooperation between AGATE and DARIAH, the Digital Research Infrastructure for the Arts and Humanities, that is developing a data reuse charter and a DARIAH Seal of Approval which aims to provide a framework that could serve as a general baseline for interactions between researchers, cultural heritage institutions, and hosting services in this respect is envisaged.[228]

Recommendation

The concrete legal framework for the future AGATE data will be fully developed during the next project phase. The Access and Data Policy should include an AGATE Data Exchange Agreement and an AGATE Data License. This task will probably require a legal expertise besides (free or pro bono) consultations for example of the European IPR Helpdesk. Such an expertise could for example be commissioned to a specialised law firm such as iRights.Law.[229] iRights.Law lawyers have already authored an expert report on the regulatory framework of IANUS and a juridical guide for DH

[225] See 'European Commission | Open Science (Open Access)', accessed 13.05.2017.
[226] See for example lessons learned from Europeana in Dekkers et al. 2013, slides 32-39, accessed 27.07.2016.
[227] See UNESCO 2015, pp. 15-16.
[228] See Romary, Mertens and Baillot 2016, pp. 11-15.
[229] See 'iRights.Law | Homepage', accessed 13.05.2017.

researchers.[230] Also members of the ALLEA Working Group on Intellectual Property Rights may be consulted for an expertise.[231]

Sufficient human and financial resources should be planned for this task, especially for the legal expertise, which is predominantly needed in the IPR area; writing contributor agreements, negotiation and agreements with partner content providers in a number of jurisdictions, and most importantly ensuring AGATE is not liable for any copyright infringement.

[230] See Klimpel and Weitzmann 2014 and Klimpel and Weitzmann 2015.
[231] See 'ALLEA | Permanent Working Group Intellectual Property Rights (IPR)', accessed 13.05.2017.

6 General Organisation and Governance of AGATE

In the following, recommendations for the organisation and governance of the AGATE project are given (suggestions for the sustainability of the project and possible organisation forms are sketched in chapter no. 7).

The organisation and governance structure of the future AGATE will depend to some degree on the actual legal and funding framework as well as on the composition of the core consortium. The following recommendations are formulated especially with an eye on the H2020 framework as a good practice for European infrastructure projects.

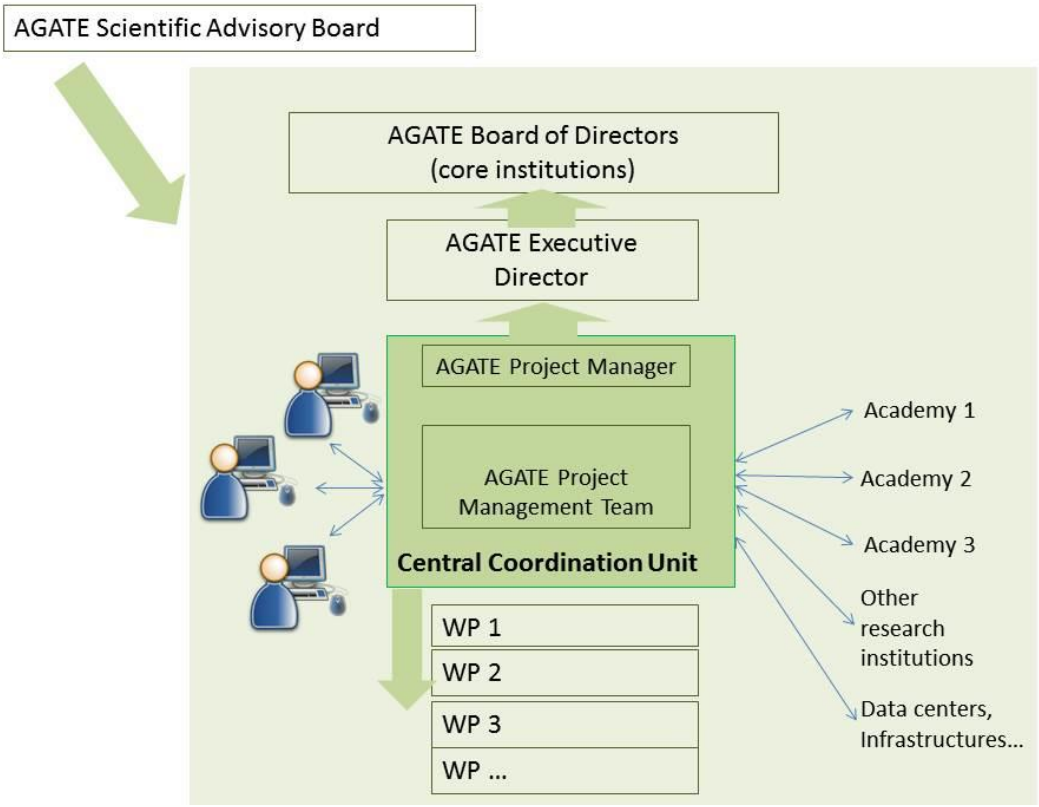

Figure 6: Organisation structure of AGATE

6.1 Central Coordination Unit

The Central Coordination Unit (CCU) will be responsible for the daily tasks and the operational steering of AGATE. It will consist of the Project Manager and the Project Management Team and will preferably be located centrally at the leading institution or in a distributed model over several institutions (according to the partners of the

Project Management Work Package (WP)). The CCU will report directly to the Board of Directors (respectively the Executive Director) as primary governing body.

The CCU will be mainly responsible for the project management and communication flows, controlling, and quality control of the WPs, and organisationally support the other bodies (for example by organisation of meetings and internal communication).

If this task is not part of a separate WP, the CCU will also be responsible for the contacts with the extended network, such as contact persons at (prospective) partner academies, individual projects or academies as data providers, other scientific organisations and infrastructures. Details will be outlined in the Consortium Agreement.

6.2 Board of Directors

A Board of Directors should be appointed as primary governing body. Because it will be the project's main decision-making body for strategic and political questions, the members of the Board of Directors should represent the major stakeholders of AGATE (that is the core consortium partners). The Board of Directors may choose for practical reasons to appoint an Executive Director of AGATE. Details on workflows, reporting strategies, and meetings will be outlined in the Consortium Agreement.

6.3 Scientific Advisory Board (SAB)

The Scientific Advisory Board will support the Board of Directors and the Project Management Team with scientific, technical and operational matters that are at the core of the activities of AGATE. The Scientific Advisory board should consist of six to eight leaders of international standing in a broad variety of fields of digital SSH and related fields.

The Scientific Advisory Board of the AGATE preparatory phase could be the nucleus for the Board of the implementation phase.[232] Details on workflows, reporting strategies, meetings will be outlined in the Consortium Agreement.

6.4 Work packages

Although the works packages will be defined in detail according to the funding programme and the consortium partners, based on the present concept and the

[232] A list of the current members of the AGATE Scientific Advisory Board can be found in the section 'AGATE Scientific Advisory Board' at the start of the report.

described features and activities of the proposed infrastructure we suggest at least the following work packages:

Work packages	
Project Management	Communication with EC, consortium partners, Scientific Advisory Board, strategic partners, coordination, monitoring
Business Model Development	Design of the governance structure and financial setting and long-term sustainability
Legal IPR setup	Legal review, cooperation partner consultations, legal agreements written, case handling procedures implemented
Technical setup	IT Infrastructure, Knowledge Map - Projects, Knowledge Map - Research Resources, Hub - Setup
Knowledge Map and Hub UX/UI Web Design	Agile methods implementation, strong relationship with Dissemination and Outreach, website builds, user consultation
Dissemination and Outreach	Online and Offline Activities (website and social media, mailing campaigns, events, travelling exhibition)
Community Management	Community engagement, development, testing and evaluation of online and offline formats and workflows
Training for data contributors	Development and testing of training resources (online material) and face to face events in cooperation with RI partners

7 Sustainability

The aim is to make AGATE a research infrastructure with European impact that is sustained by its network and that will have its own right of existence in the landscape of existing research infrastructures in the social sciences, arts and humanities. The phases as described by the Life Cycle Model of European RIs (see the following subchapter no. 7.1) leading to a fully-fledged infrastructure like AGATE will require a significant amount of human and financial resources. To ensure long-term return of these investments, attention needs to be paid to the sustainability of the infrastructure even in its initial stages.

As for all infrastructures, the sustainability of the infrastructure itself is a key issue. Especially (though not exclusively) for SSH infrastructures financial sustainability and even estimating costs for the operational phase is a problem because they are predominantly based on project funding and not (yet) institutionalised or brought into the working conventions of the academies.[233] The challenges faced are due to a lack of sustainable funding, of related sustainable business models, and that the landscape of digital infrastructure development is constantly changing.

For a pan-European infrastructure like AGATE to reach its maturity will involve developing solutions for technical, organisational, and legal questions. This process will employ a considerable amount of human and financial resources. To ensure that the future AGATE will be able to serve its user base in the long term, this chapter will point out key issues that need to be addressed and key activities that need to be carried out right from the beginning, encompassing the life phases of the RI.

Although issues of sustainability are being discussed already for a while for digital SSH and LIS projects, this issue has only recently been researched and described in a more detailed way.[234] During the last years, increasingly attention has been paid to the fact that the sustainability of a digital infrastructure needs to be addressed from the very beginning and cannot be merely addressed as an afterthought.[235] Sustainability for RIs means sustainable life cycle management. In this context, especially the lessons learned from the CENDARI project as laid out in a CENDARI report and further developed in PARTHENOS training materials are relevant.[236] Therefore we would like to recap some lessons learned from CENDARI and general sustainability issues of RIs.

[233] See Buddenbohm et al. 2015, European Strategy Forum on Research Infrastructures (ESFRI), p. 182, RfII – Rat für Informationsinfrastrukturen 2016, p. 37.
[234] See McGann 2010.
[235] See Edmond and Morselli 2016, p. 2.
[236] See Edmond and Morselli 2016. Francesca Morselli (KNAW) who had been involved in CENDARI suggested during her presentation during the AGATE Kick-Off Workshop in 2016 as one of the lessons learned from CENDARI to include already in this AGATE concept paper a first outline of a sustainability plan for the future AGATE (see Wuttke, Ott and Adrian, 2016, p. 15). A paper authored by her and Jennifer Edmond on this topic is forthcoming. For the PARTHENOS training material see 'PARTHENOS |

While developing any given digital RI it cannot be assumed that at some point, for example once the database is "ready" and filled with content and furnished with tools to access the contents, this is mostly when the project is completed and/or the funding runs out, that it can be just left by itself on the internet. Additionally, one needs to understand that a digital RI has a life cycle, meaning several phases dedicated to planning, implementation, maturity, and an end point (however far in the future it may be), that have to be kept in mind while developing a sustainability plan for a specific RI.

7.1 Life cycle model of European RIs (ESFRI):[237]

1. National Level/Design Study (develop scientific case and technical design).
2. If admitted to ESFRI Roadmap: Preparatory Phase (refinement of technical design, development of governance, definition of legal status and financial sustainability), (time window for Phase 1 and 2 for ESFRI projects on the roadmap is ten years).
3. Implementation Phase (adoption of legal status, engages substantial funding).
4. Operation Phase (several years or decades, 8-12% per year of the gross initial investment, constant upgrading, occasional major upgrades, wide impact).
5. Final Stage leading to its termination (decommissioning of the infrastructure, in case of distributed infrastructures national nodes may continue independently).

Complex digital infrastructures need maintenance, they "cannot simply be frozen in time and expected to continue to meet evolving needs"[238] and there are many other end products or infrastructural assets that need to be sustained which are often less tangible than data and tools such as knowledge networks. As most RIs are non-commercial enterprises funded with public money (mainly short-term project based research funding), they are extremely vulnerable as long-term infrastructure services if no sustainability business model has been developed at the moment the project funding runs out.[239]

Training Suite: Management Challenges in Research Infrastructures: Sustainability for Research Infrastructures', accessed 13.05.2017. Also during the infrastructure conference conducted in 2013 by the SASSH-project, the (crucial status of the) sustainability of SSH infrastructure was an issue of general concern (see Dusa, Oellers and Wolff 2014, pp. 226-227).

[237] Based on European Strategy Forum on Research Infrastructures (ESFRI) 2016, pp. 21-22, compare on a smaller scale the life cycle model of VREs used in Buddenbohm et al. 2015: Preparatory Phase, Prototype Phase, Development Phase, Operational Phase, Transfer Phase, Dismantling Phase.

[238] See Edmond and Morselli 2016, p. 1.

[239] This is a structural problem, for pros and contras, see RfII – Rat für Informationsinfrastrukturen 2016, p. 37.

7.2 Sustainability issues of digital research infrastructures

Sustainability issues of digital RIs encompass three main areas: "Research infrastructures need to be sustainable 1) organisationally and financially, 2) technically, and 3) in terms of human resources."[240] In these three main areas the following sub-aspects need to be addressed:[241]

1. **Organisational and financial issues** – a sustainable organisational structure and legal status such as ERIC or other suitable legal structure need to be developed and a wide uptake ensured, which ideally entails a wide geographical coverage and a broad user base, as well as a sustainable business model for example by guarantees of sustainment and further funding of the RI by their original founders or other appropriate bodies.
2. **Technical issues and research data** – the sustainability (long-term preservation) of research data (e.g. archives for data storage, use of PIDs for digital objects, etc.) and continuous technical development needs to be ensured to prevent the infrastructure to become obsolete. This includes aspects of the scalability of the infrastructure and deployment of dedicated authentication and authorisation services (AAI) to enable sharing of data, tools and services.
3. **Human resources** – human resources for the building, growing, and maintaining the RI itself need to be ensured to avoid (internal) knowledge loss through high staff turnover. Additionally, social networking, education, training and research support for existing and future users that ensure the uptake of the infrastructure and the collection and sharing of data. This can be by means of implementation into university curricula, development of the infrastructure as a Knowledge Sharing Infrastructure across its members, the active collection of new data and development of new tools, etc.

It is obvious that the topic of the sustainability of the future AGATE will start with the sustainability of the very data it aims to collect, aggregate, and make available via the AGATE Knowledge Map and Hub. Therefore, a detailed data management plan (DMP) should be developed as good practice for the next project phase (for example mandatory under a EU funding scheme such as H2020) and continuously updated during the following project phases. As the actual technical DMP of AGATE will be quite detailed and highly depend on the state-of-the-art at the start point of the next

[240] See European Strategy Forum on Research Infrastructures (ESFRI) 2016, p. 182.
[241] See European Strategy Forum on Research Infrastructures (ESFRI) 2016, p. 182, with additions from ESF 2011, p. 7, RfII – Rat für Informationsinfrastrukturen 2016, pp. 36-37, Edmond and Morselli 2016, p. 11.

phase, as well as the actual consortium, the rest of this chapter will be dedicated to general recommendations concerning the overall sustainability of AGATE as an infrastructure itself.

7.3 Recommendations

These following preliminary recommendations provide points of departure for more detailed sustainability planning during the next phase and on the way to the next phase because discussing options and clarifying sustainability issues at an early stage helps convince prospective data providers and other stakeholders of the project's general perspective:[242]

- Continually seek to expand group of core users (data providers).
  Recommendation – foresee outreach activities in the core community, actively engage main stakeholders in design and further development of the tools (create a sense of ownership), provide hands-on workshops and help, as well as FAQs. Seek early on to engage and involve researchers in academies' projects beyond the core consortium by inviting them to integrate information about their relevant resources in the portal, thereby increasing their visibility among research communities and the general public. See also chapters no. 3. on AGATE Hub and 4. on User Involvement).

- Foresee strategies to continually expand the group of academies and other relevant partners, invite concrete commitment.
  Recommendation – develop strategies how to formally integrate new partners and develop partner agreements for (new) partners including commitments for further uptake of AGATE, foresee areas for the prominent presentation of partners (incentive) and information how to join. See also chapters on AGATE Hub and Outreach.

- Foresee updating mechanisms, because only if the information provided via AGATE's individual components and tools is up to date (and not full of broken links) it will be of wide interest.
  Recommendation – seek cost effective workflows, seek to involve institutional and individual users, and use automatisation (flagging of time stamps for update milestone).

- Develop a business model for the Operation Phase of AGATE during the phase(s) beforehand.[243]
  Recommendations:

[242] See RfII – Rat für Informationsinfrastrukturen 2016, p. 37, Edmond and Morselli 2016, p. 12.
[243] Strategies for example in RfII – Rat für Informationsinfrastrukturen 2016, p. 38. Also of interest concerning business models for digital RIs is Benedict et al. 2015.

1. High level solution for full-fledged infrastructure: discuss and evaluate possibilities for individual commitment (financial, infrastructural, and/or human resources) from (some of the) core partner institutions, esp. the European academies, draft and implement a maintenance agreement by individual institution(s) (comparable to EHRI).

2. Low level solution: cooperate with and seek to integrate ((a)part(s) of) AGATE into existing ERICs or other infrastructures, or use their core services and organisational structures. For example joining an existing DARIAH Working group, by contributing to a pan-European digital SSH Knowledge Map (as an update of the existing DARIAH project registry, by using DARIAH's projected data deposit brokerage service, or employing established in-kind contributions, such as HAL or TextGrid).

- To reach high impact and uptake within and outside the envisaged community, the success criteria of RIs and VREs and how to reach them should be taken into account when developing engagement and outreach and dissemination activities as these form the basis for decisions about continuation and institutionalisation.[244]
  Recommendation – define concrete evaluation criteria and plan evaluation points (phases) in engagement and outreach and dissemination activities. Update and revise regularly. See also chapter no. 4.

- Develop a Data Management Plan (DMP).
  Recommendation – the DMP will record a strategy for the project to preserve all digital outputs (data and other resources).[245] It will be developed by the partners of the core consortium. The DMP should pay attention to using and creating well-documented, standardized code and choosing trustable preservation partners, as well as developing a project internal strategy to document and preserve (tacit) knowledge within the consortium. Include Milestones/Deliverables of the DMP (as checkpoints for achieved goals) in the overall project plan and plan resources accordingly. Plan regular updates.

- Think of options for end project plans.
  Recommendation – foresee reuse strategies for components, tools, and materials and the data by an ERIC, a related organisation or an infrastructure.

[244] See RfII – Rat für Informationsinfrastrukturen 2016, p. 38, Buddenbohm et al. 2014, Buddenbohm et al. 2015.

[245] For the current guidelines in H2020 see European Commission 2016. A useful tool could be the Research Data Management Organiser (RDMO) that is currently being developed by a German consortium with funding from the DFG, see 'RDMO (Research Data Management Organiser) | Homepage', accessed 13.05.2017. Details about the solutions (technical and more general) envisaged for CENDARI are contained in Edmond and Morselli 2016. These seem to provide a good basis. Not all of the envisaged measures seem to have been achieved yet, though some are part of the work in PARTHENOS, so they may be available in the future.

At the moment, AGATE's project results are envisaged as individual components and tools with the potential to be easily integrated into different contexts. Plan and negotiate an Exit Strategy for the Final Stage early and plan development accordingly (modular). Pay attention to documentation (of technical aspects, tacit knowledge) from the very beginning.

To ensure the full impact and sustainability of AGATE will be a core task that needs to be addressed from the beginning.[246] Special focus should be on:

- Activities that seek to engage the community to increase usage bottom-up (see chapter no. 2) and to enhance the project's general impact (see chapter no. 4).
- Development of the detailed sustainability plan (including institutional top-down approaches) and the DMP should be assigned to an individual Work Package (WP) or task and corresponding milestones included in the overall project plan.
- Both exercises (DMP and sustainability plan) need to be addressed right from the beginning of the next phase in close cooperation and knowledge exchange with other RI projects (especially PARTHENOS and DESIR)[247] and the SSH ERICs and engaging all relevant stakeholders inside and outside the consortium, including research funders and policy makers.

[246] Francesca Morselli during Kick-Off Workshop, see Morselli 2016, slide 12.
[247] The project DESIR (DARIAH ERIC Sustainability Refined) has just kicked off, no website is yet available. Initial information can be found on 'European Commission | CORDIS: DESIR', accessed 13.05.2017.

8 Implementation of AGATE: Suggested Components and Estimated Resources

AGATE is envisioned as a scaleable and modular infrastructure that is proposed to be built in an exploratory fashion to ensure user and stakeholder adoption. This approach is informed by the methodologies of Service Design (see: Service Design, chapter no. 2) which encourages incremental development in technology projects. Because of this modular development approach and *alternative implementation scenarios* it means we can only provide indicators of resource estimates and not a fixed budget. The term 'alternative implementation scenarios' wishes to express that AGATE can be implemented via various routes, especially concerning different configurations (members of the core consortium with various expertise and possibilities for in-kind contributions, degree of coverage, etc.) and different financial preconditions which will influence which stages and modules will be prioritised.

The next phase is a small start on national level, as the Mainz Academy will develop a German projects database expanding the contents of the current database of the Akademienunion with the findings of the AGATE project and technology already employed for the Mainz Academy database (for the German Pilot see chapter no. 2). From this pilot project experiences will be gained that will be of use for further steps.

In general, we estimate that at least a three-year time scale will be needed to reach a mature system that has undergone necessary development cycles (beta-version of the Knowledge Map and Hub, as well as accompanying activities, see chapter no. 7.1 on the life cycle of RIs). In the following an outline is given of the range of staff roles required, with overheads and associated budget for expenditure which might be also useful to similar projects. It is foreseeable that once the main development of the Knowledge Map is concluded, running costs for AGATE will be considerably lower, but resources for technical improvements and updating and especially costs for community and general outreach activities should be considered. Developing a workable business plan for AGATE will be a major task of the implementation phase, when one of the alternative development scenario paths has been set.

8.1 Estimation of the resources needed for the implementation

What is presented here in the 'resources and costs' section are indicators for activities and modules needed to carry out such a large-scale research infrastructure as opposed to fixed budgets. The reason for this *outline approach,* is that the AGATE proposal envisages a mapping exercise and service provision that has the ability to scale in terms of the breadth — the numbers of academy partners to be included and depth —

the quantity of the academies research publishing and digital research resources that would be databased.

Importantly in the area of resources and budgeting what the AGATE concept report highlights is the need for the *separations of concerns* in terms of managing modern day cataloging of research data. This separation of concerns means that the research data and metadata, creation and upkeep needs to be mainly the responsibility of the research originator, which could be either the institutions, projects or scholars. AGATE then acts as a research dissemination gateway and of equal importance as a help-point for the research community, promoting — good practice, and policy guidance — for the producers of the research metadata.

In the context of resources and costs, this pushing back of responsibility for creating *good data* is more than strategic, it is a necessity, as not only does it demand unjustifiable costs, it creates unworkable and error prone systems, which in turn put off users.

AGATE has two points of value, adding impact and dissemination value for the research it disseminates, but also helping those responsible for producing research to implement better data handling policies. Resources and budget wise the combination of databasing and promoting good practice in metadata creation is a more economical approach.

Three main areas of a prospective AGATE implementation phase have to be resourced: 1. core organisational fixed costs, 2. the Knowledge Map database, and 3. the Hub. The infrastructure would have three phases of implementation and then annual running costs. These three phases are: setup and partner liaison; module one - academy projects databasing; module two - research publication and digital resource databasing. In modules one and two, activities of the Knowledge Map and Hub are carried out.

Scaling AGATE is not only impacted by the *separations of concerns* issues related to data quality and maintenance, but also by needing to add new modular features to AGATE that would be thrown up by further research and high-level consultation with other European Open Science infrastructure providers. Needing to respond to evolving European CRIS needs is in part why AGATE has taken a strategic approach of 'Test-driven development' (TDD). TDD means that the system is being constantly assessed and improved. For AGATE this has resource implications, in that it needs a permanent development and 'user experience' (UX) in-house team.

Of special consideration for AGATE is addressing legal issues in 'open intellectual property rights' (OpenIPR) areas. This is in relationship to the data AGATE will be using in terms of negotiating the use under open licensing agreements of database content, and content enrichment in the databasing process — which is in effect authorship. Specialised legal counsel will be needed to carry out due diligence on setup and help

draft new agreements, such as ‘contributor agreements’ to compliment open licences which are in common use. Legal support will also be needed to be available on call to address queries or disputes. The types of legal agreements needed are ‘contributor agreements’ as used in cases where multiple users are adding to an aggregated body of work involving intellectual property.

The time schedule for AGATE would be as follows, a thirty-six month development run until all systems are established and running well, then an annual budget thereafter.

<table>
<tr><td colspan="4">Year 1</td><td colspan="4">Year 2</td><td colspan="4">Year 3</td><td rowspan="4">Annual running costs</td></tr>
<tr><td>Q1</td><td>Q2</td><td>Q3</td><td>Q4</td><td>Q1</td><td>Q2</td><td>Q3</td><td>Q4</td><td>Q1</td><td>Q2</td><td>Q3</td><td>Q4</td></tr>
<tr><td colspan="3">Start Q1 - year 1
Setup - 9 months</td><td colspan="4">Start Q4 - year 1
Module 1 – projects – 18 months</td><td colspan="4">Module 1 maintenance</td></tr>
<tr><td colspan="5"></td><td colspan="3">Start Q2 - year 2 - Module 2 - resources - 18 months</td><td>Module 2 maintenance</td></tr>
</table>

Table 8: AGATE development schedule

8.2 Estimation of human resources needed

See the AGATE organogram (see chapter no. 6) for areas of activity and section categories. The estimation of human resource would need to cover full staffing costs and overheads related to each staff position. Some of these positions could be covered by staff members or roles already able to members of the core consortium (e.g. as in kind contributions), such as; student assistants, PhD students, or Postdocs.

Roles (NB: these are roles and NOT full-time positions. One person could cover several roles)	**Notes**
Core organisation	
Research and project director (Director)	Oversight and public facing duties
Project Manager (PM)	
(Assistant)	Assistant to Director and PM

Administrator/office manager	
Junior administrator	
Analyst/researcher - Postdoc	
*The **AGATE Knowledge Map** and **AGATE Hub** teams work as one closely connected unit to deliver services and provide continual improvement (design) of system.*	
AGATE Knowledge Map - A database of academy projects & their digital resources	
(System Development (SysDev) and User Experience team)	
Project Manager	
(Assistant)	
Programmer	SysDev team would also support the Hub
Network and System Administrator	
Junior programmer	
UX designer	
Junior UX designer	
Copywriter	Translation, documentation
(Data management team)	
Project Manager	
Metadata specialist	
Data administrator	
AGATE Hub - a web portal for online & offline community engagement	
Project Manager	
(Assistant)	
Community manager	Working with stakeholders
Support	Direct support for users
Event manager	Workshops, conferences, user groups
Copywriter	Manuals, guides, translation

Table 9: AGATE human resources

8.3 Estimation of other costs

The following line items would have to be covered in this section of other costs:

- technical infrastructure - hardware, service contracts, specialist consultants, software licences
- specialised staff training budget
- data records use costs (such a Deutsche Nationalbibliothek)[248]
- workshops - system development, research, training
- travel costs
    - team - directly employed staff
    - consortium - key collaborators that need to be brought together
    - scientific Advisory Board
- outreach - publication costs, design, translation costs, promotional materials
- research publishing
- special legal counsel and services (IPR/Copyright)
- design sub-contacting corporate identity
- design sub-contacting web and mobile graphic/web design
- consultants pot for engagement and scenario testing
- advocacy organisation membership fund
- reading, literature and media fund

[248] See 'DNB | data costs', accessed 28.04.2017.

Bibliography

A downloadable version of the bibliography can be found here on Zotero: https://www.zotero.org/groups/1174271/agate/items/collectionKey/VR6FXPWB

Academy of Finland. “Digital Humanities (DIGIHUM) Academy Programme 2016–2019, Programme Memorandum, Academy of Finland (3.3.2015).” Academy of Finland, 2015. http://www.aka.fi/globalassets/32akatemiaohjelmat/digihum/ohjelmamuistio_digihum_en_final.pdf.

Acatech. “Acatech | Homepage.” Accessed April 30, 2017. http://www.acatech.de/.

ACDH-ÖAW. “ACDH-ÖAW | Publications,” Accessed April 18, 2017. http://www.oeaw.ac.at/acdh/de/biblio.

———. “ACDH-ÖAW | Resources & Tools,” Accessed May 14, 2017. http://www.oeaw.ac.at/acdh/en/resources.

Adrian, Dominik, Carolin Ott, and Ulrike Wuttke. “Open – Connective – Sustainable / AGATE: A Community Driven Digital Infrastructure for SSH Research(ers) at the European Academies, Report of the Second AGATE Workshop, 16.01.2017,” April 6, 2017. http://www.hsozkult.de/conferencereport/id/tagungsberichte-7102.

Agile Business Consortium. *DSDM Handbooks*, 2017. https://www.agilebusiness.org/resources/dsdm-handbooks.

ALLEA. “ALLEA | Homepage.” Accessed April 30, 2017. http://www.allea.org/.

———. “ALLEA | News entry 24.03.2017.” Accessed April 20. 2017, http://www.allea.org/allea-publishes-revised-edition-european-code-conduct-research-integrity/.

———. “ALLEA | Permanent Working Group Intellectual Property Rights (IPR).” Accessed May 13, 2017. http://www.allea.org/permanent-working-group-intellectual-property-rights/.

———. “ALLEA | SAPEA: Science Advice for Policy by European Academies.”, Accessed May 13, 2017. http://www.allea.org/asap-academies-sciences-advice-to-policy/.

———. “Open Science for the 21st Century, A Declaration of ALL European Academies Presented at a Special Session with Mme Neelie Kroes, Vice-President of the European Commission, and Commissioner in Charge of the Digital Agenda on Occasion of the ALLEA General Assembly Held at Accademia Nazionale Dei Lincei, Rome, on 11-12 April 2012,” 2012. http://www.allea.org/wp-content/uploads/2015/09/OpenScience-Rome-Declaration-final_web.pdf.

———. “The European Code of Conduct for Research Integrity. Revised edition initiated by the ALLEA Permanent Working Group on Science and Ethics,” 2017. http://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf

———. “ALLEA Statement of Enhancement of Open Access to Scientific Publications in Europe, Statement Prepared by the ALLEA Permanent Working Group on Intellectual Property Rights,” 2013. http://www.allea.org/wp-content/uploads/2015/08/Statement_ALLEA_Open_Access_2013-11.pdf.

ALLEA Permanent Working Group on Intellectual Property Rights. “Supplementary Statement on Enhancement of Open Access to Scientific Publications in Europe,” 2015. http://www.allea.org/wp-content/uploads/2015/12/Supplementary-Statement_Open-Access_FINAL.pdf.

Allianz der deutschen Wissenschaftsorganisationen. “Allianz der deutschen Wissenschaftsorganisationen | Priority Initiative “Digital Information”.“ Accessed May 13. http://www.allianzinitiative.de/en/.

Anderson, Sheila, Tobias Blanke, and Stuart Dunn. “Methodological Commons: Arts and Humanities E-Science Fundamentals.” *Philosophical Transaction of the Royal Society* 368 (2010): 3779–96. doi:10.1098/rsta.2010.0156.

———. “Methodological Commons: Arts and Humanities E-Science Fundamentals.” *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 368, no. 1925 (August 28, 2010): 3779–96. doi:10.1098/rsta.2010.0156.

Andreoli-Versbach, Patrick, and Frank Mueller-Langer. *Open Access to Data: An Ideal Professed but not Practised*. RatSWD Working Papers Series 2013. Berlin: SCIVERO, 2013. http://www.ratswd.de/dl/RatSWD_WP_215.pdf.

ARIADNE. “ARIADNE | Homepage.” Accessed May 14, 2017. http://portal.ariadne-infrastructure.eu/about.

———. “ARIADNE | Support Portal.” Accessed April, 30, 2017. http://ariadne-support.dcu.gr/

Arnold, Erik, Kate Barker, and Stig Slipersæter. “Research Institutes in the ERA: WP2 2007/S 106-12999 FORESIGHT-200702 Lot 2 WP3,” July 20, 2010. http://ec.europa.eu/research/era/docs/en/research-institutes-in-the-era.pdf.

Ayris, Paul. *Realising the European Open Science Cloud*, 2016. https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none.

BAdW. “BAdW | Digital Research Projects.” Accessed April 18, 2017. https://badw.de/akademie-digital/digitale-forschung.html.

BADW. “BAdW Forschungsvorhaben.” Accessed April 10, 2017. https://www.badw.de/forschungseinrichtung/forschungsvorhaben.html.

Baillot, Anne. “OPEN – Contribution to the AGATE Workshop January 16th, 2017.” Billet. *Digital Intellectuals*. Accessed May 5, 2017. https://digitalintellectuals.hypotheses.org/3040.

BARTOC. “BARTOC | Homepage.” Accessed May 14, 2017. https://bartoc.org/en.

BAS. “BAS | Institute for the Study of Societies and Knowledge.” Accessed May 8, 2017. http://issk-bas.org/.

BBAW. “BBAW | Digital Knowledge Store.” Accessed May 14, 2017. http://wsp.bbaw.de/.

———. “BBAW | Edoc-Server.” ST-03 Bild-Text-Block. Accessed April 27, 2017. https://edoc.bbaw.de/home.

Beall, Jeffrey. “The Weaknesses of Full-Text Searching,” *The Journal of Academic Librarianship* 34, no. 5: 438-444.

Benedict, Karl, Mairi Best, Sue Fyfe, and Clifford Jacobs. “Sustainable Business Models for Brokering Middleware to Support Research Interoperability: A Report from the Sustainable Business Models Team to the Brokering Governance Working Group of the Research Data Alliance (RDA),” 2015. https://www.rd-

alliance.org/system/files/SustainableBusinessModelsforBrokeringMiddleware%20-%20RDA%20WG%20Report.pdf.
BMBF. *Roadmap für Forschungsinfrastrukturen: Pilotprojekt des BMBF*. Bonn: BMBF, 2013. https://www.bmbf.de/pub/Roadmap_Forschungsinfrastrukturen.pdf.
Borek, Luise, Quinn Dombrowski, Jody Perkins, and Christof Schöch. "TaDiRAH: A Case Study in Pragmatic Classification" 10, no. 1 (2016). http://www.digitalhumanities.org/dhq/vol/10/1/000235/000235.html.
British Academy. "British Academy | Academy Research Projects." Accessed May 10, 2017. http://www.britac.ac.uk/academy-research-projects-0.
Budapest Open Access Initiative. "Budapest Open Access Initiative | Homepage." Accessed May 13, 2017. http://www.budapestopenaccessinitiative.org/.
———. "Budapest Open Access Initiative | View Signatures." Accessed May 13, 2017. http://www.budapestopenaccessinitiative.org/list_signatures.
Buddenbohm, Stefan, Nathanael Cretin, Elly Dijk, Bertrand Gaiffe, Maaike De Jong, Nathalie Le Tellier-Becquart, and Jean-Luc Minel. "State of the Art Report on Open Access Publishing of Research Data in the Humanities." Report. DARIAH, August 12, 2016. https://halshs.archives-ouvertes.fr/halshs-01357208/document.
Buddenbohm, Stefan, Claudia Engelhardt, and Ulrike Wuttke. "Angebotsgenese für ein geisteswissenschaftliches Forschungsdatenzentrum." *Zeitschrift für Digitale Geisteswissenschaften*, 2017. doi:10.17175/2016_003.
Buddenbohm, Stefan, Harry Enke, Matthias Hofmann, Jochen Klar, Heike Neuroth, and Uwe Schwiegelshohn. *Erfolgskriterien für den Aufbau und nachhaltigen Betrieb virtueller Forschungsumgebungen*. DARIAH-DE Working Papers 7. DARIAH-DE, Niedersächsische Staats- und Universitätsbibliothek Göttingen, 2014. http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2014-7.pdf.
———. "Success Criteria for the Development and Sustainable Operation of Virtual Research Environments." *D-Lib Magazine* 21, no. 9/10 (10 2015). doi:10.1045/september2015-buddenbohm.
CASRAI. "CASRAI | Homepage." Accessed May 12, 2017. http://casrai.org/.
CCeH. "CCeH | LAZARUS." Accessed May 13, 2017. http://cceh.uni-koeln.de/lazarus/.
CENDARI. "CENDARI | Homepage." Accessed May 13, 2017. http://www.cendari.eu/.
CESSDA. "CESSDA | Homepage." Accessed May 13, 2017. https://cessda.net/.
CIDOC CRM. "CIDOC CRM | Homepage." Accessed April 18, 2017. http://www.cidoc-crm.org/.
CLARIAH. "CLARIAH | Access to and Sharing of Research Data from Public Funding - Global Research Village OECD." Accessed May 14, 2017. http://dh-projectregistry.org/projects/review/182.
———. "CLARIAH | DH Project Registry." Accessed May 14, 2017. http://dh-projectregistry.org/.
———. "CLARIAH | Dutch Overview Digital Humanities." Accessed May 10, 2017. https://www.clariah.nl/projecten/dodh.
CLARIN. "CLARIN | Homepage." Accessed May 13, 2017. https://www.clarin.eu.
CLARIN, DARIAH-EU. "CLARIN and DARIAH-EU | DH Course Registry." Accessed May 10, 2017. https://registries.clarin-dariah.eu/courses/courses/.
CLARIN-D. "CLARIN-D | DMP Wizard." Accessed May 14, 2017. https://www.clarin-d.de/de/aufbereiten/datenmanagementplan-entwickeln.

———. “TeLeMaCo, The Linguistic Teaching Resources Hub.” Accessed May 8, 2017. https://fedora.clarin-d.uni-saarland.de/hub/.
Cockburn, Alistair. *Writing Effective Use Cases*. 16. The Agile Software Development Series. Boston: Addison-Wesley, 2006.
Cragin, Melissa H., P. Bryan Heidorn, Carole L. Palmer, and Linda C. Smith. “An Educational Program on Data Curation (Poster for 2007 STS Conference Poster Session),” 2007. http://hdl.handle.net/2142/3493.
Crossref. “Crossref | Funder Registry.” Accessed April 13, 2017. https://www.crossref.org/services/funder-registry/.
Czmiel, Alexander. “The Digital Knowledge Store. ‘Digital Infrastructures and Data Curation for Humanities Scholarship’, Presentation Slides at first AGATE workshop, 13.06.2016,” 2016. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Veranstaltungen/Workshop_AGATE/5_Czmiel_AGATE-Workshop_13.06.2016.pdf.
DARIAH. “DARIAH | Digital Methods and Practices Observatory Working Group (DiMPO) - Zenodo Community.” Accessed May 2, 2017. https://zenodo.org/communities/dimpo/?page=1&size=20.
———. “DARIAH | Dienste und Werkzeuge.” Accessed May 8, 2017. https://de.dariah.eu/list-services.
———. “DARIAH | Homepage.” Accessed May 13, 2017. http://dariah.eu/.
———. “DARIAH | #dariahTeach learning platform.” Accessed May 5, 2017. http://dariah.eu/teach/.
DARIAH-DE, and DiRT. “DARIAH-DE and DiRT | TaDiRAH (Taxonomy of Digital Research Activities in the Humanities).” Accessed July 18, 2014. http://tadirah.dariah.eu/vocab/index.php.
DASISH. “DASISH | Homepage.” Accessed May 13, 2017. http://dasish.eu/.
Data Archive. “Data Archive | Research Data Lifecycle.” Accessed May 13, 2017. http://www.data-archive.ac.uk/create-manage/life-cycle.
DataCite. ”DataCite | Homepage.” Accessed May 13, 2017, https://www.datacite.org.
de.hypotheses. “de.hypotheses | Homepage.” Accessed May 18, 2017, http://de.hypotheses.org.
Dekkers, Makx, Michiel de Keyzer, Nikolaos Loutas, and Stijn Goedertier. “Lizenzen für Daten und Metadaten, Open Data Support Trainingsmodul 2.5”, 2013. https://www.europeandataportal.eu/sites/default/files/d2.1.2_training_module_2.5_data_and_metadata_licensing_de_edp.pdf.
DFG. “DFG | GEPRIS (German Projects Information System).” Accessed May 5, 2017. http://gepris.dfg.de/gepris/OCTOPUS?language=en.
———. “DFG | Press Release no. 63, 2015.” Accessed May, 8, 2017. http://www.dfg.de/en/service/press/press_releases/2015/press_release_no_63/index.html.
———. “DFG | Subject Areas Structure.” Accessed April 30, 2017. http://www.dfg.de/en/dfg_profile/statutory_bodies/review_boards/subject_areas/index.jsp.
DCC. “DCC | Data Management Plans.” Accessed May 8, 2017. http://www.dcc.ac.uk/resources/data-management-plans.
DINI. “DINI-Zertifikat Dokumenten- und Publikationsservice 2010, DINI Schriften 3-de [Version 3.0, Juni 2010]“, June 2010.

DiRT Directory. “DiRT Directory | Homepage.” Accessed May 8, 2017. https://dirtdirectory.org/.
DNB. “DNB | “Bibliographic Services” Data Service.” Accessed April 10, 2017. http://www.dnb.de/EN/Service/DigitaleDienste/Datendienst/datendienst_node.html.
“DNB | Data Costs.” Accessed April 28, 2017. http://www.dnb.de/DE/Service/DigitaleDienste/Datendienst/datendienst_node.html.
———. “DNB | Gemeinsame Normdatei (GND).” Accessed May 12, 2017. http://www.dnb.de/DE/Standardisierung/GND/gnd_node.html.
———. “DNB | OAI-PMH Harvester Manager.” Accessed April 28, 2017. http://oaiharvestmangr.sourceforge.net/.
———. “DNB | Projekte - ORCID DE – Förderung der Open Researcher and Contributor ID in Deutschland.” Accessed April 13, 2017. http://www.dnb.de/DE/Wir/Projekte/Laufend/orcidDE.html.
DOI Foundation. “DOI Foundation | Homepage.” Accessed May 12, 2017. http://www.doi.org/.
Doorn, Peter. “Why the Social Sciences and Humanities Need Research Infrastructures.” *E-Data & Research* Special Issue 2014 (2014): 1.
Dublin Core. “Dublin Core | Dublin Core Metadata Element Set, Version 1.1.” Accessed May 14, 2017. http://dublincore.org/documents/dces/.
DuraSpace. “DuraSpace | DSpace-CRIS 5.5.0.” Accessed May 14, 2017. http://www.duraspace.org/node/2845.
———. “DuraSpace | First Release of DSpace CRIS Module for DSpace 3.” Accessed May 14, 2017. http://duraspace.org/now-available-first-release-dspace-cris-module-dspace-3.
———. “DuraSpace | Technologies.” Accessed May 14, 2017. http://duraspace.org/technologies.
Duşa, Adrian, Claudia Oellers, and Simon Wolff. “A Common Agenda for the European Research Infrastructures in the Social Sciences and Humanities.” In *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*, edited by Adrian Duşa, Dietrich Nelle, Günter Stock, and Gert G. Wagner, 225–32. Berlin: Scivero, 2014. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/Tagungsbaende/FACING_THE_FUTURE.pdf.
EASSH. “EASSH | Homepage.” Accessed May 13, 2017. http://www.eassh.eu/.
Ebert, Barbara. “CONNECTIVE (1) Pan-European Project Database (Comments), Presentation Slides at Second AGATE Workshop, 16.01.2017,” 2017. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/BMBF-Projekt/Barbara_Ebert.pdf.
Edmond, Jennifer, and Francesca Morselli. “What Does it Mean to Sustain a Research Infrastructure? Experiences and Recommendations from the Experience of the CENDARI Project (Deliverable 2.4).” CENDARI, 2016. http://www.cendari.eu/sites/default/files/CENDARI_D2.4%20Sustainability%20Plan%20final%20%282%29.pdf.
ESF. *Research Infrastructures in the Digital Humanities*. Vol. 42. Science Policy Briefing. European Science Foundation, 2011. http://archives.esf.org/hosting-experts/scientific-review-groups/humanities-hum/strategic-activities/research-infrastructures-in-the-humanities.html.
ESS. “ESS | Homepage.” Accessed May 13, 2017. http://www.europeansocialsurvey.org/.

EU-Büro des BMBF. “EU-Büro des BMBF I Homepage.” Accessed May 13, 2017. http://www.eubuero.de/.
euroCRIS. “euroCRIS | Homepage.” Accessed April 13, 2017. http://www.eurocris.org/.
———. “euroCRIS | Main Features of CERIF.” Accessed April 18, 2017. http://eurocris.org/cerif/main-features-cerif.
European Commission. “European Commission | CORDIS: DESIR.” Accessed May 13, 2017 http://cordis.europa.eu/project/rcn/207190_en.html.
———. “European Commission | European Open Science Cloud.” Accessed May 13, 2017. http://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud.
———. “European Commission | Open Science (Open Access).” Accessed May 13, 2017. https://ec.europa.eu/programmes/horizon2020/en/h2020-section/open-science-open-access.
———. “H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020, Version 3.0, 26 July 2016,” 2016. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.
———. “H2020 Programme: Multi-Beneficiary General Model Grant Agreement (H2020 General MGA — Multi), Version 4.0, 27 February 2017.” Accessed May 13, 2017. http://ec.europa.eu/research/participants/data/ref/h2020/mga/gga/h2020-mga-gga-multi_en.pdf.
European IPR Helpdesk. “European IPR Helpdesk | Homepage.” Accessed May 13, 2017. https://www.iprhelpdesk.eu/.
European Strategy Forum on Research Infrastructures (ESFRI). *Strategy Report on Research Infrastructures: Roadmap 2016*. n.p.: Science and Technology Facilities Council, 2016. http://www.esfri.eu/sites/default/files/20160309_ROADMAP_browsable.pdf.
Europeana. “Europeana | Homepage.” Accessed May 13, 2017. http://www.europeana.eu/portal/en.
———. “Europeana | Provide Data in EDM.” Accessed May 18, 2017. http://pro.europeana.eu/page/provide-data-edm.
Fedora. “Fedora Repository | Homepage.” Accessed April 20, 2017. http://fedorarepository.org/.
FORCE11. “FORCE11 | The FAIR Data Principles.” Accessed May 13, 2017. https://www.force11.org/group/fairgroup/fairprinciples.
FOSTER. “FOSTER | Homepage.” Accessed May 5, 2017. https://www.fosteropenscience.eu.
GESIS. “GESIS | Access.” Accessed May 13, 2017. http://www.gesis.org/en/services/archiving-and-registering/data-archiving/access/.
———. “GESIS | Data Archiving.” Accessed May 13, 2017. http://www.gesis.org/en/services/archiving-and-registering/data-archiving/.
Getty Foundation. “Getty Foundation | Online Scholarly Catalogue Initiative.” Accessed May 12, 2017. http://www.getty.edu/foundation/initiatives/current/osci/.
GitHub. “GitHub | Homepage.” Accessed May 13, 2017. www.github.com.
González-Blanco García, Elena. “SUSTAINABLE (Part 2) - User Involvement (Comments), Presentation Slides at Second AGATE Workshop, 16.01.2017,” 2017. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/BMBF-Projekt/Elena_Gonzalez-Blanco_Garcia.pdf.

Harrower, Natalie, ed. *Going Digital: Creating Change in the Humanities (ALLEA E-Humanities Working Group Report)*. Berlin: ALLEA, 2015. http://dri.ie/sites/default/files/files/Going-Digital-digital-version.pdf.
Heinrich, Maurice, Anne Sieverling, Felix Schäfer, and Sabine Jahn. “Digitale Forschungsdaten in den Altertumswissenschaften. Stakeholderanalyse 2013 zu Forschungsdaten in den Altertumswissenschaften. Teil 2: Kombinierte Auswertung & Interpretation. Version 1.0, 05. Mai 2015.” IANUS Forschungsdatenzentrum Archäologie & Altertumswissenschaften, 2015. Accessed June 14, 2017. doi:10.13149/000.jah37w-q.
Hippel, Eric von. *Democratizing Innovation*. 1st MIT Press paperback ed. Cambridge, Mass.: MIT Press, 2006.
H-Soz-Kult. “H-Soz-Kult | Editorial: The Status Quo of Digital Humanities in Europe (23.10.2014).” Accessed May 13, 2017. http://www.hsozkult.de/debate/id/diskussionen-2375.
Hude, Nicole von der, and Jochen Rupp. “OAI in der Deutschen Nationalbibliothek.” *Dialog mit Bibliotheken* 2013, no. 1: 83-84.
IIIF. “IIIF (International Image Interoperability Framework) | Homepage.” Accessed April 28, 2017. http://iiif.io/.
Immenhauser, Beat. “Digital Humanities – ein Thema der SAGW.” *Bulletin SAGW* 2015, no. 4 (November 2015): 38–39.
International ISRC Agency. “ISRC (International Standard Recording Code) | Homepage.” Accessed May 12, 2017. http://isrc.ifpi.org/en/.
iRights.Law. “iRights.Law | Homepage.” Accessed May 13, 2017. https://irights-law.de/.
ISAN International Agency. “ISAN (International Standard Audiovisual Number) | Homepage.” Accessed April 18, 2017. http://www.isan.org/.
ISMWC. “ISMWC | International Standard Musical Work Code - Homepage.” Accessed April 18, 2017. http://www.iswc.org/.
ISNI International Agency. “ISNI (International Standard Name Identifier) | Homepage.” Accessed April 18, 2017. http://www.isni.org/.
ISSN International Center. “ISSN (International Standard Serial Number) | Homepage.” Accessed May 12, 2017. http://www.issn.org/.
ISSP. “ISSP (International Social Survey Programme) | Homepage.” Accessed May 13, 2017. http://www.issp.org/menu-top/home/.
ISTC. “ISTC (Incunabula Short Title Catalogue) | Homepage.” Accessed April 18, 2017. http://data.cerl.org/.
Jahn, Sabine, and Dietrich von Seggern. “IANUS | IT-Empfehlungen: PDF-Dokumente.” Accessed May 13, 2017. https://www.ianus-fdz.de/it-empfehlungen/pdf-dokumente.
Jörg, Brigitte et al., ed. “CERIF 1.3 Full Data Model,” 2012. http://www.eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf.
Klimpel, Paul, and John H. Weitzmann. *Forschen in der digitalen Welt. Juristische Handreichung für die Geisteswissenschaften*. DARIAH-DE Working Papers 12. Göttingen: DARIAH-DE, Niedersächsische Staats- und Universitätsbibliothek, 2015. http://resolver.sub.uni-goettingen.de/purl/?dariah-2015-5.
———. “Lizenzierung von Materialien und Datenbank durch das Forschungsdatenzentrum IANUS, Expertise für das Forschungsdatenzentrum IANUS. Zu rechtlichen Aspekten bei

der Langzeitarchivierung digitaler Forschungsdaten von Dritten," 2014. http://dx.doi.org/10.13149/000.3h7mtr-d.
Knowledge Exchange. "Virtual Research Environments - Description." Accessed April 28, 2017. http://www.knowledge-exchange.info/index.php/projects/project/virtual-research-environments.
Kobos, Mateusz, Łukasz Bolikowski, Marek Horst, Paolo Manghi, Natalia Manola, and Jochen Schirrwagen. "Information Inference in Scholarly Communication Infrastructures: The OpenAIREplus Project Experience." *Procedia Computer Science*, 10th Italian Research Conference on Digital Libraries, IRCDL 2014, 38 (January 1, 2014): 92–99. doi:10.1016/j.procs.2014.10.016.
Lagoze, Carl, Herbert Van de Sompel, Michael Nelson, and Simeon Warne, ed. "The Open Archives Initiative Protocol for Metadata Harvesting, version 2.0," June 14, 2002. http://www.openarchives.org/OAI/openarchivesprotocol.html.
Leathem, Camilla, and Dominik Adrian. *Survey and Analysis of Basic Social Science and Humanities Research at the Science Academies and Related Research Organisations of Europe*. Berlin: Union of the German Academies of Sciences and Humanities, 2015. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/ProjectReport_SASSH_2015.pdf.
Leibniz-Institut für Astrophysik Potsdam, and University of Applied Sciences Potsdam. "AIP and FHP | RDMO (Research Data Management Organiser)," Accessed June 12, 2017. https://rdmorganiser.github.io/.
Leopoldina. "Leopoldina | Policy Advice." Accessed May 13, 2017. https://www.leopoldina.org/en/policy-advice.
Library of Congress. "Library of Congress | VRA CORE - a Data Standard for the Description of Works of Visual Culture." Accessed April 12, 2017. https://www.loc.gov/standards/vracore/.
Luise, Borek, and BARTOC. "BARTOC | Taxonomy of Digital Research Activities in the Humanities," October 15, 2015. https://bartoc.org/it/node/704.
McGann, Jerome. "Sustainability: The Elephant in the Room." In *Online Humanities Scholarship: The Shape of Things to Come*, edited by Jerome McGann, 5–14. OpenStax CNX, 2010. http://cnx.org/contents/3d5747d3-e943-4a39-acf9-beb086047378@1.3.
Morselli, Francesca. "Chances and Challenges of a European Academies Internet Gateways, Presentation Slides at first AGATE workshop, 13.06.2016." 2016. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Veranstaltungen/Workshop_AGATE/8_Morselli_AGATE-Workshop_13.06.2016.pdf.
MPG. "MPG | Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities," 2003. http://openaccess.mpg.de/Berliner-Erklaerung.
———. "MPG | Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities: Signatories," 2003. https://openaccess.mpg.de/319790/Signatories.
———. "MPG | Open Access Ambassadors." Accessed May 13, 2017. http://oambassadors.mpdl.mpg.de/.
MTA. "MTA | Centre for Social Sciences." Accessed May 8, 2017. http://www.tk.mta.hu/en/research.
———. "MTA | Homepage." Accessed April 30, 2017. http://mta.hu/.
———. "MTA | Repository of the Academy's Library." Accessed May 14, 2017. http://real.mtak.hu/.

———. “MTA | Research Documentation Centre.” Accessed May 8, 2017. http://openarchive.tk.mta.hu/.
Neuroth, Heike, Achim Oßwald, Regine Scheffel, Stefan Strathmann, and K. Huth. *nestor Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung (Online Version 2.3)*. Göttingen: nestor, 2010. http://nestor.sub.uni-goettingen.de/handbuch/nestor-handbuch_23.pdf.
OKF. “OKF | Linked Open Vocabularies (LOV).” Accessed April 18, 2017. http://lov.okfn.org/dataset/lov/.
Open Archives Initiative. “Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH).” Accessed April 28, 2017. https://www.openarchives.org/pmh/.
Open Library of Humanities. “Open Library of Humanities | Homepage.” Accessed May 12, 2017. https://www.openlibhums.org/.
OpenAIRE. “OpenAIRE | For Researchers.” Accessed May 13, 2017. https://www.openaire.eu/intro-researchers.
———. “OpenAIRE | Guidelines.” Accessed June 15, 2017. https://guidelines.openaire.eu/en/latest/.
———. “OpenAIRE | Homepage.” Accessed May 13, 2017. https://www.openaire.eu/.
———. “OpenAIRE | OpenAIRE API Documentation.” Accessed May 14, 2017. http://api.openaire.eu/.
———. “OpenAIRE | Project Factsheets - General Information.” Accessed May 12, 2017. https://www.openaire.eu/project-factsheets.
ORCID. “ORCID | Homepage.” Accessed May 12, 2017. https://orcid.org/.
ORCID DE. “ORCID DE | Post Heinz Pampel 12.01.2017.” *ORCID DE*, January 12, 2017. http://www.orcid-de.org/good-response-to-the-consortium-for-improving-contributor-identification-in-germany/.
Ott, Carolin. “AGATE Service and Information Hub Part 1 - Community, Presentation Slides at Second AGATE Workshop, 16.01.2017,” 2017. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/BMBF-Projekt/Carolin_Ott_Service_and_Information_Hub_-_Community.pdf.
PAN. “PAN | Institute of Philosophy and Sociology.” Accessed May 8, 2017. http://www.ifispan.pl/en/.
PARTHENOS. “PARTHENOS | Training Suite: Management Challenges in Research Infrastructures: Sustainability for Research Infrastructures.” Accessed May 13, 2017. http://training.parthenos-project.eu/sample-page/management-challenges/sustainability-for-research-infrastructures/.
Paskin, N. “Toward Unique Identifiers.” *Proceedings of the IEEE* 87, no. 7 (July 1999): 1208–27. doi:10.1109/5.771073.
Piwik. “Piwik | Homepage.” Accessed June 6, 2017. https://piwik.org/.
Quora. “Quora | What is the Difference between UX and UI Designer and Web Designer?” Accessed April 25, 2017. https://www.quora.com/What-is-the-difference-between-UX-and-UI-designer-and-web-designer.
RDMO. “RDMO (Research Data Management Organiser) | Homepage.” Accessed May 13, 2017. http://rdmorganiser.github.io/en/.
Re3data.org. “re3data.org | Homepage.” Accessed May 13, 2017. http://www.re3data.org/.

RfII – Rat für Informationsinfrastrukturen. *Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland*. Göttingen, 2016. http://www.rfii.de/?wpdmdl=1998.

RIA. "RIA | Digital Humanities Observatory: Digital Research and Projects in Ireland'." Accessed April 10, 2017. http://drapier.dho.ie/.

Riley, Jenny. "Seeing Standards." Accessed May 14, 2017. http://jennriley.com/metadatamap/.

Romary, Laurent. "Sustainable Data for Sustainable Infrastructures." In *Facing the Future: European Research Infrastructures for the Humanities and Social Sciences*, edited by Adrian Duşa, Dietrich Nelle, Günter Stock, and Gert G. Wagner, 187–201. Berlin: Scivero, 2014. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/Tagungsbaende/FACING_THE_FUTURE.pdf.

Romary, Laurent, and Emiliano Degl'Innocenti. "PARTHENOS Standardization Survival Kit, Deliverable D4.1 October 31, 2016." http://www.parthenos-project.eu/Download/Deliverables/D4.1_standardization_survival_kit.pdf.

Romary, Laurent, Mike Mertens, and Anne Baillot. "Data Fluidity in DARIAH – Pushing the Agenda Forward." *BIBLIOTHEK Forschung Und Praxis* 39, no. 3 (2016): 350–57.

Royal Society. "Royal Society | European Academies' Statement: Science Is Global." Accessed April 30, 2017. https://royalsociety.org/topics-policy/publications/2016/european-academies-statement-science-is-global/.

———. "Royal Society | Homepage." Accessed April 30, 2017. https://royalsociety.org/.

———. "Royal Society | Open Access Publishing." Accessed May 13, 2017. https://royalsociety.org/journals/authors/open-access/.

SAGW. "SAGW | Homepage." Accessed April 30, 2017. http://www.sagw.ch/sagw.html.

———. "SAGW | Open Access Strategie der SAGW." Accessed April 30, 2017. http://www.sagw.ch/de/sagw/laufende-projekte/open-access/oa-news/oa-strategie-sagw.html.

Sahle, Patrick. "Forschungsdaten in den Geisteswissenschaften." *Bulletin SAGW* 2015, no. 4 (November 2015): 43–45.

Schirrwagen, Jochen. "AGATE Search over digital resources, Comments at second AGATE workshop, 16.01.2017," 2017. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/BMBF-Projekt/Jochen_Schirrwagen.pdf

Schmidutz, Daniel, Lorna Ryan, Anje Müller Gjesdal, and Koenraad De Smedt. "Report about new IPR Challenges: Identifying Ethics and Legal Challenges of SSH Research (Deliverable D6.1 of DASISH)," 2013. http://dasish.eu/publications/projectreports/D6.1_final.pdf.

Schrade, Torsten. "Projektdatenbank Akademienunion. Akademienforschung sichtbar machen. Presentation Slides at the Strategy Meeting of the Executive Committee of the Union of the German Academies of Sciences, 01.03.2017". http://www.digitale-akademie.de/fileadmin/projektdatenbank/.

Science Europe. "Position Statement on Research Information Systems," 2016. http://www.scienceeurope.org/wp-content/uploads/2016/11/SE_PositionStatement_RIS_WEB.pdf.

Science.ORF.at. “Science.ORF.at | Open Access: Warum die Geisteswissenschaftler zögern (Interview with Walter Scheidel, 14.08.2013).” Accessed May 13, 2017. http://sciencev2.orf.at/stories/1723031/index.html.

Selhofer, Hannes, and Geser, Guntram. “D2.2 Second Report on Users’ Needs / Resources, Version 2.1 February 2015” Accessed May 14, 2017. http://www.ariadne-infrastructure.eu/fre/Resources/D2.2-Second-report-on-users-needs.

SHARE. “SHARE | Homepage.” Accessed May 8, 2017. http://www.share-project.org/.

Sicilia, Miguel-Ángel. “The Common European Research Information Format (CERIF) Semantic Vocabulary Specification,” 2013. http://eurocris.org/ontologies/semcerif/1.3/index.html.

The International ISBN Agency. “International ISBN Agency | Homepage.” Accessed May 12, 2017. https://www.isbn-international.org/.

The Royal Society of Edinburgh. “RSE | Becoming a Fellow.” Accessed April 30, 2017. https://www.rse.org.uk/fellows/becoming-a-fellow/.

ThoughtWorks. “ThoughtWorks | Continuous Integration.” Accessed May 3, 2017. https://www.thoughtworks.com/continuous-integration.

Travis CI. “Travis CI | Test and Deploy Your Code with Confidence.” Accessed May 3, 2017. https://travis-ci.org/.

UC3. “UC3 | DMPTool,” Accessed May 12, 2017. https://dmp.cdlib.org/.

US Department of Education. “Paradata in 20 Minutes or Less.” Google Docs. Accessed May 12, 2017. https://docs.google.com/document/d/1QG0lAmJ0ztHJq5DbiTGQj9DnQ8hP0Co0x0fB1QmoBco/edit?hl=en_US&usp=embed_facebook.

UNESCO. *Interoperability and Retrieval: Open Access for Library Schools, Module 4*. Paris: UNESCO, 2015. http://unesdoc.unesco.org/images/0023/002321/232199E.pdf.

Vandenbussche, Pierre-Yves. “Linked Open Vocabularies (LOV): A Gateway to Reusable Semantic Vocabularies on the Web | Www.semantic-Web-Journal.net,” July 2, 2015. http://www.semantic-web-journal.net/content/linked-open-vocabularies-lov-gateway-reusable-semantic-vocabularies-web.

Vierkant, Paul, Olaf Siegert, Gernot Deinzer, Agathe Gebert, Marc Herbstritt, Heinz Pampel, Regine Tobias, and Alexander Wagner. “Workflows zur Bereitstellung von Zeitschriftenartikeln auf Open-Access-Repositorien - Herausforderungen und Lösungsansätze.” *o-bib. Das offene Bibliotheksjournal / herausgegeben vom VDB* 4, no. 1 (April 7, 2017): 151–69.

VIVO. “VIVO | Homepage.” Accessed May 14, 2017. http://www.vivoweb.org/info/about-vivo.

W3C. “Sparql Implementations - W3C Wiki.” Accessed April 28, 2017. https://www.w3.org/wiki/SparqlImplementations.

Wellcome Library. “IIIF Search Inside - The Biocrats (Book).” *Wellcome Library*. Accessed April 28, 2017. http://wellcomelibrary.org/item/b18035978.

Wissenschaftsrat. “Wissenschaftsrat | Kerndatensatz Forschung.” Accessed April 13, 2017. http://www.kerndatensatz-forschung.de/.

Worthington, Simon. *Book to the Future - a Book Liberation Manifesto*. Hybrid Publishing Consortium, 2015.

Worthington, Simon, and Christina Kral. *A Publication Taxonomy*. Hybrid Publishing Consortium, 2014. https://research.consortium.io/docs/a_publication_taxonomy/a-publication-taxonomy.html.

Wuttke, Ulrike, Stefan Buddenbohm, and Claudia Engelhardt. “Kooperationspartner in Wissenschaft und Technik (HDC-Bericht 5.1.),” 2014. http://humanities-data-centre.org/wp-content/uploads/Public_2014_10_16_DeliverableTP3_AP5_IdentifikationKooperationspartner_v7.pdf.

Wuttke, Ulrike, Carolin Ott, and Dominik Adrian. “AGATE: Chances and Challenges of a European Academies Internet Gateway: Kick-Off Workshop of the Project “Elaboration of a Concept for a European Academies Internet Gateway (AGATE), Workshop Report 1.” Union of the German Academies of Sciences and Humanities, 2016. http://www.akademienunion.de/fileadmin/redaktion/user_upload/Publikationen/BMBF-Projekt/AGATE_Erster_Workshop_Bericht_23.08.2016.pdf.

Zenodo. “Zenodo | Homepage.” Accessed May 14, 2017. https://zenodo.org/.

Appendix

1. Partners and supporters of AGATE 1.0
2. AGATE Consortium, H2020-INFRADEV_01_2017 Design Studies call of the European Commission
3. AGATE workshops
4. Comparison Research Databases

1. Partners and supporters of AGATE 1.0

The table shows the European academies who have expressed their interest in the AGATE project in form of knowledge exchange, participation in the AGATE workshops or by formal letters of intent. The group of potential partners was approached at the beginning of the project by two letters to the Presidents of the ALLEA member academies: from the President of the Union of the German Academies, Professor Hatt; and from the President of ALLEA, Professor Stock. Interested academies were invited to participate in the two workshops organised by AGATE, to give feedback on the preliminary project description published on the Akademienunion's website and to contact the scientific coordinator and — in case they were interested to join the core consortium — to send a letter of intent. The table also shows which of these academies have participated in the H2020 design study proposal (AGATE 2.0) as consortium members or by letters of support.

2. AGATE Consortium, H2020-INFRADEV_01-2017 Design Studies call of the European Commission

The table shows the consortium participants of the AGATE proposal, submitted in response to the H2020-INFRADEV_01_2017 Design Studies call of the European Commission, which builds on results of AGATE 1.0 and will develop them further into the design of an Open Innovation Research Infrastructure. The lead organisation of the proposal is the Austrian Academy of Sciences, as the Austrian Center for Digital Humanities had taken over the task of coordinating the application process. The proposal was supported by forty institutions and infrastructures such as ALLEA, CLARIN, DARIAH, euroCRIS, and OpenAIRE, as well as by several national ministries, metaLab@Harvard, and CERN. The participating academies had already demonstrated strong interest in the AGATE project and concept 1.0 and were approached according to their expertise for participation in the H2020 Design Study call. The Union of the German Academies of Sciences and Humanities was considerably involved in the application process via the AGATE project and holds work package 6 "Dissemination and Outreach".

3. AGATE Workshops

AGATE organised two workshops which took place on June 13, 2016 and January 16, 2017 in Berlin, at the Berlin-Brandenburg Academy of Sciences and Humanities. The programme, a short review, the presentation slides and reports are provided on the website of the Union of the German Academies of Sciences and Humanities.

For the resources of the first workshop with links to the programme and workshop report, see:

http://www.akademienunion.de/en/working-groups/cooperation-projects/enagate/1-workshop/

For the resources of the second workshop with links to the programme and workshop report, see:

http://www.akademienunion.de/en/working-groups/cooperation-projects/enagate/2-workshop/

4. Comparison Research Databases

Fields and features compared in a tabulated list from the research databases being used to benchmark the AGATE database. The comparator databases include the following: The database of the Research Projects of the Academies' Programme of the Union of the German Academies of Sciences and Humanities; the database of the academy's projects of the Mainz Academy of Sciences and Literature; the Wissensspeicher (Digital Knowledge Store) of the Berlin-Brandenburg Academy of Sciences and Humanities; the DRAPIer-database (Digital Research and Projects in Ireland) of the Royal Irish Academy; the Digital Humanities Registry by CLARIAH; the OpenAIRE project database.

Appendix 1: Partners and supporters of AGATE 1.0

	Academy / Institution / Research Infrastructure	Acting member institution / department / project	Support
1	Academy of Athens		Participation in 2nd AGATE workshop
2	Academy of Sciences of Moldova		Letter of Support
3	Academy of Sciences of Turin		Letter of Support
4	Austrian Academy of Sciences	ACDH (Austrian Center for Digital Humanities)	Letter of Support, participation in 1st and 2nd AGATE workshop, coordination of H2020 proposal
5	British Academy		informal exchange
6	Bulgarian Academy of Sciences	Institute for the Study of Societies and Knowledge (ISSK), Institute of Information and Communication Technologies (ICCT), Institute for the Bulgarian Language (IBL)	Letter of Support
7	Croatian Academy of Sciences and Arts		Letter of Support
8	German National Academy of Sciences Leopoldina		participation in 2nd AGATE workshop
9	Hungarian Academy of Sciences	Library and Information Center of the Hungarian Academy of Sciences (MTA KIK)	Letter of Support, participation in 1st and 2nd AGATE workshop, István Monok member of Scientific Advisory Board
10	National Academy of Sciences of Belarus		informal exchange
11	Polish Academy of Sciences	Institute of Literary Research (IBL PAN) and Poznan Supercomputing and Networking Center (PSNC)	Letter of Support, participation in 2nd AGATE workshop
12	Polish Academy of Arts and Sciences Krakow		Letter of Support
13	Royal Irish Academy		Natalie Harrower member of AGATE Scientific Advisory Board
14	Royal Netherlands Academy of Arts and Sciences		Francesca Morselli participation in 1st AGATE workshop, knowledge exchange with DANS director Peter Doorn and others at DANS
15	Royal Swedish Academy of Sciences		participation in 2nd workshop
16	Royal Society of Edinburgh		informal exchange

Appendix 1: Partners and supporters of AGATE 1.0

17	Slovenian Academy of Sciences and Arts	Research Centre of the Slovenian Academy of Sciences and Arts (ZRC SAZU)	Letter of Support, participation in 1st and 2nd AGATE workshop
18	Swiss Academies of Arts and Sciences	Data and Service Center for the Humanities (DaSCH), Diplomatic Documents of Switzerland, Swiss Inventory of Coin Finds	Letter of Support, participation in 1st AGATE workshop
19	Union of the German Academies of Sciences and Humanities	Academy of Sciences and Literature, Mainz (Development of a German pilot)	Letter of Support
20	ALLEA		participation in 1st and 2nd AGATE workshop, cooperation and support in outreach and dissemination
21	CESSDA		knowledge exchange, invitation
22	CLARIN ERIC		participation in 1st and 2nd workshop
23	DARIAH-EU		Letter of Support, participation in 2nd AGATE workshop
24	euroCris		invitation of AGATE to Athens, participation in 2nd workshop
25	Europeana		invitation of AGATE to network workshop, participation in 1st workshop
26	Max Weber Foundation		participation in 2nd workshop
27	EGI		knowledge exchange
28	OpenAIRE		Letter of Support, participation in 1st and 2nd AGATE workshop

Key	Member of AGATE 2.0. consortium		

Appendix 2: AGATE Consortium, H2020-INFRADEV_01_2017 Design Studies call of the European Commission

	AGATE Partners	Website	Main Contact
1	Österreichische Akademie der Wissenschaften (OeAW)	http://www.oeaw.ac.at/acdh	Eveline Wandl-Vogt
2	Università degli Studi di Catania (UNICT)	http://archivio.unict.it/home	Roberto Barbera
3	Instytut Chemii Bioorganicznej PAN (PSNC)	http://www.man.poznan.pl/online/en/	Cezary Mazurek
4	Instytut Badan Literackich Polskiej Akademii Nauk (IBL)	http://chc.ibl.waw.pl/	Maciej Maryl
5	Institute of Information and Communication Technologies BAS (IICT- BAS)	http://www.iict.bas.bg/EN/	Ivan Georgiev
6	Institute for the Study of Societies and Knowledge (ISSK)	http://issk-bas.org/	Emilia Chengelova
7	Ibercivis	www.ibercivis.es	Fermin Serrano
8	Copenhagen Business School, Department of Innovation and Organizational Economics (CBS) (Denmark)	http://www.cbs. dk/en/research/departments	Marion Poetz
9	Ludwig Boltzmann Gesellschaft GmbH (LBG)	http://www.lbg.ac.at/en	Lucia Malfent
10	Znanstvenoraziskovalni center Slovenske akademije znanosti in umetnosti (ZRC SAZU)	http://www.zrc-sazu.si/en/node	Jerneja Fridl
11	winnovation consulting gmbh (winnovation)	http://www.winnovation.at	Gertraud Leimueller
12	Union der deutschen Akademien der Wissenschaften (Akademienunion)	http://www.akademienunion.de/en/	Ulrike Wuttke

Key	Project Coordinator
	Work Package Leads
	Participants

Appendix 4: Comparison Research Databases

Field name and functions (AGATE recommendation)	***Suggested type of input data***	***Example field content***	**Databases**	**1. Projects of the German academies programme on the website of the Union of the German Academies of Sciences and Humanities**	**2. Project database of the Academy of Sciences and Literature, Mainz (Status June 2017)**	**3. Digital Knowledge Store of the Berlin-Brandenburg Academy of Sciences and Humanities (Beta version, status May 2017)**	**4. DRAPIer (http://drapier.dho.ie/project.html#1)**	**5. Digital Humanities Registry (CLARIAH) (Information drawn from an exemplary project entry: http://dh-projectregistry.org/projects/review/182)**	**6. OpenAIRE project entries**	Challenges/ needs for research or coordination
User interface in English or more languages				*no (only German)*	*no (currently only in German; localization technically possible)*	*no (currently only in German; English is planned)*	*yes*	*yes*	*yes*	*translation into English*
MUST HAVE										
Project title	free text	Deutsche Inschriften des Mittelalters		yes	yes	yes	yes	yes	yes	-
Project acronym	free text	AGATE		none	none	yes	none	none	none	-
Project identifier	alpha/numeric	II.A.22		yes (Vorhabennummer)	yes (currently internal)	yes (not shown on the website), RDF-URI	none	1.) Project ID (number) 2.) External Identifiers (Currently the database only recognizes so-called NARCIS project identifiers (OND[number])	contract (GA) number	apply unique identifier for projects
Project maintaining academy	standard	BBAW		yes	yes	yes	none	Institutions	organizations	no complete unique identifier (UIDs) for organisations.
Institute/ office site	standard	University of Greifswald		yes (Arbeitsstelle)	yes (Arbeitsstelle)	no	none	1.) Institutes involved (free text) 2.) Institution Comment (free text)	via link to CORDIS project database	no complete unique identifier for organisations and sub sites and locations. A combination of UIDs and ontologies would be needed. CORDIS is one EU option http://cordis.europa.eu
Home country/ federal state	standard	Lower Saxony		yes (federal state)	yes (federal state)	yes	none	none	via link to CORDIS project database	Location and region: competing classifications and UID means mapping between different systems needed.
Permanent cooperation partner (joint research project)	standard	CCDH		none	yes (via institutional role model)	not shown yet, but planned in form of project relations	1.) Parent Project (drop down menu/list) 2.) Parent Project not listed (free text) 3.) Sub-projects (free text)	none	Organizations	no complete unique identifier (UIDs) for organisations, or relationships over time. A semantic relationship would need to be made in RDF.
Running time	standard (e.g. Start Date/ end Date (yyyy-mm-dd)	2014-2030		none	yes (currently internal)	yes	Start Date / End Date (yyyy-mm-dd)	Start Date (yyyy-mm-dd), End Date (yyyy-mm-dd)	start date (yyyy-mm-dd) end date (yyyy-mm-dd)	-
Project status	selection: running/ completed/ interrupted	running		considers only running projects	yes (running and completed projects)	yes (running and completed projects)	yes (running and completed projects)	none	yes (running and completed projects)	Gathering of the necessary data of completed projects, otherwise only running projects could be included.

Appendix 4: Comparison Research Databases

Field name and functions (AGATE recommendation)	*Suggested type of input data*	*Example field content*	**Databases**	**1. Projects of the German academies programme on the website of the Union of the German Academies of Sciences and Humanities**	**2. Project database of the Academy of Sciences and Literature, Mainz (Status June 2017)**	**3. Digital Knowledge Store of the Berlin-Brandenburg Academy of Sciences and Humanities (Beta version, status May 2017)**	**4. DRAPIer (http://drapier.dho.ie/project.html#1)**	**5. Digital Humanities Registry (CLARIAH) (Information drawn from an exemplary project entry: http://dh-projectregistry.org/projects/review/182)**	**6. OpenAIRE project entries**	Challenges/ needs for research or coordination
Summary of the project	description (free text)	Ziel des Projektes sind die Dokumentation und die historische Analyse des gesamten italienischen Wortschatzes von den Anfängen bis heute. Damit leistet das LEI einen zentralen Beitrag zur Kenntnis der italienischen Sprachgeschichte und zur Wahrung der sprach- und kulturgeschichtlichen Traditionen.		yes	yes	yes	description (free text)	description (free text)	More information	texts ideally have about the same length (at least define a max. number of characters), preferably short, medium and long descriptions would be best. Descriptions of different length are needed as they are used for different purposes. E.g., Google search result uses a 'metadescription' field, 140 characters long. To be equivalent to bibliographic description a description will be 400 characters long.
Photo/ images	upload			yes	yes	none	none	none	none	copyright clearance, get a picture
Project website	link	www.bbaw.de/forschung/dwds/uebersicht		URL to project description on the website of the respective academy and to project website	yes	yes	Website	Links	via link to CORDIS project database (but unclear where to find and not available in uniform way)	maintain links (prevent broken links), apply PURLs
Digital research resource(s)	standard (PID) or free link entering	www.dwds.de/		URL to main digital resource (digital edition or dictionary)	URLs to related institutions (office sites e.g. head of project, location for post, admin office etc.)	URLs to collection/ edition etc.	none	Links: Project Data (not systematised as beng added to a general category)	Research Data	maintain links (prevent broken links), apply PURLs
Digital publications	link to respective entries	1. Daniel Schmidutz et al., "Report About New IPR Challenges: Identifying Ethics and Legal Challenges of SSH Research," 2013		none	yes (depicted as further information)	yes	none	none	Publications	maintain links (prevent broken links), apply PURLs, DOIs
Project/ Publication type	standard	edition		yes, project type. Overall 11 categories, three main types: 1. dictionaries (2 subcategories), 2. editions (7 subcategories), 3. basic research in social sciences and cultural studies)	yes, publication type. 5 categories: database, text edition, edition of works, encylopedia, dictionary	yes, 5 categories: digital edition, database, website, publication, other	none (file types and formats)	none	Document type	Publication type: Choose a standard/Define a dropdown list

Appendix 4: Comparison Research Databases

Field name and functions (AGATE recommendation)	***Suggested type of input data***	***Example field content***	**Databases**	**1. Projects of the German academies programme on the website of the Union of the German Academies of Sciences and Humanities**	**2. Project database of the Academy of Sciences and Literature, Mainz (Status June 2017)**	**3. Digital Knowledge Store of the Berlin-Brandenburg Academy of Sciences and Humanities (Beta version, status May 2017)**	**4. DRAPIer (http://drapier.dho.ie/project.html#1)**	**5. Digital Humanities Registry (CLARIAH) (Information drawn from an exemplary project entry: http://dh-projectregistry.org/projects/review/182)**	**6. OpenAIRE project entries**	Challenges/ needs for research or coordination
Disciplinary classification	standard (e.g. RVK) (multiple specifications are possible)	ND 6400 - Mittelalterliche Numismatik Europas		yes, 18 disciplines (multiple specifications are possible)	yes, 12 disciplines (multiple specifications are possible)	yes, 26 disciplines (multiple specification are possible)	Disciplines	NWO Disciplines (drop down menu/list)	none	Completing ontologies, allow free text input, (Academic disciplines have unified 'ontologies' or classifications and at the same time there are separate field specific 'discipline' classifications). E.g. In physics it's called PACS (Physics and Astronomy Classification System) and is a 4 level deep taxonomy with 4000+ subjects
Chronological classification	standard (e.g. GND, BSZ periods) (multiple specifications are possible)	1. Drittel 2. Jh. v. Chr.		none	yes (7 epochs)	yes (4 epochs)	Temporal categories	none	none	Different period classifications would need to be allowed for different disciplines and granular time period needs.
Spatial classification	standard (e.g. Geonames) (multiple specifications are possible)	Europe		none	yes (German federal states; other classifications possible)	none	Irish Geographic Names	none	none	Use spacial classifications from a variety of systems with OpenStreetMap (OSM) acting as a resolver.
Digital methods	standard (multiple specifications are possible)	3D modelling		none	none yet (technically possible via generic classification system)	none	Methods and Techniques	1.) Research Activities (drop down menu, list based on TaDiRAH) 2.) Research Techniques (drop down menu, list based on TaDiRAH) 3.) Research Objects (drop down menu, list based on TaDiRAH)	none	Standards for methods have to be defined/indicated for use, such us TaDIRAH, decide if to allow free text description
Digital tools	standard (multiple specifications are possible)	oXygen XML Editor		none	none yet (technically possible via generic classification system)	none	none	none	none	categories have to be defined, e.g by Digital Research Tool directory http://dirtdirectory.org/, decide if to allow free text descriptions with a auto-complete options lookup name list.
Digital formats	standard (multiple specifications are possible)	XML-TEI		none	none yet (technically possible via generic classification system)	none	1.) Data Formats 2.) Metadata Formats	none	none	categories have to be defined, e.g. like http://drapier.dho.ie/projects.html, allow free text descriptions with a lookup name list.

Appendix 4: Comparison Research Databases

Field name and functions (AGATE recommendation)	*Suggested type of input data*	*Example field content*	**Databases**	**1. Projects of the German academies programme on the website of the Union of the German Academies of Sciences and Humanities**	**2. Project database of the Academy of Sciences and Literature, Mainz (Status June 2017)**	**3. Digital Knowledge Store of the Berlin-Brandenburg Academy of Sciences and Humanities (Beta version, status May 2017)**	**4. DRAPIer (http://drapier.dho.ie/project.html#1)**	**5. Digital Humanities Registry (CLARIAH) (Information drawn from an exemplary project entry: http://dh-projectregistry.org/projects/review/182)**	**6. OpenAIRE project entries**	Challenges/ needs for research or coordination
Funder	standard	DFG		not required as all projects have same funder	yes (differentiation between Akademienprogramm and others)	yes (not filled in yet)	Funding	1.) Funding Body(ies) (free text) 2.) Size of Grant (number) 3.) Grant Currency (drop down menu/list)	Funding	Decide if this information should be included/negotiate details
Social media activities (Twitter Account, Blog etc)	standard and links	https://twitter.com/leichenpredigt https://twitter.com/DFDmainz		none	none	none	none	none	none	-
Licences and data property rights	standards, as well as option for 'other'.	Creative Commons CC BY 4.0 DE		none	none yet (technically possible via generic classification system)	none yet (is planned)	none	none	Open Access mandate	-
NICE TO HAVE										
Further information (related to funder)	free text			yes	none	none	none	Hyperlinks to relevant web-based resources (free text)	none	Decide if necessary
Related cooperating projects, project title	free text	Deutsche Inschriften online (German OCR literature digitisation)		none	yes	yes, via project relations	none	none	via link to CORDIS project database	Decide if this information should be included/negotiate details
Contact information	input mask (title, name, institution, e-mail, telephone, ...)	Dr. Hans Mustermann Berlin-Brandenburgische Akademien der Wissenschaften, h.mustermann@xyz.de +49 12345 / 67890 orcid.org/0000-0001-2222-3333		contact (at the moment Sebastian Zwies as coordinator of the academies' programme is contact person for all projects)	at least one contact person (if project completed), in running projects, several persons that might be contacted; contact information is modelled according to vCard standard and embedded as microfromat; schema.org output also possible)	none	one contact person	none	via link to CORDIS project database	projects have to decide main contact person
Persons/ roles	input mask for role assignment, standards (ISNI, ORCID)	project lead/full-time personel/research associate/etc.		none	yes	yes	none	1.) Persons involved (free text) 2.) Persons Comment (free text)	via link to CORDIS project database (but unclear where to find and not available in uniform way)	ISNI and ORCIDs person identifiers need for coordination between these different sources.
Related cooperating project, Cooperation partners	standard	The Academy of Hebrew Language, CLARIN		none	none yet (technically possible via institutional role model)	none	none	none	via link to CORDIS project database	Decide if this information should be included/negotiate details
Related cooperating project, Project acronym	free text	OCR-D		none	none	none	none	none	via link to CORDIS project database	Decide if this information should be included/negotiate details
Further information (related to cooperations)	free text			none	none yet (technically possible via institutional role model)	none	Keywords	none	More information	-

Appendix 4: Comparison Research Databases

Field name and functions (AGATE recommendation)	*Suggested type of input data*	*Example field content*	**Databases**	**1. Projects of the German academies programme on the website of the Union of the German Academies of Sciences and Humanities**	**2. Project database of the Academy of Sciences and Literature, Mainz (Status June 2017)**	**3. Digital Knowledge Store of the Berlin-Brandenburg Academy of Sciences and Humanities (Beta version, status May 2017)**	**4. DRAPIer (http://drapier.dho.ie/project.html#1)**	**5. Digital Humanities Registry (CLARIAH) (Information drawn from an exemplary project entry: http://dh-projectregistry.org/projects/review/182)**	**6. OpenAIRE project entries**	Challenges/ needs for research or coordination
Reuse statements	free text	A PhD student has used research data from the project for her PhD project and has won an award		none	none yet (technically possible)	none	none	none	Link to project (feature allowing user to relate a publication to a project)	-
PhD projects possible (early stage researchers)	yes/no	none		none	yes (strategic effort of Mainz academy)	none	none	1.) Is PhD Project (yes/no) 2.) PhD Students involved (number)	none	-
Citizen's participation in the project (Citizen Science)	yes/no	none		none	none	none	none	none	none	-