

Ulrike Wuttke, Carolin Ott, 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

Version: 23.08.2016

Date: 13.06.2016
Time: 11.30 a.m. – 5:30 p.m.
Venue: Workshop room 1
Berlin-Brandenburg Academy of Sciences and Humanities
Jägerstraße 22/23
10117 Berlin – Germany

Project title: “Elaboration of a Concept for a European Academies Internet Gateway (AGATE)”
Project coordination: Union of the German Academies of Sciences and Humanities
Research Funding: Federal Ministry of Education and Research (BMBF)
Funding period: October 2015–March 2017

Citation: Ulrike Wuttke, Carolin Ott, 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).



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

General

Since October 2015 the Union of the German Academies of Sciences and Humanities coordinates the project “Elaboration of a Concept for a European Academies Internet Gateway for the Social Sciences and Humanities (AGATE)”, conducted in close collaboration with ALLEA. On 13 June 2016, about 30 experts from European science academies, infrastructure projects in the fields of Digital Humanities and Social Sciences, and related infrastructures, such as DARIAH, CLARIAH(NL)/CLARIN, Europeana, OpenAire, and the German Council for Scientific Information Infrastructures met at the Berlin-Brandenburg Academy of Sciences and Humanities (BBAW) for the project’s Kick-Off-Workshop, to discuss the “Chances and Challenges of a European Internet Gateway for the Social Sciences and Humanities (AGATE)”. The workshop’s objective was to share experiences and know-how, identify common chances and challenges of the proposed AGATE, and to pinpoint synergies and possibilities for future collaboration.

Executive Summary

The workshop opened with a session on the “Status Quo of digital Research Practices and Publications at the European Academies today”. Several representatives of European academies gave insights into the challenges the digitisation process poses to their academies’ SSH-projects, available solutions, and desiderata.

The presentation about the Hungarian Academy of Sciences (MTA) had a strong focus on digital publications and (digitised) sources. In Hungary Open Access has a good standing in the Humanities and Social Sciences. There are an extensive number of academic repositories, which currently poses the problem of harmonising and combining data sets in central databases in order to seize their full potential. The MTA is the major national player of the Hungarian academic sector and already a strong advocate of AGATE in this early stage of the initiative.

Data linkage and data preservation is one of the big challenges for the Swiss Academy of Humanities and Social Sciences (SAGW). While all “born digital” data are openly accessible, there are still many restrictions for more traditional forms of research output. In the past years the SAGW has increasingly assumed a leading role in the digital turn and as a networker for the Swiss SSH. By now guidelines for data management and data maintenance are being developed to foster harmonisation and coordination between the relevant institutions.

In Croatia a strategy for the digital turn is currently being mapped out by the Croatian Academy of Sciences and Arts (HAZU) partly based on the desiderata of its researchers which have been revealed in a survey. There is a strong demand for more digitalisation in the humanities, but also a significant lack of knowledge about the existing relevant initiatives and their actual services in particular. Education in this field is thus highly welcome and seen as a desired feature of a future AGATE. While Open Access is well established for Croatian scientific journals via a common platform, the openly accessible number of monographs is not very significant.

The Austrian Academy of Sciences (ÖAW) provides with the Austrian Center for Digital Humanities (ACDH) an advanced advice and support service for the SSH research projects run by its institutes. Through and beyond the technical support the ACDH also seeks to improve the visibility of SSH to reach a broader public, e.g. by fostering Open Access and Open Science. The academy is internationally well connected and runs several repositories.

The discussion focussed on the following issues:

- The diversity and heterogeneity of the academies’ SSH data and metadata will define the efforts needed to integrate them and shape (and limit) the form and features of the AGATE-search function.

- The participants agreed that the long-term perspective and relevance of the research output in the SSH pose particular challenges to the sustainability of digital research methods and publication practices. Therefore, AGATE should in this context be more than a technical toolbox, but a platform that advocates education and awareness rising.
- AGATE is also expected to increase the visibility of the SSH research projects performed at the academies and function as a lobbying tool for them.
- There was a strong tendency among the participants to define the target groups of AGATE. For the design of AGATE it makes a difference if it primarily addresses the needs of researchers or the general public.
- Generally, the presentations and discussions of the panel suggested that the European academies - albeit their heterogeneity in institutional forms and in function - play a key role for the SSH. Joining forces, integrating resources and strengthening the networking activities through a platform like AGATE thus are very promising activities also for the general standing of SSH in Europe. The presentations showed that individual academies have certain foci of expertise in the digital SSH that complement each other, thus creating potential synergies for AGATE.

In the afternoon session, fundamental aspects of the proposed AGATE were presented and discussed:

The Digital Knowledge Store of the Berlin-Brandenburg Academy of Sciences and Humanities (BBAW) served as an example for an elaborate search engine that allows browsing and searching through the research output of a wide range of different projects of one academy, wherefore it seems to have great potential to be a role model for the envisioned search function of AGATE. The project has integrated digital resources of the academy's projects in a painstaking process that included the manual adjustment of metadata and API's of very heterogeneous resources. A similar process for the output of every single project of the European Academies would apparently be impossible. It occurs that certain requirements would have to be met to enable an automated harvesting of the respective (meta-)data. These requirements could be determined in guidelines. Such guidelines are currently being developed for the Digital Knowledge Store. Also the Research Center of the Slovenian Academy of Sciences and Arts (SRC ZASU) runs an advanced database of current and completed programs and projects that could serve as a model for AGATE. SRC ZASU runs also several repositories and has a well-developed culture of Open Access.

The added value of developing a space that facilitates trainings, services, and information for the academies' research community as part of AGATE were examined thoroughly during the afternoon session. Like in the first panel, several participants stressed the expectation that AGATE should function as a platform for educational efforts and be a driving force of change. Therefore it was advised that the researchers' needs should clearly be identified in order to offer the right services. To avoid reinventing the wheel, the existing infrastructure initiatives and their activities must be known very well and wherever possible integrated.

Several participants underlined that conceptualizing and implementing a platform like AGATE should be seen as a change process that requires a suitable change management considering time, diplomacy, and persuasive power. It should not be underestimated that the context of the academies among the research institutions is unique. They form stable institutions with long traditions, a feature that is certainly useful and eventually needed for the permanent curation of data in the future. At the same time, their structure is rather closed and hierarchical which can also be an obstacle for the implementation of change.

The participants agreed repeatedly that AGATE has to be more than a technical solution. Its relevance lies equally in the support of digital research practices and network-building amongst the

academies. Therefore, the second workshop on 16 January 2017 at the BBAW will focus on practical information, especially the knowledge of existing offers by European SSH-infrastructure and possibilities for cooperation. It will address researchers, ICT and digital librarian experts, and other interested parties from European Academies.

Detailed Workshop Report

Short introduction of the Project – Objectives of the Workshop

The workshop was officially opened by the President of the Union of the German Academies of Sciences and Humanities (short: Akademienunion), Hanns Hatt, who recounted the project's background. From 2013 to 2015, the Akademienunion, in close cooperation with ALLEA, had conducted a large *Survey and Analysis of Basic Social Science and Humanities Research at the Science Academies and Related Research Organisations of Europe* (SASSH). This survey was funded by the German Federal Ministry of Education and Research (BMBF). The study identified significant potential for improving access to digitised research publications from academies' research, like digitisations of primary sources, digital editions, databases, but also electronic publications in the form of articles or monographs. Another finding was that researchers and experts from the academies' IT-infrastructure and libraries require better access to services around digital tools and technologies, standards, licences, and long term preservation. In addition, the study called for an improved exchange of knowledge, experiences and information. It further highlighted the need for enhanced cooperation and collaboration between academies' projects, as well as with European digital infrastructure initiatives such as CLARIN, DARIAH or Europeana.

The follow-up project, again sponsored by German Federal Ministry of Education and Research, is based on the conviction that many of these challenges could be addressed by a pan-European digital infrastructure for the academies' research, the future AGATE. The vision behind AGATE is to enable the exploration of the enormous stores of knowledge that European academies create, as well as to encourage deepened national and international collaborations between academies and beyond.

Hereafter, Ulrike Wuttke gave a short introduction to the AGATE-project and its objectives. She underlined that the objective of AGATE has to be understood as three stages that build on one another and are closely interlinked. The purpose of AGATE is 1) to increase the general visibility of the academies' SSH research, 2) to improve the findability and accessibility of digital research results (aka digital resources), 3) to foster scientific cooperation and collaboration amongst the academies and between the academies and relevant infrastructure partners - that means relevant national, European and international infrastructures and initiatives - to ensure the sustainability of the digital research and publication practices.

The proposed AGATE-Search Function would provide a unified source of information on the research undertaken at respectively by the academies to make it visible to non-specialists. This central research projects database in combination with a connection to the digital research output, for example in the form of a meta catalogue, will showcase and promote the richness, quantity, quality, and broad relevance of the SSH-research at the European academies and open the doors for reuse and collaboration. She explained that as visibility is not an end in and of itself and digital research results such as databases or digital editions are more relevant if used, it is not only the academies' duty, but also part of their self-interest to justify the huge efforts that are necessary to produce and preserve these data, by making them as visible and accessible as possible to a vast audience for reuse. Therefore progress has to be made in the fields of 1) open access in general, 2) giving access

to data sets (in contrast to giving access to electronic publications in the form of monographs or books, etc.), and 3) interoperability.

Wuttke illustrated the similar challenges that many academies as institutions and individual academies' projects are facing: 1) an increased need for support to master the switch from analogue to digital, 2) low general awareness for the sustainability of digital research methods and publication practices on the side of the researchers as well as on the side of the decision makers, and to conduct active risk management by knowledge transfer and documentation, 3) tight research budgets that currently lead to constrained resources for ensuring sustainable access to digital research results and long term preservation. She stressed that a sensible solution to meet these challenges is to bundle respectively pool the academies' efforts and activities in these fields in close cooperation with infrastructure partners.

Morning Session: The Status Quo of digital research practices and publications at the European Academies Today – Experience reports

CHAIR: Laurent Romary

István Monok (MTA, Hungarian Academy of Sciences): The Status Quo of Digital Research Practice and Publications at the Hungarian Academic World Today

Slides

Text of the presentation

ISTVÁN MONOK first explained the Hungarian academic research structure, which differs significantly from Germany and other European countries. The structure is very centralized and the state is the main research funder, with the Hungarian National Research, Development and Innovation Office as main funder. He stated that at the moment the Social Sciences and Humanities (SSH) are in a disadvantageous position in the competition for research funds, they receive only approx.10% of the total budget. The Hungarian Academy of Sciences (MTA) is a major player in Hungarian academic research. Around 98% of its research budget is spent on the sciences, only about 2-3% on SSH, while around 10% of its researchers at institutes are in the field of SSH.

Then, Monok focussed on digital research methods and publication practices. He introduced the MTMT (Store of Hungarian Scientific Publications), the central Hungarian database for scientific output (<https://www.mtmt.hu/>) which is maintained by the MTA, respectively the Library and Information Center of the Hungarian Academy. Since the MTMT was initially installed by institutes of natural sciences, it includes a citation index for each publication (science metrics). This receives serious resistance from humanities researchers. Monok also explicated that in Hungary Open Access is by default mandatory for scientific publications which have to be uploaded as pdf-versions to the Repositories of the Academy's Library (REAL) (Open Access Mandate of the Hungarian Academy of Sciences). However, a publisher may establish an embargo period, but the embargoed resources are nevertheless visible at the researchers' workplace. Problems are caused by the multiplicity of systems for bibliographic databases, which exist in Hungary and which have to be unified in the future. The broad acceptance of Open Access in Hungary has in Monok's opinion something to do with the rather gloomy situation of the Hungarian scientific publishing system for books in the Humanities and in Hungarian. As most of the books are anyway self-published, the willingness to publish them Open Access is high in order to reach a broader audience, especially amongst younger researchers. Recently, a need for a centralized Open Journal System platform has been identified, which would also lead to standardisation (use of international ID numbers (DOI, ORCID) and automatisation).

Concerning research data, Monok introduced COMPASS, a subscribed database of databases. It contains research databases which have been created in Hungary. He posed the question, how we can consider databases (which are collaborative work) as scientific output in the humanities, and offered as a solution the model of the natural sciences, where a publication can have many authors. Finally, he expressed that AGATE would be of utmost importance and usefulness to the MTA to aggregate Hungarian content to a European level and that the academy's library has the required organisational and technical experience to play an active role.

Beat Immenhauser (SAGW): Swiss Academy of Humanities and Social Sciences: Kick-Off Workshop of the project “Elaboration of a concept for a European Academies Internet Gateway (AGATE)”

Slides

First BEAT IMMENHAUSER gave a general overview about the SAGW. The SAGW (as member of the Swiss Academies of Arts and Sciences) is an umbrella organisation of 60 learned societies with approximately 30.000 members. It does not directly fund research projects, because this is a task of the Schweizer Nationalfonds, but it funds Research Infrastructures and the diffusion of research findings.

Immenhauser explained that the digital turn, in this case the extension of digital research practices and publications in SSH-projects, is an important challenge for all Swiss research institutes. He exemplified main challenges of the digitisation of the workflow on the basis of databases (e.g. Historical Dictionary of Switzerland), digital editions (Diplomatic Documents of Switzerland), and digitisation projects, especially the enrichment of content.

The SAGW is very active concerning international cooperation in the fields of infrastructures and initiatives for digital SSH-research. It is member of the AG Datenzentren des DHd (Working Group Data Centres of the German language based Digital Humanities association “Digitale Geisteswissenschaften im deutschsprachigen Raum”), it is as an institution part of the DARIAH ERIC (Switzerland is not a member country), Europeana data provider, and Beat Immenhauser is an active member of the ALLEA-Working Group E-Humanities. Many of the research projects at the individual institutes cooperate (inter-)nationally as well, for example in portals (or gateways) such as the Biographie-Portal (<http://www.biographie-portal.eu/>).

Immenhauser highlighted that at the SAGW all digital born data are open access by default, but this is not the case with “traditional” editions which are published as printed volumes. He discerned four dissemination scenarios for research output at the individual research institutes: 1) exclusively printed output (traditional text editions), 2) balanced hybrid output (dictionaries, text editions), 3) primacy of digital output, but some printed output (historical source editions), and 4) exclusively digital output (databases), and exclaimed that the first model has in his opinion no future. Concerning digital tools, Immenhauser explained that the individual research institutes are independently responsible to cover their needs, but that DH-tools are exchanged if possible. As common challenges for all research institutes of the SAGW Immenhauser identified data linkage and data preservation and elaborated on the central institutional solutions offered and developed at respectively in cooperation with the SAGW: metagrid.ch (a web service for data linkage, which is also used by several international partners) and DaSCH (Data and Service Center for the Humanities). DaSCH is conceived as technical backup for the academy to handle the digital turn, but it is also open to others. DaSCH is not only a Humanities Data Centre, but also a kind of DH centre in general.

In the last years the SAGW has assumed a leading role in the digital turn. Before this, the initiative has solely been on the side of the institutes and some bottom-up initiatives such as histHub. Immenhauser underlined that the common challenges of the digital turn require a more coordinated approach to prevent major technology gaps. Therefore, the SAGW will soon also provide guidelines for data management and maintenance. Through their network activities the SAGW is quite well informed concerning the needs at Swiss institutes and researchers, but there is still more need for international exchange, as the Digital Humanities are a quite young discipline in Switzerland.

Milena Žic-Fuchs (HAZU): The Croatian Academy of Sciences and Arts and Digitalization

Slides

MILENA ŽIC-FUCHS first gave a short impression of the structure of the Croatian Academy of Sciences and Arts. Although the Croatian Academy is primarily a learned society, it has 35 research units mainly in SSH under its roof. At the moment a strategy document is being finalized to set the course for the digital turn at the Croatian Academy. For this paper several consultations were conducted at archives, libraries, and research units. One of the findings was that in general there is a high acceptance of the digitisation process (90%), but that the concrete implications of this process are not clear to all consulted parties. However, the need for the digitisation of documents, collections, and archival data was stressed by many parties, because many of the research units have invaluable documents that should be brought to the eyes of the public, but this will be a huge task for the future. The strategy paper does not refer to scientific publications, because a national solution is already available with HRČAK (Portal of Scientific Journals of Croatia). HRČAK is the central portal of Croatian scientific journals and offers access following the Open Access initiative. In Croatia therefore Open Access is well established for journals however, for monographs which are still the gold standard in the humanities, the acceptance is much lower.

There is already some international cooperation between the Croatian Academy or its individual research units in the field of SSH infrastructures. The Central Library of the Croatian Academy is a Europeana data provider, some research units have expressed their wish to be included in DARIAH-HR (Croatia is not member of CLARIN). However, on the level of the individual researchers it seems not always to be clear what the objectives of the Pan-European SSH infrastructures (such as DARIAH or Europeana) are, they are more or less the same to them.

Žic Fuchs concluded her talk with four main challenges for the Croatian Academy: 1) the need to educate the researchers on Research Infrastructures, 2) the need for a higher level of coordination and collaboration between the research units themselves, as well as their connections with the Croatian Academy Library and Archives, 3) the need to raise the awareness of the necessity for ensuring visibility in all aspects of SSH.

Karlheinz Mörth (ÖAW, ACDH): “Austrian Academy of Sciences: Digital Research Practices and Publications”

Slides

KARLHEINZ MÖRTH first introduced the Austrian Academy of Sciences (ÖAW) and his home institute, the Austrian Center for Digital Humanities (ACDH). The ÖAW is a national academy and Austria’s central non-university research and science institution. He described the dissemination of new insights with the aim of contributing to progress in science and society as a whole as being one of the main tasks of the academy. The research conducted at the 28 research institutes of the ÖAW covers equally the arts and humanities and the social and natural sciences. Still, regarding funding the arts and humanities are under pressure. Besides the ACDH, also the ITA (Institute of Technology Assessment) is interdisciplinary in the special way that it belongs to the sciences as well as the arts and humanities.

Although the ÖAW is not mainly a funding body with *go!digital* it runs a program for promoting digital research practices in the SSH. The projects funded by this program become automatically part of CLARIN/DARIAH Austria (as in-kind contribution). Go!digital is only one of the many activities which is conducted under the umbrella of the recently founded ACDH, which consists at the moment

of four working groups and a young international team of approx. 45 researchers. The ACDH's main objective is to support humanities by fostering digital methods, which means that the humanities are at the focus and digital methods are nowadays an integral part. Important aspects of the ACDHs mission are: 1) building up digital infrastructures, 2) developing technical tools and services, 3) hosting and publishing digital data, and 4) intensifying knowledge transfer. Mörth underlined that the ACDH's vision of Digital Humanities is not only technical, but that social aspects play a huge role. Another central aspect is the dissemination of research results also reaching beyond academe to make SSH visible (citizen science, collaboration, interdisciplinary research). Therefore, not only Open Access is promoted actively, but even Open Science (which goes further than Open Access and includes free access to research data (results and data)).

The ÖAW and especially the ACDH are part of many national and international project collaborations (e.g. Parthenos or dariahTeach). CLARIN and DARIAH cooperate on a national Austrian level as CLARIAH-AT which is coordinated by the ÖAW at the ACDH. Concerning Austrian DH-infrastructure Mörth explained that re3data.org shows 23 repositories for research data, that are quite heterogeneous. Two are certified with the DSO, six support persistent identifiers, and seventeen have publishing policies. Of these repositories he highlighted GAMS (University of Graz), PHAIDRA (University of Vienna), and the CLARIN Centre Vienna (ÖAW) and the institutional repository of the ÖAW, EPUB.OEW, which can also deal with xml-files.

Summary of the discussion:

In the subsequent discussion the question arose how the speakers would understand the terminus technicus of "digitisation" and whether it refers to access to metadata, data in general or machine readable texts. The speakers underlined that digitisation means much more than just putting books on a scanner (Mörth), but contains a whole range of topics including DMP, and that data - given to the special situation in the humanities range from metadata, texts and pictures, to the kind of research data similar to the sciences (Monok). When speaking with researchers it becomes clear that everybody understands something different (Zic Fuchs).

The chair of the session LAURENT ROMARY pointed at the lessons learned from the projects Cendari and EHRI that aimed at the digitisation of archive material and where data of all kinds of flavours had to be integrated. Both projects were confronted with the question of how to guarantee a continuum (sustainability) after the end of the main funding period, because the data collection process was not automated, which posed the problem how to "refresh" the content. Romary identified the heterogeneity of the digitisation process and the importance to investigate how others dealt with this problem as one major challenge for AGATE.

The participants of the workshop then returned to the point made by Mörth concerning the right toolbox for digital humanities research and wanted to know how it is possible to decide what is the right tool. Mörth explained that this questions surmounts the level of tools, but touches also the question of standards, and that it can be only answered in constant dialogue with the researchers.

Then, answering the question if the concept of Citizen Science has already reached the academies and in which ways research results are disseminated to the broader public, Mörth referred to the PR-department of the ÖAW, which takes an active role in communicating research results to the broader public, not only to broaden the dissemination, but also to avoid misunderstandings and by pointing at relevant topics. Immenhauser and Zic Fuchs commented that at their academies more traditional dissemination channels are used (conferences, publications for broader public) and that the concept of Citizen Science does not play a role. It was subsequently argued that the digitisation process respectively the digital turn of the SSH should be more research driven. Because this development

seems at the moment mainly driven by the SSH-infrastructures, the academies need to ask themselves which role they could play. Immenhauser underlined that the SAGW as an institution needs a digital strategy which should be developed together with the individual institutes and researchers, and that a digital agenda is also an important PR tool. In general, the academies should play a more active role. Mörth added that communication with researchers and infrastructures is very important to decide on the strategy, because it is so difficult to foresee developments, even for the next 5 years. Zic Fuchs concluded that AGATE should incorporate into the wider landscape and become a driving force to get the "wealth" of the academies on board, while keeping in mind the different levels of the individual research institutes and academies. Thus one objective should be the integration within the academies' SSH as a community of practice, and the link to the other infrastructures.

Romary concluded the morning session with the remark that the objectives of AGATE surpass building a technical platform, but that education and awareness rising concerning digital methods and publication practices in SSH amongst the academies should play a huge part.

Afternoon Panel: Afternoon Session: Panel Discussion with Speakers and Workshop Participants: A European Academies Internet Gateway – Chances and Challenges

CHAIR: Eveline Wandl-Vogt

Alexander Czmiel (BBAW): The Digital Knowledge Store: “Digital Infrastructures and Data Curation for Humanities Scholarship”

Slides

ALEXANDER CZMIEL started this presentation with an overview of the tasks and achievements of TELOTA (The Electronic Life of the Academy), the Digital Humanities initiative of the BBAW since 2001, which supports the whole lifecycle of BBAW-projects from planning the digital part of projects, the development of DH-tools, to the publication of digital research results. TELOTA is an active part of the DH-community and has developed more than 40 projects until today.

The aim of the TELOTA-project “Digital Knowledge Store (DKS)” (Digitaler Wissensspeicher, 4 researchers, 4 student assistants) is to provide centralized access to all digital resources of the BBAW to increase their visibility, because these resources are only relevant if they can be seen. Therefore all this diverse material needed to be structured and connected semantically. The DKS does not store the data, but "harvest the projects" by connecting metadata manually and semi-manually. This approach differs from Google, because it can go "inside" the data/databases, even if general access to the content is restricted. Czmiel pointed at named entity recognition as an area for further development for the DKS as well as a metadata management tool, which is currently being developed, based on the experience that everything has to be done manually and which in the future should be also usable for others. As further areas for improvement he stressed the relevance of the search results.

For now, the DKS provides access to 1.267.512 records from 175 projects while its metadata can be harvested automatically via OAI-MHP. The DKS has a user interface which allows “knowledge browsing”, in contrast to a simple search interface. This means that the user can explore semantically

connected projects. Advantages of the approach chosen by the DKS are that it is possible “to look” into the structure of the data and that it provides a (re)-search entrance to curated scholarly knowledge.

Czmiel concluded his presentation with a list of tasks including finding solutions to keep old projects alive and to implement more standardisation in several areas, for example by establishing data and project description formats and standard data and metadata exchange interfaces, and further developing the visualisation of the data.

Questions and Answers

The discussion first took up the question who is responsible for the sustainability of the URL's. Czmiel explained that the DKS does not directly show the resources themselves, but only points at them and that in fact the project itself is responsible for the stability of the URL. He identified guaranteeing the sustainability of the digital resources as a common task for the academies. The next question evolved around the users of the DKS. Czmiel explained that for these kind of infrastructures it is rather difficult to say something concrete about the users, but in case of the DKS he would discern three kinds of users: 1) the very broad public as well as researchers, 2) the projects that provide the data, 3) others institutions, etc., who wish to reuse the system developed by the DKS. He added that there are plans to develop a solution to track and scientifically analyse the user interactions and to allow users to make annotations to their search results. For AGATE it was pointed out from this part of the discussion that it is very important to clarify who the intended users are (broader public, researchers, ...) and which value AGATE adds compared to already existing solutions (Use cases). It has to be more clearly worked out that the objective and added value of AGATE - like the DKS - is much more than a general dissemination task. AGATE could stimulate more interoperability, standardisation, and sustainability, because new projects will have to plan their project specific solutions with an eye on interoperability with a cross-platform solution, which is an intrinsic part of the publication process. The last point of the discussion evolved around the feasibility to scale the DKS-approach up to a European level, which was considered worth further investigation, especially concerning the estimation of the resources needed for the central implementation, but also at the individual academies, respectively projects for implementing interfaces for automated exchange, workflows and documentation, and solving issues concerning multilingualism. The tenor was however positive as AGATE emerges at a time, where many lessons are learned and there is a lot of experience around. One possibility would be to develop a prototype to present lighthouse projects such as a cluster of networked dictionary resources to show the extra value of the concept as well as to have a proof of concept, offering also a possibility to track what users are doing with the resource and to learn how to improve it.

Miha Seručnik (ZRC SAZU): Projects of the Research Centre of the Slovenian Academy of Sciences and Arts

Slides

First MIHA SERUČNIK gave a general overview about his home institution, the Research Center of the Slovenian Academy of Sciences and Arts to set the scene. The ZRC SAZU was founded in 1981 and links 18 institutes in the sciences and humanities. Nowadays it is one of the leading research and educational centres in Slovenia. The ZRC SAZU conducts national research programmes, national and international research projects, applied fundamental research and acts as publisher of journals and book series. Current and completed programs and projects at individual institutes of SRC SAZU can be consulted via a central database: <http://www.zrc-sazu.si/en/programi-in-projekti>.

Seručník then focussed on several humanities projects to illustrate the different kinds of digital result output occurring at individual institutes. Access to archives and collections via web interfaces, for example to textual sources (Archaeological Cadastre of Slovenia, ARKAS), video and photo archive catalogues (Fototeka ISN), and linguistic resources (Fran: Dictionary and Text Corpora of the Slovenian Language). The resources in Fran are dictionaries and corpora as a combination of digital born materials and retrodigitisation, also the combination of dictionaries and corpora is interesting. He then introduced the Etnofolk project (<http://www.etnofolk.eu/en>), an international cooperation under the lead of the Institute of Ethnology of the Academy of Sciences of the Czech Republic, which provides a unique source of information about folk culture in Central Europe to researchers, teachers, students, museum employees, and local and regional administration representatives. The aim of this project is to integrate the digital resources in a machine readable way. Another international cooperation is the Slovenian biography which is part of the Biographie-Portal, a cooperative project of the Bavarian State Library, the Historical Committee at the Bavarian Academy of Sciences and Humanities, the Austrian Academy of Sciences, and the Foundation Historical Dictionary of Switzerland. For the use of TEI Seručnik highlighted the project “Sermon of bishop Slomšek”.

Seručník pinpointed that a considerable amount of projects has no digital or publicly accessible digital output. Therefore, it is planned to develop a repository (based on Fedora) which is mainly intended for publications, but prepared for datasets. This would prepare ZRC SAZU to handle long-term preservation and the publishing of research data, a field in which rethinking is taking place at the moment in Slovenia. As central access portal to the digital resources of ZRC SAZU he introduced Arzenal (Virtual treasury of national heritage), a project completed in 2010, which has been designed and constructed as a web model and which seeks to provide search and browse content in a uniform manner. A new project underway is the SI-DIH aggregator/search tool (OAI-MHP) which is part of the DARIAH Slovenia (SI-DIH) portal. At the moment more work is needed to win new content providers for the portal. He concluded that although there are a broad variety of DH initiatives under the roof of ZRC SAZU, there is no central strategy yet.

Questions and Answers

The discussion evolved first around the different levels of access meta search engines/portals are facing which leads to very heterogeneous results. It is necessary to decide where the focus is, either deep level linking or broad picture. It was exemplified that for example for the Digital Knowledge store it is not necessary to rewrite the viewer for complex digital content, because the resource is used via the specific viewer built for the project that means in its original environment which was built with a certain user in mind. One question concerned the difficulties to win new content providers for the SI-DIH portal. Seručnik explained that for example museums are not so much

interested to adapt their data to the project's needs, simply because of a lack of resources and funding. Other contributors stressed that the academies because of their structure are quite special concerning their digitisation needs, therefore it is also a psychological problem to get them going, and that it is important to keep libraries and archives in mind (concerning standards) when building AGATE.

Elena González-Blanco García (LINHD): Training, Information and Service Hub

Slides

ELENA GONZÁLEZ-BLANCO GARCIA began her talk with the statement that the problems that had been discussed so far are broadly relevant. She illustrated this with the specific situation of digital research at the academies in Spain where the scene is quite scattered and where also often a language problem plays a role, many of the academies' websites are not in English. As an example she referred to the Instituto de España (Spanish Institute) of which 9 academies are part. At the moment it is not possible to get a central overview of all information about the kind of research activities which are going on at these academies. Concerning the technical infrastructure there are several handicaps: the very low state of internationalisation, the absence of standards which leads to fragmentation, a lack of experience with e-research and infrastructures.

She identified three steps of a necessary digital innovation process to facilitate and promote sustainable digital research and publication practices in the context of long-term academies' SSH research which could be facilitated by AGATE:

- 1) Identifying and discovering the project's needs, such as information about similar projects, shared problems and answers, tools that can be adapted, expert advice and consulting (at least provide a first answer or liaise with an expert).
- 2) Be a driving force for the implementation of the change, which means in the context of digital transformation; promote open standards, collaborative and interdisciplinary research, and the training of researchers, developers and common users. Therefore it is very important to offer tutorials on TEI, and other digital good practises. The main focus of the development of the service and information hub has to be on the user interface (easy navigation (less text more icons), multilingual, UX-IU+ design). While developing the user interface and choosing the fields of information the different types of users have to always be kept in mind, because they will only return to such a service, if they are content.
- 3) Making it sustainable: Interoperability between different projects, work on a common preservation e-infrastructure strategy, and use open code, open access, open communication strategy, helped by community support.

González-Blanco Garcia then continued to sketch the ideal information and service hub for AGATE with mainly two types of users in mind: researchers and the broad public. For researchers she underlined the need for collaborative team-working spaces with unified login systems, a long-term repository to upload and tag contents with permanent ID-systems and URIs, and semantic-based standard metadata system to classify content (using existing ontologies). For the public at large she gave prominence to a semantic based search engine (e.g. Isidore) and underlined the importance to draft a concrete communication (and dissemination) plan to link "AGATE" to the academies' world and beyond (blog, newsletter, and social networks. Because the interaction with users is so important it should be considered to provide mailing lists for different topics, for example metadata.

The service and information hub will be an important aspect of AGATE, but it is highly necessary to develop it with an eye on what other infrastructures have already developed and are using, in short “take the best out of the existing solutions” and to collaborate on content.

Questions and Answers

The participants agreed that it would be highly desirable to take the best from already existing materials to reach a win-win situation especially with the relevant national, community-driven, or European Infrastructures, but that the academies via AGATE should also consider more focussed what they can contribute to broader ongoing discussions. Other points were made concerning the necessity to provide material in several languages, and that a mixed approach between live and online offers aimed at different kinds of user would be ideal (e.g. a summer school on dictionary research in European perspective, a wiki, but also material developed by projects, for example for educational purposes). UNED and LINHD are a good example how distant teaching can reach physical and virtual participants. The advantage is that people who cannot attend live, can at least have advantage from all the materials (videos, slides, etc.), this is last but not least an aspect of sustainability.

Francesca Morselli (KNAW/DANS): Chances and Challenges of a European Academies Internet Gateway

Slides

First FRANCESCA MORSELLI gave an overview of the activities of DANS (Data and Archiving and Networked Services), an institute of the Royal Dutch Academy of Sciences (KNAW), that is specialized in promoting and providing permanent access to digital research information. DANS is permanently funded since 2005 by the KNAW and the Dutch national research foundation (NWO) and can resort to 50 years of experience in this field. The core services of DANS are EASY (a long-term Electronic Archiving System for self-deposit), NARCIS (the gateway to scholarly information in the Netherlands), and DataverseNL (a system for short/intermediate-term storage). Amongst the additional services are persid (an initiative to establish an infrastructure for Persistent Identifiers), training and consultancy, and the publication of the journal *Research Data Journal* for the Humanities and Social Sciences. DANS is part of national and international digital SSH research networks and Infrastructures, such as CLARIAH (NL), the eHumanities Group of the KNAW, DARIAH, CLARIN, CESSDA, EHRI (European Holocaust Research Infrastructure), or ARIADNE (Digital Infrastructures for Archaeological Research).

Concerning the feasibility of the future AGATE, Morselli pointed out that is important to investigate first thoroughly the landscape of existing research infrastructures in the arts and humanities as well as the risks and possibilities. In this context, she highlighted OpenAire, NARCIS, and the DODH project registry. Next she discussed the focus of AGATE. It has to be made clear where the focus of AGATE is, or which parts of the concept are aimed at which user respectively purposes: Is the main focus the support of scholarly practices of a network of academies that promote their SSH-research. She pinpointed at the risk of fragmentation of the infrastructure landscape and the need to keep an special eye of interoperability, the possibilities of LOD, and the need to fall back to existing standards such as CERIF. As lesson learned from CENDARI, she underlined that from the beginning a sustainability plan (for tools, data, repository, funding) should be an integral part of the concept. All in all, AGATE should learn from the current research infrastructures and define where - concerning its focus and end-users - is room for development and clearly flesh out the opportunities as well as the risks and to collaborate and integrate available solutions from the start.

Questions and Answers

A huge part of the following discussion was dedicated to questions how to deal with the heterogeneous quality of the resources, concerning their accessibility, metadata, etc. In this context it was suggested that more efforts would be needed to encourage the standardisation of the metadata of the academies' repositories for example by developing AGATE-guidelines and to map acceptable metadata formats in order to improve automatic harvesting. The more harvesting can be done automatically and the more the stability of access points is ensured (by using for example URNs) the less is the risk of lots of broken links which increases the sustainability. A portal like AGATE could in fact function as an incentive (the metaphorical carrot): if you standardise (and or built or adapt access points) your digital resources will be more visible, in AGATE and beyond. This would actually be win-win situation: more academies' repositories could be for example harvested by OpenAire, but because the individual academies' resources would be somehow drowning in the abundance of material accessible via OpenAire, it would be advantageous if they would also be findable via the specialised AGATE-portal (as curated content of a community).

Under the umbrella of the European Science Academies there is a big and specific SSH community which still in its majority has to fully grasp the chances and challenges of the digital turn. Therefore, the risk of fragmentation by establishing AGATE as a new infrastructure has to be considered in relationship to the chances of a more channelled communication between the academies' SSH-researchers and the identification, implementation and development of standards and services in close cooperation/discussion between researchers and SSH-infrastructures. While the European SSH-infrastructures will play a huge role in these processes, the academies on a national level should be encouraged to work more closely with libraries. Libraries as some of the oldest information infrastructures should be the natural partners for questions concerning digital preservation in general and to ensure the findability, accessibility, and citability of digital resources. Last but not least the question of intellectual property rights was discussed and agreed that given the complex and diverse national situation, a pan-European infrastructure like AGATE has to pay extra attention and can learn a lot from other pan-European infrastructures, therefore this topic needs to be an integral part of the concept.

Résumé of the day and prospects: Eveline Wandl-Vogt and Ulrike Wuttke

At the end of this very fruitful day EVELINE WANDL-VOGT and ULRIKE WUTTKE gave their perspective on the presentations and discussions and gave an outlook to the future. They underlined that it has to be understood that AGATE is not just a technical infrastructure, but that it becomes clear that the implementation of AGATE will also require a change process that has to be approached with methods of change management. This process will need time, diplomacy, and persuasive power because the academies' world is rather closed and hierarchical and also very heterogeneous. It also becomes obvious that AGATE may focus on three groups of users: researchers, academies' management, and the broad public. For the researchers not only access to the curated resources is relevant, but also the identification of similar topics and or digital practices to foster the exchange of knowledge and unearth potentials for cooperation. Therefore it is important to include possibilities to add this kind of information in a structured way into the technical concept.

Although the relevance of open access and open-science as underlying principles of the objectives of AGATE were discussed by the participants in the broader context of dissemination, the moderators mentioned as missing in the discussion a more thorough consideration of the place of AGATE in relation to the reward-system for the researchers and how it could act as a trigger to full heartedly

embrace a culture of sharing. In this context also economic opportunities for the ERA which could arise from business companies using the academies' data - an untapped potential - have to be considered (and which feature prominently on the EC agenda). While conceptualising AGATE also aspects of crowdsourcing and co-creation (involvement of the users in the development of the infrastructure) have to be kept in mind. One idea would be to form small decentralised groups of expertise, to get the process started.

The following points were identified as central aspects of the workshop's discussions concerning the AGATE-concept: the content of AGATE (what and how, aspects to be considered contain standards, interoperability, and the long-term perspective). Last but not least, the challenge of permanent data curation was underlined by the moderators. There is a lack of stable institutions dedicated to the long-term accessibility of complex digital objects resp. enhanced publications (living data), but also file based storage. If this problem is not solved in the long term according to the academies' needs, sooner or later, this will lead to a lot of broken links (not only) in AGATE and loss of scientific research results.

— The moderators concluded that AGATE offers a great opportunity to make the academies SSH-research generally more visible and to improve the findability and accessibility of the digital resources in order to underline and improve their relevance. Increased relevance makes not only a stronger case to fund new projects in these fields, but to also take dedicated actions and allocate money to preserve the results as part of our digital heritage. Finally the moderators underlined that AGATE will offer a platform to foster knowledge exchange and cooperation in order to improve the sustainability of digital research and publication practices at the European science academies.

Contact:

Dr Ulrike Wuttke

Scientific coordinator of the project

Union der deutschen Akademien der Wissenschaften

c/o Berlin Brandenburgische Akademie der Wissenschaften

Jägerstraße 22/23

10117 Berlin

Tel.: 030-20370369

E-mail: wuttke@akademienunion-berlin.de

Web:

<http://www.akademienunion.de/en/working-groups/cooperation-projects/agate>