



K O N I N K L I J K E N E D E R L A N D S E  
A K A D E M I E V A N W E T E N S C H A P P E N

**Teaching Digital Humanities Around the World: An Infrastructural Approach to a Community-Driven DH Course Registry**

Wissik, Tanja; Edmond, Jennifer; Fischer, Frank; de Jong, Franciska; Scagliola, Stef; Scharnhorst, Andrea; Schmeer, Hendrik; Scholger, Walter; Wessels, Leon

2020

**document version**

Peer reviewed version

[Link to publication in KNAW Research Portal](#)

**citation for published version (APA)**

Wissik, T., Edmond, J., Fischer, F., de Jong, F., Scagliola, S., Scharnhorst, A., Schmeer, H., Scholger, W., & Wessels, L. (2020). *Teaching Digital Humanities Around the World: An Infrastructural Approach to a Community-Driven DH Course Registry*. HAL Open Archive . <https://hal.archives-ouvertes.fr/hal-02500871v1>

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the KNAW public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the KNAW public portal.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

**E-mail address:**

[pure@knaw.nl](mailto:pure@knaw.nl)

# Teaching Digital Humanities Around the World: An Infrastructural Approach to a Community-Driven DH Course Registry

Tanja Wissik<sup>1</sup>, Jennifer Edmond<sup>2,3</sup>, Frank Fischer<sup>3,4</sup>, Franciska de Jong<sup>6,10</sup>, Stefania Scagliola<sup>7</sup>, Andrea Scharnhorst<sup>3,8</sup>, Hendrik Schmeer<sup>6,1</sup>, Walter Scholger<sup>9</sup>, Leon Wessels<sup>5,6</sup>

<sup>1</sup>Austrian Academy of Sciences (Austria), <sup>2</sup>Trinity College Dublin (Ireland), <sup>3</sup>DARIAH-EU, <sup>4</sup>National Research University Higher School of Economics (Russia), <sup>5</sup>Utrecht University (The Netherlands), <sup>6</sup>CLARIN ERIC, <sup>7</sup>University of Luxembourg (Luxembourg), <sup>8</sup>Royal Netherlands Academy of Arts and Sciences (The Netherlands), <sup>9</sup>University of Graz (Austria), <sup>10</sup>Erasmus University Rotterdam (The Netherlands).

## Abstract

The digital humanities (DH) enrich the traditional fields of the humanities with new practices, approaches and methods. Since the turn of the millennium, the necessary skills to realise these new possibilities have been taught in summer schools, workshops and other alternative formats. In the meantime, a growing number of Bachelor's and Master's programmes in digital humanities have been launched worldwide. The DH Course Registry, which is the focus of this article, was created to provide an overview of the growing range of courses on offer worldwide. Its mission is to gather the rich offerings of different courses and to provide an up-to-date picture of the teaching and training opportunities in the field of DH. The article provides a general introduction to this emerging area of research and introduces the two European infrastructures CLARIN and DARIAH, which jointly operate the DH Course Registry. A short history of the Registry is accompanied by a description of the data model and the data curation workflow. Current data, available through the API of the Registry, is evaluated to quantitatively map the international landscape of DH teaching.

## Keywords

Digital Humanities, Teaching, Education, Registry, Research Infrastructure

## 1 Introduction

The use of digital methods in humanities research, commonly referred to with the umbrella term 'digital humanities' or, more simply, 'DH', is increasingly becoming a more general and accepted approach to knowledge creation in these disciplines. From the outset, however, the new conditions from which the digital humanities emerged (which required such things as specialised software, new standards, and different methodological approaches) gave rise to requirements for new skill sets, and the invention of alternative pedagogies to deliver them. These emergent practices have led to a bifurcation between programmes being developed specifically to support the acquisition of digital methods for humanities research and existing, more traditional programmes in fields such as history or modern languages. The digital humanities have consolidated and expanded since the field first emerged, in terms of research as well as training, and students can now choose worldwide from digital humanities Bachelor's and Master's programmes with different focus areas, as well as summer schools, workshop series and other parallel forms of education. This expansion has greatly increased the number and variety of

programmes available, but it has also made it increasingly difficult to maintain an overview, or to feel confident as a potential student of DH that one has found the optimal programme for one's needs. One response to this challenge has been the DARIAH/CLARIN DH Course Registry, for which this article describes the context, history, development and features, as well as drawing some conclusions from the information it gives access to.

## 2 The Digital Humanities: A Brief Overview

These days it is hard to picture an office, or any work environment where data is processed, without computers. Universities are no exception. We use email to communicate with colleagues, text editors to write papers and web browsers to find information. Yet some researchers also use their computers to access, analyse and visualise their research data. When this occurs in the fields of the humanities, this activity is, as stated in the introduction to this article, generally referred to as the *digital humanities* (DH). However, the existence of this single umbrella term word does not mean that we are talking about a homogeneous methodological framework or community of researchers. Instead of describing DH as a field in the conventional sense, it is better to speak of "an array of convergent practices" (Digital Humanities Manifesto 2.0., 2009), which is situated between ICT/computer/information science, collection providers and cultural heritage institutions, and the researchers working with this content (Akdag Salah et al., 2015). Studies and projects by historians, linguists, musicologists, literary scientists and other humanities scholars all have their place in DH, as well as contributions from technologists and librarians working with digital data.

So, what is it DH scholars or *digital humanists* do? What research questions do they try to answer and why do they use computational methods for that? Although the various disciplines involved each have their own specific focus, some common approaches can be distinguished. First of all, DH researchers tend to work with datasets that are larger than a corpus typically used for traditional research projects in the humanities. Google Books (<https://books.google.com>) is a widely used online corpus that consists of several million books, more than any individual person could read in their lifetime. Using an approach called 'distant reading,' a term coined by literary scholar Franco Moretti, it is possible to analyse such large bodies of text in a short amount of time in order to study linguistic changes and cultural phenomena (Michel et al., 2011; Moretti, 2013) at a macro scale. In this way, DH allows scholars to address big questions using big data.

Clearly, one of the main drivers for the advancement of DH has been the vast number of freely accessible, online collections of cultural heritage data. Researchers, students and citizen-scientists owe a debt of gratitude to galleries, libraries, archives and museums (which we will henceforward refer to as the GLAM sector) for their massive contribution to this development. Digitised or born-digital books, periodicals, manuscripts, art objects and other types of cultural heritage objects rendered as data can be accessed and analysed by people from all over the world. In addition, the GLAM sector provides DH tools and offers training (Terras, 2012; White and Gilbert, 2016). In Europe, initiatives from the GLAM sector to provide digital access to cultural heritage and support DH research, are stimulated by the Association of European Research Libraries or LIBER (<https://libereurope.eu>) and Europeana, a platform for digital cultural heritage (<https://www.europeana.eu/>). But DH is not just about large numbers. Traditional hermeneutic inquiry still plays a vital role in DH research. Many digital humanists constantly switch between levels of analysis, ranging from the macro (big data) to the micro level (the smallest unit of study),

combining distant reading with close reading (Budrick et al., 2012). To interpret trends at these different levels, it is crucial to have an understanding of the sources themselves. Source criticism is as important for DH as it is for the humanities at large. Over the last few years, digital humanists have developed several approaches for critical use of sources and tools (inter alia Drucker, 2011; Wettlaufer, 2016; Koolen et al., 2018).

DH is a relatively young field of study, yet there are clear signs that it is maturing rapidly. The number of humanities scholars using computational methods to collect, analyse and visualize data is growing. Many universities across the world have established their own DH centres, forging collaborations between humanists and technologists. International DH projects facilitate cross-national collaboration and research, and receive significant financial support from research funding agencies at national and European levels. Finally, DH has founded its own journals and conferences to share results and best practices (Kirschenbaum, 2012). Paradoxically, the success of DH poses a number of challenges, however. How can researchers keep track of the vast amount of data made available? How can the adoption of common standards be stimulated? How can DH be integrated in the curriculum? How can access to the tools and knowledge required to design and manage a DH project be shared beyond the teams of those lucky few able to secure significant project funding? Research infrastructures active in supporting the adoption of digital methods in the humanities are seeking to address these challenges.

### **3 CLARIN and DARIAH: Two European Research Infrastructures**

Research infrastructures offer resources and services to the scientific community. They stimulate international, cross-disciplinary collaboration, promote the use of common methods and standards, and educate new generations of scholars.<sup>1</sup> In Europe, the European Strategy Forum on Research Infrastructures (ESFRI) recognises over fifty different research infrastructures covering a wide range of academic disciplines, including biomedical science, astrophysics, heritage science, and many others. In this section we will focus on two such organisations recognised by ESFRI as established ‘landmarks’ serving the humanities and social sciences: the *Common Language and Technology Infrastructure* (or CLARIN, <https://www.clarin.eu>) and the *Digital Research Infrastructure for the Arts and Humanities* (or DARIAH, <https://www.dariah.eu/>) (ESFRI, 2018: 12).

CLARIN and DARIAH have distinct and complementary objectives underpinning their infrastructure models. CLARIN’s mission is to create and maintain an infrastructure to support the sharing, use and sustainability of language data and tools for research in the humanities and social sciences. DARIAH aims to enhance and support digitally-enabled research across the arts and humanities. Both CLARIN and DARIAH are European Research Infrastructure Consortia (or ERICs), a form of legal entity created in order to allow pan-European initiatives like CLARIN and DARIAH to be formalised on a non-economic basis. (EC, 2019) They are independent organisations, with their own governing bodies, primarily funded through the fees paid by their respective member-countries.

As noted above, over the last decades vast numbers of digital collections and tools have been created for DH research. While this abundance enables new ways of doing research, it also

<sup>1</sup> For a general introduction to research infrastructures for the humanities, see: <http://training.parthenos-project.eu/sample-page/intro-to-ri/> .

brings a number of challenges. How can all these data be stored in a sustainable way? How can researchers find the collections and tools relevant for their research? Research infrastructures offer solutions for these challenges by hosting data repositories to make collections and tools sustainable (Edmond, 2015) and providing platforms for finding and sharing data. To mention but two examples, CLARIN recently launched an initiative to provide user-friendly overviews of so-called *Resource families* throughout Europe, such as parliamentary records, newspaper corpora, corpora of spoken language, second language learner corpora, and several others. Resource families are selected based on their availability in multiple languages and their relevance for a wide range of disciplines and methodological approaches and therefore offer lots of potential for cross-disciplinary and cross-border research. The overview page for each resource family<sup>2</sup> gives information on the structure of the corpus, licenses and download links (Fišer et al., 2018). DARIAH is currently constructing an open *Marketplace* for the humanities and social sciences, which will grow into a rich environment for humanities researchers to share data, tools and services (Edmond et al., forthcoming). These services stimulate the reuse of existing data and tools, which not only saves time and money but also enables researchers to study phenomena regardless of their place of residence. To make it easier for researchers to use and combine different datasets, CLARIN and DARIAH have adopted common standards for (meta)data. Both research infrastructures are also committed to promoting and exemplifying the movement toward making research data Findable, Accessible, Interoperable and Reusable (FAIR) (De Jong et al., 2018).

Facilitating opportunities for cross-disciplinary and cross-border collaboration is another key feature of research infrastructures. CLARIN and DARIAH actively stimulate this kind of collaboration by offering physical and virtual platforms where researchers from various countries and disciplines meet each other. Both research infrastructures organise conferences, workshops and other events to exchange best practices, forge collaborations and enhance the infrastructure. Within DARIAH, a number of *working groups* have been established dedicated to strategic and emerging areas such as Artificial Intelligence and Music, GeoHumanities, and Women Writers in History. These working groups form a vital link between the infrastructure and the community of users (Romary and Edmond, 2019). CLARIN maintains a *Knowledge Sharing Infrastructure*, with knowledge centres specialised in Speech Analysis, Data Management, Digitisation, and many other fields of expertise.

In short, CLARIN and DARIAH support DH by offering access to data collections and tools and the means to share them by providing opportunities for both collaboration and individual development. This whole endeavour relies upon the contributions, in terms of research, development, curation and teaching, carried out within the member-countries and universities, research institutes and GLAM organisations involved in CLARIN and DARIAH, yet the scientific impact of research infrastructures is larger than the sum of their parts. Their supra-national scale allows researchers to access knowledge and best practice on topics like common standards, cross-border collaboration and comparative research agendas. CLARIN and DARIAH provide an international platform for research, collaboration and teaching in DH, but in this last element it differentiates itself from other players in a diverse field by providing distinctly macro-level,

<sup>2</sup> See this webpage for the relevant links: <https://www.clarin.eu/resource-families>

infrastructural added value. To allow the DH to continue to flourish, it is crucial to educate new generations of students and researchers, but education in DH is not limited to transmitting concrete and practical skills related to data architecture, analytics, and programming, it also includes learning to think critically about these computational methods (Murphy and Smith, 2017) and apply them within the often messy contexts of humanities datasets. In order to support this transition as something distinct from either traditional humanities, librarianship or computer science training, CLARIN and DARIAH have developed a number of initiatives to support teaching in DH, which will be explained in more detail below.

#### **4 DH Teaching/Education and RIs**

Because of its integrative, interdisciplinary and collaborative nature and emergence into disciplines with very different existing methodologies, the digital humanities have, almost from their outset, had to consider a new set of challenges regarding skills acquisition. The first iteration of the longest-running DH training school, the University of Victoria's (UVic, Canada) Digital Humanities Summer Institute, occurred already in 2001 and, in that year, included an explicit track on Digital Humanities pedagogy.<sup>3</sup> The proliferation of DH training schools and workshops continued on a strong trajectory from here. The UVic Summer Institute was soon joined by the now 10-year old European Summer University in Culture and Technology at the University of Leipzig (Germany) and the Digital Humanities at Oxford Summer School (UK), and other, more formal, programmes began to appear as well. Though none of the posts or collections linked to it exist any longer, a thread on the DH Questions and Answers forum from 2011 indicates a strong growth in programmes at many levels and in many places throughout the United States of America and starting in Europe.<sup>4</sup> The growth did not seem to have been matched by a similar strong development in the level of self-reflection regarding pedagogy seen in that first UVic Institute, however. More than a decade after those first sessions in Victoria, Matt Gold still reflected that: "the digital humanities, as a field, would benefit from a more direct engagement with issues of teaching and learning than it has exhibited thus far." (Gold, 2012: 153). This growth in practice without an equally strong development of a theoretical base may be a result of the digital humanities' strong and enduring tradition of peer learning, which has been identified as the most desirable and effective way digital humanities skills are transferred (Antonjevic, 2015).

Given the definitions and descriptions of research infrastructure provided above, it may seem surprising that consideration of training and education takes place within them. And, indeed, it seems to have taken some time for this idea to embed itself in policy and practice. A 2016 survey of user needs assessments with regards to skills development carried out in the context of the European PARTHENOS projects<sup>5</sup> comes to the following conclusion: "The most striking observation is that research infrastructure projects seldom strategise or theorise explicitly about their training interventions, and how they interact with the wider environment of digital humanities."

<sup>3</sup> As per <http://dhsi.org/archive.php> selecting the year 2001.

<sup>4</sup> N.N. 2011: "Is there a list anywhere of all the graduate programs that study DH? (17 posts) (9 voices)" <http://digitalhumanities.org/answers/topic/is-there-a-list-anywhere-of-all-the-graduate-programs-that-study-dh> (accessible via [Internet Archive](#))

<sup>5</sup> PARTHENOS is a European cluster project that is bringing together research infrastructures serving the (digital) humanities to build efficiencies by sharing knowledge and practices among them. More information can be found at <http://www.parthenos-project.eu/>.

(Edmond et al., 2016) In other words, the founding attitude of the builders of research infrastructure toward training seemed to reach only as far as what a user might require to specifically and narrowly deploy the RI's own local tools, and no further. As these organisations have grown and developed their understanding of both the supportive and disruptive roles they play in their research ecosystems, however, this attitude may be shifting. In particular, research infrastructures are coming to terms with the unique role that they play in the cycle of knowledge production, working at a more applied end of the spectrum that privileges the mode of learning that Rockwell and Sinclair refer to as 'acculturation,' providing the bridging competences required to use digital humanities skills in an applied context (Rockwell and Sinclair, 2012). This mode of development leverages the RIs' emphasis on the mobility of people as a part of the richness of the knowledge landscape, in particular in niche areas. But the momentum is not only coming from the RIs themselves. In fact, a pair of recent reports on the sustainability of research infrastructures (ESFRI, 2017; European Commission, 2017) have highlighted that having the right people with the right skills in the right places is one of the key success criteria for such organisations. As such, the growing awareness among research infrastructures for the importance of developing and sustaining human capital, fostered by projects such as PARTHENOS or SSHOC<sup>6</sup>, is being met and amplified by top-down pressure from the research policy community.

The landscape features that have emerged as a result of these forces have been diverse, to say the least. Within the DARIAH/CLARIN ecosystem alone, three different platforms have been created and are actively used to support specific kinds of skills acquisition: the CLARIN VideoLectures collection (<http://videolectures.net/clarin/>) gives access to a wide variety of material, including lectures, tutorials and interviews with researchers, while #dariahTeach (<https://teach.dariah.eu>), provides a full online environment for DH courses and workshops tuned for use in university courses so as to strengthen alliances and foster innovative teaching and learning practices (Schreibman et al., 2016). In addition, the PARTHENOS Training Suite (<https://training.parthenos-project.eu/>) provides open educational resources for use by trainers and self-learners focused on competencies more common in the environment of research infrastructures, such as research data management, knowledge organisation, research questions and digital collections, and citizen science. But for all of their engagement in making the acquisition of skills accessible for new cohorts of scholars, the RIs generally do not engage in formal certification of these skills. This leaves a certain gap, for while many researchers do indeed simply want to understand their tools in order to carry out their analyses, others see the opportunities within DH as having a wider potential impact on their career paths and competency profiles, and would like a more elaborate, structured or validated pathway. Rather than reinvent themselves as universities, DARIAH and CLARIN felt they could instead contribute to addressing this need with a platform that highlights such opportunities in a manner that promotes researcher mobility and visibility across systems and countries. This was the originating vision of the Course Registry.

<sup>6</sup> The Social Sciences and Humanities Open Cloud (SSHOC) project is a European endeavor, which brings together a cluster of research infrastructures serving the social sciences and humanities to build services related to the European Open Science Cloud. More information can be found at <https://sshopencloud.eu>.

## 5 DH Course Registry

The DH Course Registry (DHCR)<sup>7</sup> started, as described below in section 5.1, as a DARIAH initiative, but has been sustained since September 2016 as a joint effort between DARIAH and CLARIN. As already explained above, increasing the visibility of DH training activities is a major concern of the DH community. The DH Course Registry was conceived to address this issue by increasing the visibility of DH teaching activities beyond the usual university networks. It was designed as a hub to collect information on digital humanities programmes, courses, lectures and summer schools, to make them available and explorable (Wissik, 2019). The main audiences or target groups of the DH Course Registry are:

- students who want to take part in a university programme in digital humanities or related fields or who want to find a student exchange or semester abroad opportunity,
- lecturers or researchers who want to promote their teaching activities,
- programme administrators who want to promote and facilitate students and staff exchange,
- researchers interested in the proliferation of DH training,
- decision makers who need quantitative evidence to support decision making about the funding of DH activities.

### 5.1 The Origins and Early Development of the Course Registry

The origins of the DH Course Registry lie in an initiative of Manfred Thaller and Patrick Sahle, who in 2011 published a list of all the course descriptions related to digital methods, linguistics, information and computer science in German-speaking countries (Sahle et al., 2011). Later on it was presented to an international audience at a [pre-conference workshop during the DH 2012 conference in Hamburg](#) (Thaller, 2012), and gradually expanded from an Excel spreadsheet in German to a multilingual interactive web resource with a interactive map of Europe (and, eventually, a map of the world) based on the available location information for the institutes offering the courses. Over the next many years, numerous individuals and institutions contributed code and expertise to the development of the Course Registry under the umbrella of the DARIAH Virtual Competency Centre for Research and Education. Finally, with financial support from CLARIAH-NL, a Dutch programme for digital research infrastructures in the humanities, the beta-version of the DH-Course Registry was presented at DARIAH's annual conference in Rome in 2014 to be taken up by the community. After several further years of convergent development, the official relaunch of the registry as a joint initiative of DARIAH and CLARIN was announced in 2017.<sup>8</sup> The next year, at a meeting of the ADHO SIG<sup>9</sup> for Digital Humanities Pedagogy during the DH2018 conference in Mexico City, the decision was taken to use the DH Course Registry as a tool to showcase DH courses worldwide. Currently the DH Course Registry features, beside courses from Europe, also courses from the United States, Australia and South Africa. The DH

<sup>7</sup> The DH Course Registry can be accessed via <https://dhcr.clarin-dariah.eu>.

<sup>8</sup> An overview of the individuals and institutes that have contributed to the evolution of the DH-Course Registry since 2014 can be found here: <https://dhcr.clarin-dariah.eu/info#credits..>

<sup>9</sup> The Alliance of Digital Organizations (ADHO) is an international umbrella organization to promote and support digital research and teaching across arts and humanities disciplines. <http://adho.org> It also invites its members to form Special Interest Groups (SIGs).



Course Registry is open to contributions from new countries, e.g., China, and we expect the list of countries to grow in the future.

## 5.2 Data Collection and Data Curation

The idea behind the DH Course Registry is that it is a community-driven effort. Therefore, data is collected not by a central organisation or individual, but contributed by teaching institutions adding their own DH-related activities. The collected information regarding the teaching activities consists of metadata that is ingested in English, regardless the language the course is taught in. Sometimes this implies that course titles and descriptions etc. have to be translated, while other metadata fields can be filled in by choosing from pick lists. Regarding the course information, the DH Course Registry aims to be as inclusive as possible, however there are very basic restrictions in place in order to keep the information in the database DH-focused and relevant for the envisaged audiences: for example the educational or training activities should fit into one of the following categories: *Bachelor Programme, Master Programme, Research Master, PhD Programme, Module, Course, Summer School, Continuing Education*. Furthermore, there should be a balance between *humanities* content and *digital* or *computational* components within the programme or course. This means that basic computer literacy courses (e.g. courses on how to use MS Office applications) are out of the scope of the DH Course Registry, as well as basic library information literacy courses (e.g. courses on how to search in a library catalogue). Furthermore, humanities courses with no digital component are likewise out of the scope for the DH Course Registry, as are programming courses or workshops that are not targeting humanities students or humanities researchers.

Usually such collections of information or databases are at high risk to become outdated very quickly, and it was challenging in the beginning to find a data curation model that would work for the Course Registry and to prove its relevance to the course providers. To realise the initial objectives of the registry required patience and endurance, rewards for collaborating, and direct contact with those course providers directly involved in upskilling lecturers and offering students study trajectories (Scagliola, 2019).

In order to assure and monitor the quality and timeliness of submitted data, the Course Registry now relies on national moderators for each country who are tasked with curating the content entered by individual lecturers or programme directors (Safradin and de Jong, 2017; Wissik, 2019). The national moderators are volunteers from the DARIAH and CLARIN communities. For each country that is represented in the Course Registry, a dedicated national moderator, or in some countries even several national moderators, is/are appointed to monitor and curate the database entries for the relevant country. In the case that no national moderator is yet in place for a specific country, the user administrators take over the task of national moderators until the appointment of a national moderator can occur. At the beginning of the academic year, or whenever the national moderator gets a notification from the system (for example, when a new course has been submitted), the national moderator assesses the course information and initiates the required updates by contacting the owner of the entry (course maintainer), which in most cases corresponds to the lecturer delivering the course or the programme director. If a course has not been updated for a long time, it gets labelled as “inactive” and will not be visible to the public, although the data is still in the database. Once it is updated,

it becomes visible again. Through these various mechanisms, the Course Registry is able to balance the needs for authoritative information with ease of use and management of the resource.

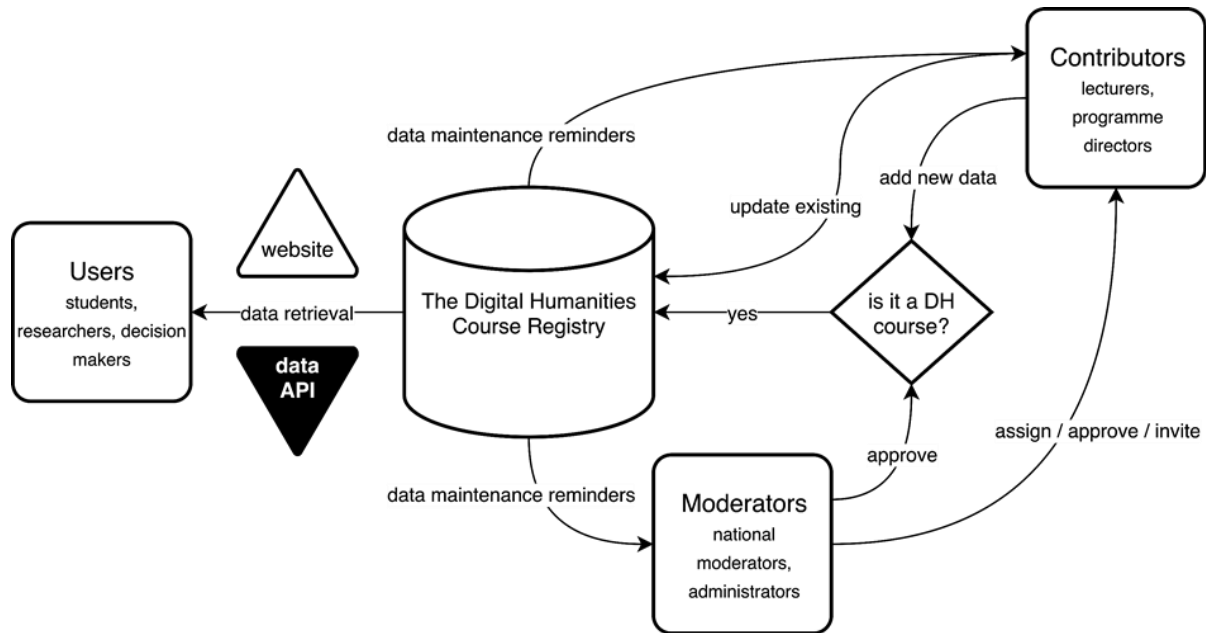


Figure 1: DH Course Registry data collection and data curation workflow

Curating the data is not the only task of national moderators, they also play a huge role in promoting the DH Course Registry, reaching out to new lecturers and new communities and inviting them to contribute to it. The DH Course Registry is open to contributions from new countries world-wide, also from countries with non-Latin scripts (for example we already list courses from Russia), since the DH Course Registry only stores the English metadata about the teaching activities and not the course materials themselves, as described above.

### 5.3 The DH Course Registry Data Model and Search Interface

The central entities of the data model are, not surprisingly, the courses (see Figure 2). All the metadata, related to the courses can be grouped into two types of metadata: metadata related to the course content (course type, course parent type, language, disciplines, objects and techniques), metadata related to the provider of the courses (institutions, cities, countries). The metadata related to the course content and the provider of the courses is the basis for the filter options that are explained later on. The entities objects and techniques come from the TaDiRAH Taxonomy (Borek et al., 2016).

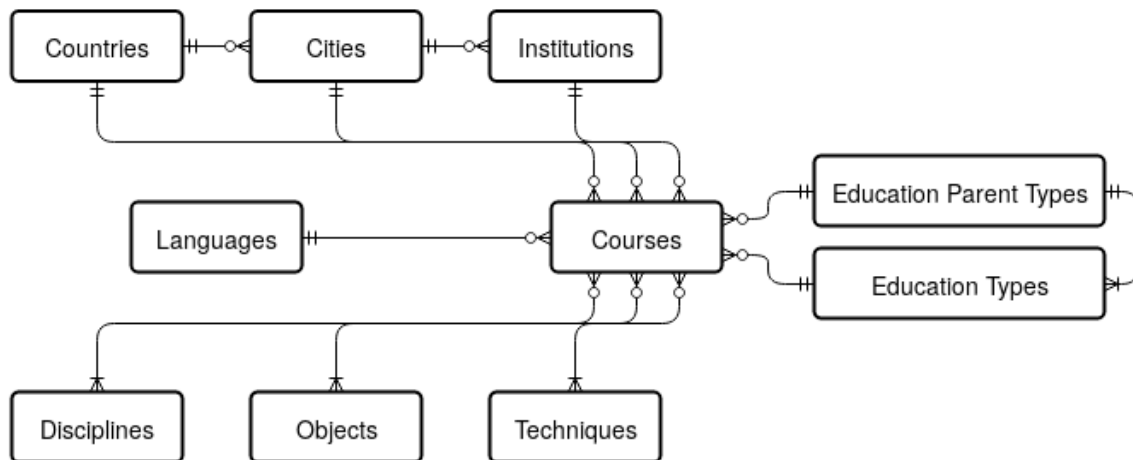


Figure 2: DH Course Registry data model

The DH Course Registry search interface presents a number of filter options to narrow down search results. It is possible to search by location (country and city) as well as institution. These search options are enhanced with an interactive map, presenting the listings also graphically by location. Furthermore, it is possible to search by educational programme type. Since the DHCR aims to collect information on DH training and education in Europe and beyond and given the heterogeneity of the education landscape, it was difficult to agree on a common terminology to classify the types of education and training activities to be included. Very general and broad activity types are therefore used as facets, as described above.

Some of these facets, such as *Bachelor Programme*, *Master Programme*, *Research Master* and *PhD Programme* are whole degree programmes, usually offered at a higher education institution. *Modules* and *Courses* are parts of such degree programmes. A *Course* includes the teaching lessons held over one semester, trimester or a year as part of a university degree programme. It could be a lecture or a seminar, for instance. A *Module* is a grouping of several courses that are thematically related. Regarding education types outside degree programmes we have *Summer Schools* and *Continuing Education*. As *Continuing Education*, we understand training events such as workshops that are not leading to a degree.

It is also possible to filter the data based on SSH disciplines. The disciplines were first either based on the disciplinary categorisation as applied by the Dutch Scientific Council for Academic Research (NWO) or NARCIS<sup>10</sup> (Safradin and de Jong, 2017). Since the DH Course Registry is a “living system”, which can be adjusted based on community input and experience” (Safradin and de Jong, 2017), several changes and adjustments have already been made, for example to the discipline filter.

<sup>10</sup> The National Academic Research and Collaborations Information System (NARCIS, <https://www.narcis.nl>) is a portal to research information in The Netherlands, hosted by Data Archiving and Networked Services, an institute of the Netherlands Royal Academy of Arts and Sciences and NWO.

By integrating the TaDiRAH Taxonomy<sup>11</sup> (Borek et al., 2016) as metadata entities into the DHCR, it became possible to browse through the data by techniques and objects as classified there, given that the (meta)data was properly filled out by the course maintainer when submitting a training activity (see Figure 2).

To sum up, the current data model and technical implementation are the basis of the data curation workflow and the filter options, which are essential for the DHCR to fulfil its goal as a community-driven registry and caters for the needs of the different target groups.

#### **5.4 Coordination of the DH Course Registry**

The community-driven approach is a strength, but does present certain challenges, especially for the coordination and maintenance of the DH Course Registry. As mentioned in section 3, DARIAH hosts a variety of working groups formed around disciplinary, methodological or audience-facing interests. One of these working groups, the WG “DH Course Registry”<sup>12</sup> is dedicated to the coordination, the user management and the technical maintenance of the registry. The working group consists of two co-chairs and all the national moderators of the DH Course Registry, as well as other researchers interested in the further development of the DH Course Registry. Besides the coordination, the working group also reaches out to the community in order to motivate them to contribute to the registry and to disseminate it. In this context, events like the Metadatathon<sup>13</sup> were organised, the DH Course Registry was presented at relevant conferences and local initiatives, such as the Call for Danish Contributions<sup>14</sup> (an initiative of the DIGHUMLAB, a national research infrastructure in Denmark, to collect Danish courses for the DH Course Registry), were launched. Such events or initiatives can be used as examples to create similar campaigns also for other countries that want to facilitate contributions to the DH Course Registry.

The working group together with the technological as well as staff support of CLARIN and DARIAH will guarantee the sustainability und further development of the DH Course Registry.

#### **6 Exemplary Data Analysis with DH Course Registry Data**

The proof of the Course Registry’s relevance and success lies, of course, in its use. In May 2019, the DH Course Registry contained 209 active courses and programmes from 27 countries.<sup>15</sup> The following countries in Europe and beyond are covered within the DH Course Registry (ordered according to the total number of training activities as of May 2019) as shown in Table 1.

<sup>11</sup> The TaDiRAH (Taxonomy of Digital Research Activities in the Humanities) can be accessed at <http://tadirah.dariah.eu/>.

<sup>12</sup> <https://www.dariah.eu/activities/working-groups/dh-course-registry/>

<sup>13</sup> [http://videlectures.net/DHcourse2018\\_paris/](http://videlectures.net/DHcourse2018_paris/)

<sup>14</sup> <https://dighumlab.org/digital-humanities-course-registry-call-for-contributions-from-denmark/>

<sup>15</sup> As mentioned above, we also archive courses not actively maintained, which are not visible when browsing the website. If we include those we count 310.

Country	Number of training activities
Germany	33
France	28
Switzerland	20
Netherlands	18
Italy	18
Austria	12
United States of America	12
United Kingdom	11
Russian Federation	9
Finland	7
Ireland	4
Greece	4
Norway	3
Spain	3
Israel	3
Hungary	3
Sweden	2
Belgium	2
Croatia	2
Slovenia	2
Lithuania	2
Czech Republic	2
Denmark	1
Latvia	1
Estonia	1
Poland	1
Portugal	1

Table 1: Countries covered within the DH Course Registry (as of May 2019)

Between 2017 and 2018 a steadily increasing number of visitors could be observed. The DH Course Registry website had 2424 visits from 78 countries in 2017<sup>16</sup> and 3607 visits from 90 countries in 2018. From January 2019 to May 2019 the website had 1911 visits from 103 countries. The most commonly submitted training activities among those active in the DH Course Registry are full Master's programmes, of which there are 75 active entries, followed by 60 Courses, and 45 Bachelor programmes. The other categories are submitted less often, but do maintain a presence in the Registry.

<sup>16</sup> For the time before 2017 we do not have web analytics numbers.

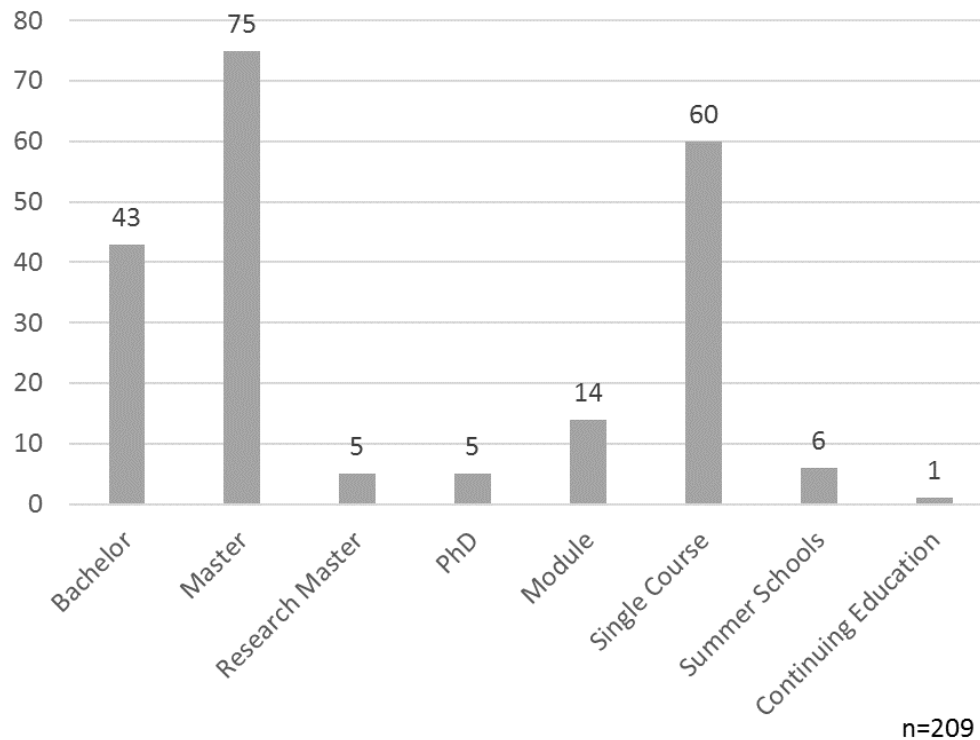


Figure 3: Distribution of education types among the active training activities<sup>17</sup>

It is also interesting to note which departments or faculties offer these training and teaching activities. It is not surprising that most of these activities are offered by Humanities faculties (171 activities out of 310). But also other faculties like Arts (33 activities), Social Sciences (11 activities) and Natural Sciences (21 activities) are offering DH-related teaching activities. Furthermore, it can be seen that in some institutions digital humanities teaching activities are organised at an interfaculty level (14 activities).

<sup>17</sup> This figure only displays the 209 teaching activities, which were active and visible as of May 2019.

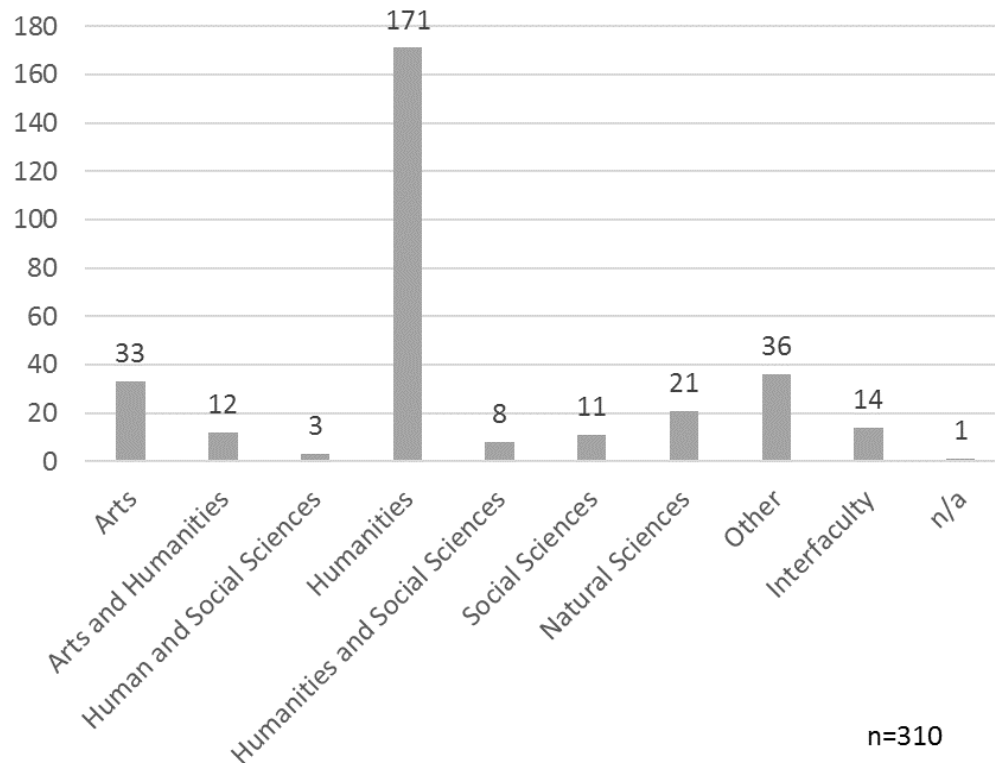


Figure 4: Distribution of faculties offering the submitted teaching activities<sup>18</sup>

The data contained in the DH Course Registry does not show a complete picture of DH-related training activities since it is a community-driven platform and we depend on individuals being willing to provide their data. In spite of this, the resource remains the most representative resource we know of to demonstrate change over time in the growth and delivery of DH courses internationally. The data available within the DHCR is not only searchable via the website but also accessible via an API (see Figure 1) in line with the recent trend of increasing development of APIs in digital humanities projects (Tasovac et al., 2016) in order to make data and metadata more accessible. However, the analysis of the available data had some limitations for third parties. The exposed basic .json and .xml API only returned a selection of actively maintained metadata, older archived data sets remained hidden for the audience. But, in the ongoing development, a new API<sup>19</sup> version, provides various query scenarios, filters and finally enables research on historical data as well. The new API version<sup>20</sup> – its data model and documentation – is compliant with OAS 3.0, the industrial standard for REST API description (Miller et al., 2018).

<sup>18</sup> In this figure we work with the total of submitted training activities, not only the active ones, so there is a total of 310.

<sup>19</sup> Within the DHCR Sustain Project (<https://www.oeaw.ac.at/acdh/projects/dhcr-sustain/>), funded within the DARIAH Theme funding call 2018/2019 the API and its documentation will be improved. The general technical development work and maintenance is funded by PARTHENOS (H2020 Grant Agreement n. 654119) and CLARIN-ERIC.

<sup>20</sup> The new API version can be accessed at <https://dhcr.clarin-dariah.eu/api>.

## **7 Concluding Remarks**

This kind of community-driven information gathering is useful both for individuals looking to find or promote DH training opportunities as well as those looking to understand the development of DH over time and internationally. An initiative like the DH Course Registry is not simple to sustain, however. The community-driven approach it has implemented brings with it both advantages and disadvantages, in particular when it comes to motivating the community to continuously add new courses and to keep the information up-to-date for already existing entries. Ensuring technical maintenance and keeping the future development steps for the resource in line with general technological possibilities provide further layers of complexity. These challenges would perhaps be insurmountable in the context of a term-limited research or infrastructure project with a national or even regional scope, but the scale and durability of the two infrastructures hosting the Course Registry are able to manage and minimise (although not remove entirely) any risks inherent in this model. Regarding technical maintenance and further development, the DH Course Registry here also benefits from its hosting organisations, as it enjoys the technological and staff support of both CLARIN and DARIAH. From this stable basis, the DH Course Registry can plan and propose further developments, harnessing community ideas and support as well as project funding to make significant iterations to the platform. For example, the DH Course Registry is now being integrated into DARIAH-CAMPUS<sup>21</sup>, a recently-launched discovery framework and hosting platform for learning resources, and has leveraged funding to improve user experience and other aspects from a variety of sources managed from within the partner research infrastructures and beyond.

As such the DHCR has established itself as an offering of two European infrastructures to the community world-wide, creating and managing research information transnationally, via a distributed, community-focussed process. In this sense, the DHCR presents an inviting open door to the international community, in particular researchers in countries that do not participate in the Course Registry yet, to participate, to cooperate, to benefit from, and contribute to it.

## **Acknowledgements**

This work was supported by different institutions and project grants: CLARIN ERIC, DARIAH-EU, PARTHENOS (H2020 Grant Agreement n. 654119), CLARIN-PLUS (H2020 Grant Agreement n. 676529), DHCR Sustain – Improving Sustainability through Usability (DARIAH Theme 2018 (S34D)), CLARIAH-AT, CLARIAH (NL). And different institutions contributed to the initiative over the years: Austrian Academy of Sciences (ÖAW), Royal Netherlands Academy of Arts and Sciences (KNAW): eHumanities Group, Data Archiving and Networked Services (DANS), Erasmus University Rotterdam, University of Cologne, Pôle Informatique de Recherche et d'Enseignement en Histoire (PIREH / University Paris 1), University of Göttingen, University of Graz.

<sup>21</sup> <https://campus.dariah.eu/>



## References

- Aasman, S. and Scagliola, S. (2017), Strategies for integrating Digital Humanities Skills and Practices into the Humanities Curriculum, DH Benelux 2017. <https://dhenelux2017.eu/2017/07/04/manifesto-for-integrating-dh-in-teaching-programs/>
- Akdag Salah, A., Scharnhorst, A. and Wyatt, S. (2015), Analysing an academic field through the lenses of Internet Science : Digital Humanities as a Virtual Community. in T Tiropanis, A Vakali, L Sartori & P Burnap (eds), *Internet Science: Proceedings of the Second International Conference, INSCI 2015, Brussels, Belgium, May 27-29, 2015*. vol. 9089, Lecture Notes in Computer Sciences, vol. 9089, Springer. DOI: 10.1007/978-3-319-18609-2\_6
- Antonijevic, S. (2015), *Amongst Digital Humanists: An Ethnographic Study of Digital Knowledge Production*, Palgrave.
- Borek, L., Dombrowski, Q., Perkins, J., Schöch, C. (2016), TaDiRAH: a Case Study in Pragmatic Classification, *Digital Humanities Quarterly*, 2016, Volume 10, number 1.
- Digital Humanities Manifesto (2009), "The Digital Humanities Manifesto 2.0.", available at: <http://manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/> (accessed 18 March 2019).
- Drucker, J. (2011), "Humanities Approaches to Graphical Display", *Digital Humanities Quarterly*, vol. 5, no. 1, 2011
- EC (2019), What is an ERIC. <http://ec.europa.eu/research/infrastructures/index.cfm?pg=eric>
- Edmond, J. (2015), Collaboration and Infrastructure. In: Susan Schreibman, Ray Siemens and John Unsworth ed., *A New Companion to Digital Humanities*. John Wiley & Sons, 2015. pp. 54-65.
- Edmond, J., Garnett, V., Burr, E., Laepke, S., Jenny Oltersdorf, J., Goulis, H. (2016). PARTHENOS Deliverable D7.1 Initial Training Plan, available at <https://zenodo.org/record/2575380#.XXe8BX-xVaQ> (accessed 21 August 2019).
- Edmond, J., Fischer, F., Romary, L., Tasovac, T. (forthcoming), *Springing the Floor for a Different Kind of Dance: Building a 21st Century Research Infrastructure for the Arts and Humanities*.
- ESFRI (2018), "ESFRI roadmap 2018" available at <http://roadmap2018.esfri.eu/media/1066/esfri-roadmap-2018.pdf> (accessed 21 June 2019).

ESFRI (European Strategy Forum on Research Infrastructures Long-Term Sustainability Working Group) (2017). Long-Term Sustainability of Research Infrastructures available at: [https://ec.europa.eu/info/sites/info/files/esfri\\_scripta\\_vol2\\_2017.pdf](https://ec.europa.eu/info/sites/info/files/esfri_scripta_vol2_2017.pdf) (accessed 22 August 2019)

European Commission (2017) Sustainable European Research Infrastructures: A call for action. Commission staff working document — Long-term sustainability of Research Infrastructures available at: [https://ec.europa.eu/info/sites/info/files/research\\_and\\_innovation/research\\_by\\_area/documents/swd-infrastructures\\_323-2017.pdf](https://ec.europa.eu/info/sites/info/files/research_and_innovation/research_by_area/documents/swd-infrastructures_323-2017.pdf) (accessed 22 August 2019).

Fišer, D., Lenardič, J. and Erjavec, T. (2018), CLARIN's Key Resource Families. Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018). More information can be found on: <https://www.clarin.eu/resource-families>.

Gold, M. K. (2012). Looking for Whitman: A Multi-Campus Experiment in Digital Pedagogy. *Digital Humanities Pedagogy: Practices, Principles and Politics*. Open Book Publishers, 2012, pp. 151–176.

Google Books (2009), "Google Books.", available at: <https://books.google.com> (accessed 21 August 2019).

de Jong, F., Maegaard, B., De Smedt, K., Fišer, D. and Van Uytvanck, D. (2018), CLARIN: Towards FAIR and Responsible Data Science Using Language Resources, Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)

Kirschenbaum, M. (2012), "What is Digital Humanities and what's it doing in English Departments?", in: Matthew K. Gold ed., *Debates in the Digital Humanities*, University of Minnesota Press, Minneapolis/London, pp. 3–11.

Koolen, M., van Gorp, J. and van Ossenbruggen, J. (2018), "Toward a model for digital tool criticism: Reflection as integrative practice", *Digital Scholarship in the Humanities*, 12 October 2018. <https://doi.org/10.1093/llc/fqy048>

Darrel Miller, D., Whitlock, J., Gardiner, M., Ralphson, M., Ratovsky, R., Sarid, U. (2018), OpenAPI Specification, Version 3.0.2. <http://spec.openapis.org/oas/v3.0.2> (accessed 21 August 2019).

Moretti, F. (2013), *Distant Reading*. Verso, London, 2013.

Murphy, E. C. and Smith S. R. (2017), Introduction to "Imagining the DH Undergraduate: Special Issue in Undergraduate Education in DH". *Digital Humanities Quarterly*, vol. 11, no. 3, 2017.

Michel, J., Yuan Kui Shen, Y. K., Aiden, A. P., Veres, A., Gray M. J., The Google Books Team, Joseph P. Pickett, J. P., Dale Hoiberg, D., Dan Clancy, D., Peter Norvig, P., Jon Orwant, J.,

Steven Pinker, S., Martin A. Nowak, M. A., Lieberman Aiden, E. (2011). "Quantitative Analysis of Culture Using Millions of Digitized Books", *Science*, vol. 331, 2011, no. 6014, pp. 176–182.

Rockwell, G. and Sinclair, S. (2012), "Acculturation and the Digital Humanities Community" in: Brett Hirsch, ed., *Digital Humanities Pedagogy*. Open Book, pp. 177–212.

Romary, L. and Edmond, J. (2019), A Tangential View on Impact for the Arts and Humanities through the lens of the DARIAH-ERIC. In: Maegaard, B. and Pozzo, R. ed., *Stay Tuned to the Future: Impact of the Research Infrastructures for Social Sciences and Humanities*, Leo S. Olschki Editore, pp. 149–158. Available at [https://www.clarin.eu/sites/default/files/STAY\\_TUNED\\_TO\\_THE\\_FUTURE.pdf](https://www.clarin.eu/sites/default/files/STAY_TUNED_TO_THE_FUTURE.pdf) (accessed 21 August 2019).

Sahle, P., Johanna Puhle, J., Lisa Rau, L. (2011), *Digitale Geisteswissenschaften*, available at <http://www.cceh.uni-koeln.de/Dokumente/BroschuereWeb.pdf> (accessed 21 August 2019).

Safradin, B. and de Jong, F. (2017), CLARIN-PLUS D5.2 Operational course and education material registry. Available at [https://office.clarin.eu/v/CE-2017-0986-CLARINPLUS-D5\\_2.pdf](https://office.clarin.eu/v/CE-2017-0986-CLARINPLUS-D5_2.pdf) (accessed 21 August 2019).

Scagliola, S. (2019), Integrating the digital into the academic curriculum - how to engage your audience – Blog C2DH – University of Luxembourg, 2019, available at <https://www.c2dh.uni.lu/thinking/integrating-digital-academic-curriculum-history-how-engage-your-audience> (accessed 21 August 2019).

Schreibman, S., Benardou, A., Clivaz, C., Durco, M., Huang, M., Papaki, E., Scagliola, S., Tasovac, T., Wissik, T. (2016), #dariahTeach: online teaching, MOOCs and beyond. In *Digital Humanities 2016: Conference Abstracts*. Jagiellonian University & Pedagogical University, Kraków, pp. 354–356.

Terras, M.(2012), "Digitization and digital resources in the humanities", in: Claire Warwick, Melissa Terras and Julianne Nyhan eds., *Digital Humanities in Practice*, Facet, 2012, pp. 47–70.

Thaller, M. (2012), Towards a reference curriculum for the Digital Humanities. Proceedings of DH2012, Hamburg. Available at <http://www.dh2012.uni-hamburg.de/conference/programme/abstracts/towards-a-reference-curriculum-for-the-digital-humanities.1.html> (accessed 21 August 2019).

Tasovac, T., Barbaresi, A., Clérice, T., Edmond, J., Ermolaev, N., Garnett, V., Clifford Wulfman, C. (2016), APIs in Digital Humanities: The Infrastructural Turn. *Digital Humanities 2016*, Jul 2016, Cracovie (Poland), pp. 93–96.

Wettlaufer, J. (2016), "Neue Erkenntnisse durch digitalisierte Geschichtswissenschaft(en)? Zur hermeneutischen Reichweite aktueller digitaler Methoden in informationszentrierten Fächern", in: Zeitschrift für digitale Geisteswissenschaften, 2016. text/html Format. DOI: 10.17175/2016\_011

White, J. W. and Gilbert, H. ed. (2016), Laying the Foundation: Digital Humanities in Academic Libraries, vol. 7, Purdue University Press, West Lafayette, Indiana.

Wissik, T. (2019), The role of research infrastructures in the area of DH education, training and competence building. In: Stefano Allegrezza (ed.), AIUCD 2019 – Book of Abstracts, Udine 2019, pp. 233–235.