The International Journal of Digital Curation
Issue 1, Volume 5 | 2010

Understanding the Information Requirements of Arts and Humanities Scholarship

Agiatis Benardou,
Digital Curation Unit-IMIS, *Athena* Research Centre
Panos Constantopoulos,
Digital Curation Unit-IMIS, *Athena* Research Centre &
Department of Informatics, Athens University of Economics and Business
Costis Dallas,
Digital Curation Unit-IMIS, *Athena* Research Centre,
Communication, Media and Culture Department, Panteion University &
Faculty of Information, University of Toronto
Dimitris Gavrilis,
Digital Curation Unit-IMIS, *Athena* Research Centre

Abstract

This paper reports on research of scholarly research practices and requirements conducted in the context of the Preparing DARIAH European e-Infrastructures project, with a view to ensuring current and future fitness for purpose of the planned digital infrastructure, services and tools. It summarises the findings of earlier research, primarily from the field of human information behaviour as applied in scholarly work, it presents a conceptual perspective informed by cultural-historical activity theory, it introduces briefly a formal conceptual model for scholarly research activity compliant with CIDOC CRM, it describes the plan of work and methodology of an empirical research project based on open-questionnaire interviews with arts and humanities researchers, and presents illustrative examples of segmentation, tagging and initial conceptual analysis of the empirical evidence. Finally, it presents plans for future work, consisting, firstly, of a comprehensive re-analysis of interview segments within the framework of the scholarly research activity model, and, secondly, of the integration of this analysis with the extended digital curation process model we presented in earlier work.[1]

[1] This paper is based on the paper given by the authors at the 5th International Digital Curation Conference, December 2009; received November 2009, published June 2010.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. ISSN: 1746-8256 The IJDC is published by UKOLN at the University of Bath and is a publication of the Digital Curation Centre.

Introduction

The need to develop a sound understanding of the research process in the Arts and Humanities, as a special kind of "business process", and as a prerequisite for the development of appropriate digital infrastructures, tools and services for scholarship has been identified since the 1990s (American Council of Learned Societies, 1998; Bearman, 1996). Measures suggested included "a broader examination of the methodology and practice of the humanities, and of the function of information resources and scholarly communication", and, more concretely, the identification of "scholarly tasks corresponding with specific 'modes' of research [to be] matched with a tool-set of systems and interface capabilities (e.g. annotation and attribution, comparison and presentation, synthesis)" (Dallas, 1999).

A similar call for a closer focus on the requirements stemming from actual information work in scholarship emerges now as a crucial part in the agenda of major initiatives to develop appropriate, timely and interoperable infrastructures, services and tools to serve the current and emerging needs of humanists; at a time of large-scale digitisation programmes in the field of cultural heritage, producing Web-accessible digital resources of primary interest to arts and humanities research, as well as major investments announced for digital research infrastructures across the Atlantic (Blanke & Dunn, 2006; Crane, Babeu, & Bamman, 2007; Green & Roy, 2008), the need to ensure fitness for purpose of the planned repositories, services and tools cannot be underestimated.

This paper reports on work conducted in the context *of Preparing DARIAH: Preparing for the construction of the Digital Research Infrastructure for the Arts and Humanities*, a collaborative European project co-funded by the ESFRI e-Infrastructures programme, aiming at providing the foundations (strategic, financial, legal, technological and conceptual) for the timely design and construction of the digital infrastructure requisite for scholarly research in the arts, humanities and cultural heritage in Europe (Constantopoulos et al., 2008). The Digital Curation Unit-IMIS, *Athena* Research Centre is currently engaging in a two-pronged research programme within the conceptual modelling work-package of DARIAH consisting: a) of the formulation of a conceptual model for scholarly research activity suitable for the representation of actual information practice in scholarly work, and b) of an empirical study of scholarly research activity, based on the elicitation, transcription, encoding, analysis and interpretation of open-question interviews with humanities scholars across Europe. We present here the background, motivation and conceptual framework, the plan of work and methodology of our empirical research, and a brief conclusion with planned work.

Background

Earlier studies of scholarly work range from ethnographically-based research and epistemologically informed monographs from the field of science, technology and society studies (STS), such as, notably Knorr-Cetina (1999) and Latour & Woolgar (1986), to information behaviour studies stemming from an information science perspective, including Stone, who as early as 1982 had noted the solitary nature of humanistic research; the borrowing of methods drawn from other disciplines by humanists; their reliance on monographs rather than journals; the importance of

retrospective materials (both primary sources and literature) for research; and, the lack of adequate bibliographic tools (Stone, 1982). In the 1990s, Bates and colleagues investigated information work by humanities scholars visiting the Getty Research Institute, noting in their results the major importance of named entities (proper names, places, titles of works) as entry points to resource discovery, but also the fact that the introduction of digital tools did not supplant traditional methods, but co-existed synergistically with them (Bates, Wilde, & Siegfried, 1995); further work by Bates explored the reliance of humanities scholars, as well as artists, on multimedia resources (Bates, 2001). While these studies dealt primarily with the needs of humanists *vis-à-vis* scholarly objects (i.e. outcomes of scholarly research), such as in a library setting, further studies accommodated also information work on primary and secondary sources, focusing on disciplines such as history (Dalton & Charnigo, 2004; Delgadillo & Lynch, 1999; Duff, Craig, & Cherry, 2004; Tibbo, 2003) and art history (Beeman, 1994; Hemmig, 2008; Odum, 1998). Short reports, based on the consultations and research concluded by the AHRC Methods Network in the UK, summarise needs and likely scenarios for ICT use in humanities research in particular fields, including history, art history, archaeology, and museums and heritage; a conceptual overview of ICT research methods employed by researchers has been developed by the Methods Network in the form of a thesaurus ("methods taxonomy").

Further empirical research pointed to:

1. the increasingly interdisciplinary nature of humanistic research, bearing significantly on the information service characteristics needed to cater for them (Palmer & Neumann, 2002; Searing, 1996).
2. the increased reliance of humanists on an accommodative process of assembling information resources relevant to the task at hand though browsing, "craft[ing their] arguments over timelines", rather than on comprehensive searching (Covi, 1999).
3. the importance of strategies for finding and collocating relevant information other than searching, mainly through berry-picking, or chaining, illustrated, in the case of scholarly objects, through strategies of footnote chasing, citation searching, journal run, area scanning, subject and author searches, but also relevant to primary source information use (Bates, 1989).
4. the importance of annotation for scholarly work (Bradley & Vetch, 2007).
5. the frequent reliance on serendipity and the possibility of non-deterministic contextual "discovery" by scholars (Chan, 2007; Duff & Johnson, 2002; Foster & Ford, 2003).

Ellis and his team at Sheffield University spearheaded qualitative work, fuelled by grounded theory research on four different research communities across the sciences, social sciences, and the humanities (Ellis, 1993). They identified six common processes across disciplines: *starting, chaining, browsing, differentiating, monitoring* and *extracting*, further updated a decade later by Meho and Tibbo with processes of *accessing, networking* and *verifying* (Meho & Tibbo, 2003). Alternative models appear to focus more narrowly on searching, or information seeking, behaviour, or on the motives rather than the activities of researchers (Foster, 2004; Kuhlthau, 1991).Brockman, Palmer and associates presented a broadly based conceptual framework of the nature of scholarly work, focusing on processes of

reading, collaborative networking, researching and searching, and *ways of writing*, and emphasising the differences in information work in the humanities *vis-à-vis* other disciplines (Brockman, Neumann, Palmer, & Tidline, 2001).

The notion of "scholarly primitives", initially advocated by Unsworth with reference to processes employed by literary scholars (Unsworth, 2000), was found to be useful as a concept to identify common, low-level scholarly activities in other humanistic disciplines such as history and historical geography (Mostern, 2006). A recent study, aimed at defining appropriate infrastructures and services at the Minnesota University Libraries (University of Minnesota Libraries, 2006), is based on the identification of four groups of scholarly information activities (*discover, gather, create* and *share*).

The latest synthesis, by Palmer and associates, defines five "scholarly activities": *searching, collecting, reading, writing, and collaborating.* These, as well as a bucket of "cross-cutting primitives" are further refined to a more detailed, and useful, list of twenty granular "scholarly primitives"; of these, *browsing, collecting, re-reading, assembling, consulting* and *notetaking* were found to be particularly common in the humanities, while *chaining, accessing, assessing, disseminating* and *networking* were seen as equally applicable to the humanities as well as other disciplines. *Chaining,* in particular, was identified as the most notable activity among humanists as they seek information. In addition, *probing* and *translating* activities were found to be most common in interdisciplinary research, a noteworthy finding considering the frequently interdisciplinary nature of work in the arts and humanities (Palmer, Teffeau, & Pirmann, 2009). On a broader perspective, a statement of strategic direction with regard to the digital infrastructure for the humanities in the US, the 2006 report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences, provides a useful consensus statement on what constitutes scholarly practices that need to be addressed by ICT infrastructure and tools (American Council of Learned Societies, 2006). Finally, Borgman's recent monograph provides an overarching framework in order to conceptualise the relationship between disciplinary practices in the humanities, and documents and data, with technological infrastructures and tools (Borgman, 2007).

Motivation and conceptual framework

The vision of the *Preparing DARIAH* project is to provide the mechanisms – strategic, legal, financial, organizational and technological – necessary in order to ensure the long-term availability and access to resources needed by European researchers in the arts and humanities, at all stages of the scholarly research process, by means of an appropriate digital infrastructure spanning across Europe. When implemented, the DARIAH infrastructure will hopefully help researchers become aware of resources available in repositories across Europe, and provide them with the means – the services and tools - to locate and access these resources, be it primary data, documentary evidence, or secondary, scholarly objects.

An important challenge when planning systems that do not yet exist, especially relevant in the case of innovation and knowledge-related domains such as the continuously-evolving world of academic research, consists of the difficulty in eliciting requirements merely through observing actual practice; the problem is familiar in the field of digital curation, where an important challenge in the face of

longevity is to assure the future fitness for purpose, in other words to avert epistemic failure, of stored representations of information objects (Dallas, 2007; Giaretta, 2008; McCarthy, 2007). The emergence of digital humanities as a methodological current that is radically interdisciplinary in approach, often formal or quantitative in method, and making intensive use of the data management and visualisation capabilities of information technology, is a case in point why historical practice is not a sure guide of future developments. On the other hand, shifts in research interests and methodological changes in the arts and humanities, still relevant today, may be dated to decades even centuries, and some fundamental traits do not seem to have changed in centuries of scholarship: research remains often idiographic rather than nomothetic, method is as a rule inter-textual rather than analytical, and the value of evidence and scholarly literature persists with time, rather than drastically decreasing as in the case of the sciences (Dallas, 1999). In this light, the specification of digital infrastructure for the arts and humanities needs obviously to address the historical practices, needs and perceptions of actual researchers, rather than be merely driven by the interests and priorities of technology and service providers.

Understanding the nature and information requirements of scholarly research, notwithstanding differences between disciplines, research fields and methodological approaches, emerges, thus, as an important motivation, and a prerequisite for the definition of affordances of infrastructures, services and tools fit for the purpose of current and future scholarly research. On the other hand, requirements need to be substantiated by actual evidence, drawn from the domain of arts and humanities, and encompassing not only which digital resources, services and tools researchers use, and how, but, more generally, how they interact with the whole spectrum of information and conceptual entities – digital as well as non-digital – involved in the research process. These include not only entities typically thought of as primary research data or as scholarly literature, but also finding aids, ways of organizing and sequencing research activities, terminological and encyclopaedic resources, standard procedures, and tools. A key reason for conducting this work is to produce, as part of *Preparing DARIAH*, an account of arts and humanities information practice which can be operationalised through the definition of specific functionalities of tools and services in the planned infrastructure (also taking into account strategic and policy considerations), and which can provide an evidence-based rationale of why and how DARIAH tools and services will correspond, and address, real, recurrent and important work patterns and needs of arts and humanities research.

A conceptual perspective for the identification of pertinent categories and properties representing scholarly research is provided for by cultural-historical activity theory, a school of thought and set of principles developed originally in psychology and later found useful in fields as diverse as developmental research, organisations, work and ergonomics, social aspects of technology, Human-Computer Interaction, and digital curation (Bannon & Bødker, 1991; Dallas, 2007; Engeström, 1987, 2000; Kaptelinin & Nardi, 2007; Leont'ev, 2007; Nardi, 1996). An *activity,* understood as "purposeful interaction of a subject with the world"; is always directed toward some *object*, a physical or conceptual entity (or entities); this object embodies, also, the fulfilment of some objective or *motive,* which in turn is intended to meet a specific *need*. Activity systems are composed as a hierarchy of *activities,* constituted by conscious *actions* designed to meet hierarchically structured goals. Subjects can be

individuals, but also *communities* with shared needs and motives. Purposeful *interaction* between subjects and objects takes place by means of *tool mediation,* whereby tools include not just physical, but also procedures, computer programs, languages and signs (Kaptelinin & Nardi, 2007).

In order to provide an operational framework for the actual representation, analysis and understanding of evidence related to scholarly information practice in the arts and humanities, we used an analysis of relevant literature, in order to develop a conceptual model for scholarly research activity which we checked for relevance on the basis of an initial analysis of the empirical research presented below. The model (Figure 1) can be seen as an application of the CIDOC Conceptual Reference Model, an established and stable international standard for cultural information (Crofts, Doerr, Gill, Stead, & Stiff, 2009). Entities in the DCU conceptual model for scholarly research activity, such as *Actor, Research Goal,* and *Procedure,* are specialisations of CIDOC CRM entities (listed in parentheses in the model, prefixed by "E") and are thus endowed by all properties of these entities as defined in the standard; relationships between entities, presented as arc labels in the model, are equivalent to CIDOC CRM properties or specialisations thereof. Entities such as *Research Activity, Procedure, Method, Information Object, Proposition, Research Goal, and Tool/Service,* and properties connecting these entities, correspond to notions relevant for the conceptualization of scholarly research process by humanists, such as those sought by our empirical research.

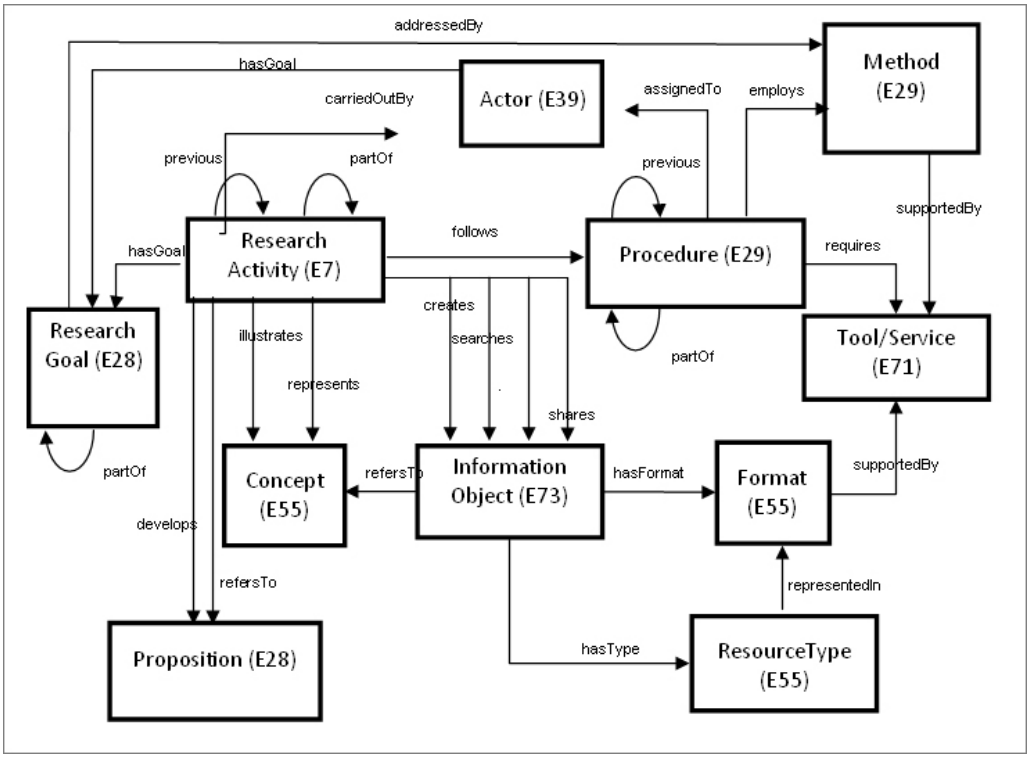

Figure 1. Scholarly research activity model.

The DCU model of scholarly research activity is intended to facilitate the elicitation of requirements, and the design and development of information repositories and services in digital infrastructures that support research in the arts and humanities. For this purpose, the model should be able to represent not just actual-

historical information on a structured set of events (*what, where, when was done?*), but also encompass notions of subject (*who did it?*), method (*how was it done?*) and purpose (*why was it done?*). This necessitated a refinement to the CIDOC CRM *Activity* entity, so as to introduce a distinction between *Process* and *Procedure* while maintaining corresponding (though not necessarily isomorphic) descriptions of the two; this is a duality often encountered in conceptual models of task-oriented systems, such as enterprise information systems (Dietz, 2006; Hay, 1996). The model is still work in progress by the DCU.

Empirical research

Following a survey of earlier research on scholarly information work, a series of semi-structured interviews with arts and humanities scholars from across Europe was planned. From January to June 2009, twenty four interviews were conducted; most interviews were conducted face-to-face in Greece, either in Greek or in English; a further round is planned across Europe, based on Skype and/or telephone interviews. Survey participants had to be researchers doing advanced research in the field of the arts and humanities. They were selected according to expertise and were members of academic departments or research institutes.

All interviews were recorded with the consent of the participants, transcribed into written form, segmented according to topic, and tagged by two members of the research team (one analyst and one classical archaeologist / ancient historian). The duration of the interviews ranged from forty-five to ninety minutes, depending on the interviewees' personal interests, research methodology employed, and reports of other activities which we considered would be relevant to the research, such as, notably, academic teaching: an activity that frequently coexists with scholarly research, in as much as academic teachers are, as a rule, involved in active research. Interviews were conversational, and based on an open questionnaire, encouraging the elicitation of further information when this was justified by the scope of the present research. Despite known difficulties arising in the analysis of free text interviews, the open questionnaire format helped identify important differences between specific methodological perspectives and thematic interests within and across disciplines.

Participants

Interviewees covered a wide scope of disciplines within the arts and humanities, ranging from history (ancient, modern, contemporary), history of art, Byzantine studies, archaeology (iconographic research, experimental archaeology, archaeological site-based research, archaeological survey) and anthropology (anthropology of material culture, ethnomusicology). They were at different career stages, ranging from doctoral candidates to full professors, and held both academic and non-academic positions. They displayed widely varying familiarity and intensity of use of ICT tools, ranging from laggards to early adopters and innovators according to the Rogers Adoption/Innovation curve (Rogers, 1995). Thirteen of the interviewees were male, the rest female. About two thirds were Greek. The largest groups of interviewees by discipline were archaeologists (11, albeit of widely diverse specialties, from iconography to site survey), historians (5), humanities/cultural studies (3) and literature scholars (3).

Questionnaire

The typical interview scenario consisted, firstly, of a short introduction about the DCU, followed by a short introduction about DARIAH and its goals; secondly, of a set of open questions regarding both conventional and digital collections and tools, and, thirdly, of additional questions specific to the use of digital collections by scholars. Questions were organised as indicated by the themes presented in Table 1.

Theme	Details
Point of departure	This question concerns the way in which a research topic is conceived and formulated, as well as the initial pursuit of primary evidence and information used to document it. Key issues regarding research planning are raised here.
Annotation	Discussion on annotation issues focuses on the form of notes and the way each researcher chooses to organize them. Furthermore, researchers are encouraged to expand on the practicalities, ranging from photocopying and photographing to underlining and scanning, as well as the organization of their own databases and the methods they employ to compose their material at various stages.
Terminology	This question deals with terminology used during research (either commonly used within a discipline or individually developed and used by the scholar).
Raw data – Sampling – Comparison	This key area explores the notion and use of comparison and juxtaposition as research methods in the field of arts and humanities. Moreover, it touches upon conceptions of data constitution, as in data selection and sampling.
Personal Communication – Grey Literature	As personal communication is a key feature of research and constitutes a viable method of data collection, this area seeks to investigate the methods undertaken and the individuals involved in it during various stages of research.
Collaboration	Collaboration is a controversial issue for arts and humanities research. In this section of the interview, we attempt to understand how collaboration is conceived, and what are key drivers and barriers.
Work Saving	This question refers to the methods and frequency with which individuals save their work.
Workplace	This theme aims to identify and describe the physical environment where research is preferably conducted, as well as equipment and material used.
Available digital resources	Discussion focuses on the usability of digital resources available for research in the field of the arts and humanities. Individuals expand on the stages of research in which they consult digital resources, while making references to specific resources and comment on them.
Varia	This theme concerns activities which take place concurrently, or are otherwise congruent with research work, which may or may not influence actual research.

Table 1. Interview topics.

The list of eleven initial open-ended themes correspond to questions asked in various forms, in order to elicit responses that match specific activities, procedures, methods, types of information objects etc. employed by individual researchers. Themes explored touched upon both methodological and conceptual questions. Questions on methodology included annotation (Q.2), as a continuous activity rather than a separate stage of research process, searching and information seeking (including serendipitous discovery, chaining, browsing), databases, data management and organization (Q.3), terminology (Q.4), raw data – sampling (Q.5), work saving (Q.7), workplace (Q.9), as well as availability and usability of digital resources and user needs (Q.10).Conceptual topics embraced issues like the initial conception and approach of research topic, and the beginning of a research project's life-cycle (Q.1) (Meho & Tibbo, 2003), the notions of comparison (Q.5) and collaboration (Q.8) as well as observations on grey literature (Q.6). Further topics varied according to the interviewee (Q.11).

It should be noted that the order of the questions/topics loosely followed the order of activity of scholarly research; interviews began with the conception and development of a research topic, continued with issues of primary source/data and secondary source searching and chaining, serendipitous finding, note taking and annotation, information and data management and organization, as well as usability of and requirements for available digital resources, grey literature and collaboration. In short, the interview followed the research lifecycle, allowing for individual researchers to expand on issues they considered as key to their work.

Examples of interview transcripts

Interviews were transcribed and segmented to parts that correspond with the broadest possible research activity, even if such activity combines more than one "scholarly primitive"; in fact, collocation of such "primitives" may offer valuable information on the sequence structure of scholarly research activities.

Representative interview segments, indicating the variety of answering styles and interests, are presented in Table 2.

Postdoctoral researcher / Greece / Classical Archaeology
To begin with, I make catalogue cards. These cards may be in an electronic form, or may be printed. When you work in a museum it is sometimes more convenient and handy to have the cards printed. Therefore I make the cards and then print them out to complete them manually. What is in those cards largely depends on what you ask for each time. That is, there must be a description of the material, its shape and decoration for example, or the clay with its technical characteristics. This is the initial stage. Largely this leads to the catalogue, which also includes the dimensions of the objects. In other works I make my own database, in which I also include my sketches and drawings.
Doctoral candidate / Greece / Cultural Anthropology
I have worked with people in several projects and co-authored numerous articles. The internet is surely very helpful. Lately, we usually upload a text and work on different versions of it or work on a wiki. This is extremely time-saving and does not cancel the personal discussion. There are things I would like to be able to do on such digital environments, as peer-to-peer or wiki ones. There is gap in that I can work on a text real-time, however our field is largely representational/visual, so I sometimes need to be shown or to show things myself, such as a map or a place. Working with drawings, images, etc., on which the other person can interfere and work, is difficult. There are no such satisfactory internet platforms which can support such an activity. They are still being developed. I need mixed and combined means and ways of working texts and images. Anything that can be as similar as possible to working around a table.
Professor / UK / Classical Archaeology
As far as possible I use established terms as clearly as possible. I would rather try to describe what I'm looking at and see how it sits within the framework of discussion in the literature. I think if you have to call a new term you could have to be really sure what you are doing. If you do use an existing one in a different sense you need to be very precise about why you're doing it. When you are doing interdisciplinary research you have to be very careful. And also have to be very careful when you think that any term you use is going to be cross referred by other European languages. If you use a term with capital letters and quotes around it by the time it is in German it will have a different significance. So one does need to be careful. Where one does have to create a new term it needs to be glossed with the kind of definition that you hope will then get into the secondary literature in its own right.

Table 2. Examples of interview segments.

Tagging and structuring

The answers fell into the eleven top level questions as indicated above. The results, each answer was linked to each respective question and tagged using free text tags. The analysis process involves the following steps:

1. The extraction of free text tags from the transcribed answers (high entropy words, high relevance phrases). The tagging process was carried out in parallel by two different researches (an archaeologist and an engineer) in order to cover different viewpoints.
2. The extracted tags were clustered together (per interview) in order to reduce dimensionality.
3. Similar tags were identified across all interviews and were merged (where possible).

This process allowed the creation of a tree-like concept map (Figure 2) used to help identify a) important tasks researchers are engaged in, and b) possible ways to expand / improve the interview process. For example, *berry-picking, chaining, combining, annotation, thematic organization, translation,* and *database development* are some of the activities that may provide useful input for the specification of functionalities of a digital tool supporting scholarly research process.

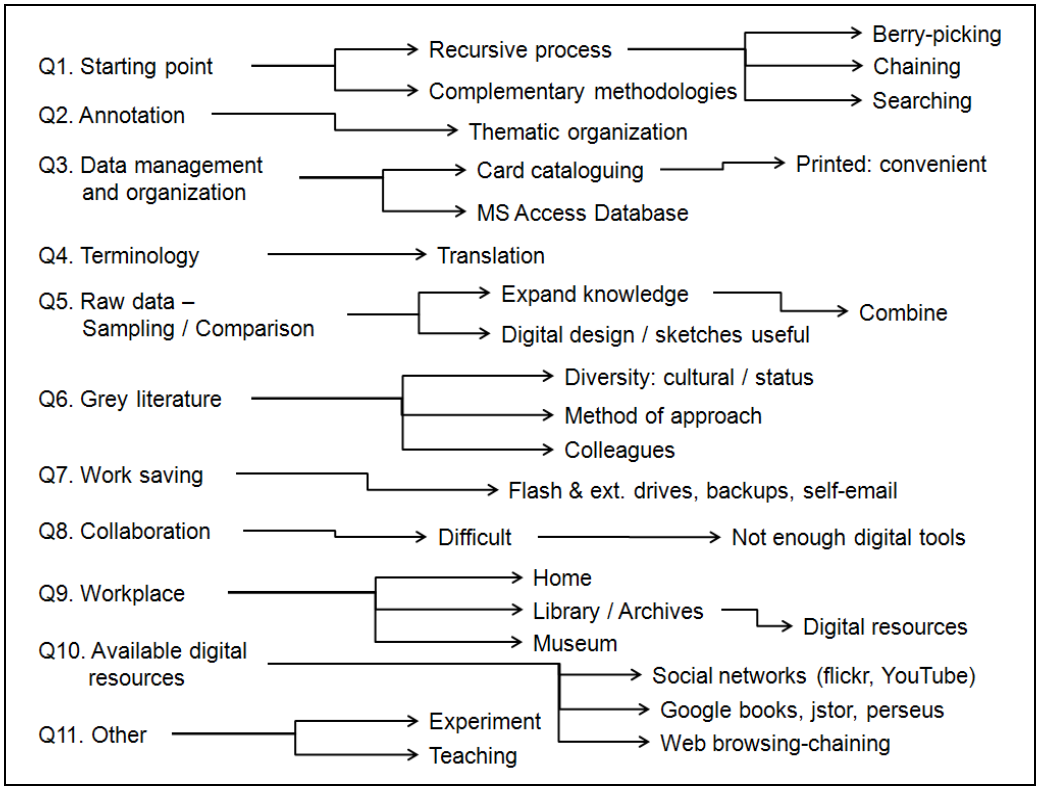

Figure 2. Sample tag map per question.

Conclusion

This paper presents our approach to understanding, eliciting and analysing user requirements for information in scholarly research, a required step for the evidence-based evaluation and definition of functional specifications of the planned digital research infrastructure for arts and humanities research, conducted as part of the *Preparing DARIAH* European e-Infrastructures project. It summarises the findings of earlier research, primarily from the field of human information behaviour as applied in research and scholarly work, it presents a conceptual perspective informed by cultural-historical activity theory, it introduces briefly our conceptual model for scholarly research activity (which constitutes the first concrete output of this research), it

describes the plan of work and methodology of the empirical research project, and presents illustrative examples of segmentation, tagging and initial conceptual analysis of the empirical evidence.

All in all, our initial analysis so far indicates that arts and humanities scholars engage in, and value highly, not only information seeking activities, but also research activities related to the *curation* of information objects such as primary and secondary data, and epistemic objects; arts and humanities researchers, in that sense, are curators *par excellence* of scholarly information, playing a key part in transforming "raw" (primary) into "institutional" (secondary) facts (Searle, 1997), through augmenting information objects semantically through annotation and edition, and through transforming them into knowledge objects by means of scholarly writing and publication. This conclusion, if confirmed, may have important repercussions on the specification of e-infrastructures, and also on our understanding of digital curation process with regard to research resources in the arts and humanities.

In the next stage of the project, we plan to use the scholarly research activity model as a framework for formal representation of the information identified in the interviews and for further conceptual analysis of the results of the empirical research. We will do so via a second iteration of the analysis in the first phase of twenty-three interviews, and in a combined process as we collect the second phase of (Skype or telephone-based) interviews with European researchers. We will be able at this stage to propose a substantive discussion on evidence-based requirements for arts and humanities research needs relevant to the planned digital research infrastructure. Apart from its direct utility for scoping user requirements, this exercise will provide a practical validation mechanism for the power and pertinence of the conceptual model.

A further goal is to also represent the conceptual content of interview results on the basis of the extended digital curation process model, combining earlier models by DCC and DCU (Constantopoulos et al., 2009). This is a natural step, as research activities such as those noted in the literature, and identified in our tagging process, can be mapped well within the framework of the extended process indicated by our original digital curation model (Constantopoulos & Dallas, 2008), encompassing goals, and pre-ingest stages, as well as use/experience, and later migrated into the DCC&U model. This will hopefully provide a further common conceptual ground between projects in digital infrastructures, such as DARIAH, and the field of digital curation.

References

American Council of Learned Societies (1998). *Computing and the humanities: Summary of a roundtable meeting*. Chicago: American Council of Learned Societies.

American Council of Learned Societies (2006). *Our cultural commonwealth: The report of the American Council of Learned Societies Commission on cyberinfrastructure for the humanities and social sciences*. American Council of Learned Societies. Retrieved April 15, 2009, from http://www.acls.org/uploadedFiles/Publications/Programs/Our_Cultural_Commonwealth.pdf]

Bannon, L. J., & Bødker, S. (1991). Beyond the interface: Encountering artifacts in use. In *Designing interaction: psychology at the human-computer interface*. Cambridge, MA & New York, NY: Cambridge University Press.

Bates, M. J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, *13*(5), 407-424.

Bates, M. J. (2001). *Information needs and seeking of scholars and artists in relation to multimedia materials*. Retrieved April 22, 2009, from http://www.gseis.ucla.edu/faculty/bates/scholars.html

Bates, M. J., Wilde, D. N., & Siegfried, S. (1995). Research practices of humanities scholars in an online environment: The Getty online searching project report no. 3. *Library and Information Science Research*, *17*(1), 5-40.

Bearman, D. (1996). Overview and discussion points. In *Research Agenda for Networked Cultural Heritage* (pp. 7-22). Santa Monica, CA: Getty AHIP.

Beeman, W. O. (1994). Stalking the art historian. *Work and Technology in Higher Education: The Social Construction of Academic Computing*, 89.

Blanke, T., & Dunn, S. (2006). The arts and humanities e-Science initiative in the UK. In *Second IEEE International Conference on e-Science and Grid Computing, 2006. e-Science'06* (pp. 136–136).

Borgman, C. L. (2007). *Scholarship in the digital age: information, infrastructure, and the Internet*. Cambridge, MA: MIT Press.

Bradley, J., & Vetch, P. (2007). Supporting annotation as a scholarly tool: Experiences from the Online Chopin Variorum Edition. *Literary and Linguistic Computing*. doi: 10.1093/llc/fqm001.

Brockman, W. S., Neumann, L., Palmer, C. L., & Tidline, T. J. (2001, December 1). Scholarly work in the humanities and the evolving information environment. *The Digital Library Federation.* Retrieved April 22, 2009, from http://www.diglib.org/pubs/dlf095/index.htm

Chan, S. (2007). Tagging and searching - serendipity and museum collection databases. In *Museums and the Web 2007: Proceedings*. Toronto: Archives and Museum Informatics. Retrieved from http://www.archimuse.com/mw2007/papers/chan/chan.html

Constantopoulos, P., & Dallas, C. (2008). Aspects of a digital curation agenda for cultural heritage. In V. Marík, J. M. Bradshaw, J. Meyer, W. A. Gruver, & P. Benda (Eds.), *DHMS 2008: IEEE SMC International Conference on Distributed Human-Machine Systems* (pp. 317-323). Prague: IEEE Systems, Man and Cybernetics Society; Czech Technical University in Prague.

Constantopoulos, P., Dallas, C., Androutsopoulos, I., Angelis, S., Deligiannakis, A., Gavrilis, D., et al. (2009). DCC&U: An extended digital curation lifecycle model. *International Journal of Digital Curation*, *4*(1), 34-45.

Constantopoulos, P., Dallas C., Doorn, P., Gavrilis, D., Gros, A., & Stylianou, G. (2008). Preparing DARIAH. In *Proceedings of the International Conference on Virtual Systems and MultiMedia (VSMM08)*. Nicosia, Cyprus. Retrieved from http://www.dcu.gr/dcu/Documents/documents/preparing-dariah/en/attachment

Covi, L. M. (1999). Material mastery: situating digital library use in university research practices. *Information processing and management*, *35*(3), 293-316.

Crane, G., Babeu, A., & Bamman, D. (2007). eScience and the humanities. *International Journal on Digital Libraries*, *7*(1), 117–122.

Crofts, N., Doerr, M., Gill, T., Stead, S., & Stiff, M. (Eds.). (2009). *Definition of the CIDOC Conceptual Reference Model (version 5.0.1)*. ICOM/CIDOC CRM Special Interest Group. Retrieved from http://cidoc.ics.forth.gr/official_release_cidoc.html

Dallas, C. (1999). Humanistic research, information resources and electronic communication. In J. Meadows & H. Boecker (Eds.), *Electronic Communication and Research in Europe* (pp. 209-239). Luxembourg: European Commission.

Dallas, C. (2007). An agency-oriented approach to digital curation theory and practice. In J. Trant & D. Bearman (Eds.), *The International Cultural Heritage Informatics Meeting Proceedings*. Toronto: Archives & Museum Informatics. Retrieved August 5, 2009, from http://www.archimuse.com/ichim07/papers/dallas/dallas.html

Dalton, M. S., & Charnigo, L. (2004). Historians and their information sources. *College and Research Libraries*, *65*(5), 400-426.

Delgadillo, R., & Lynch, B. P. (1999). Future historians: Their quest for information. *College and Research Libraries*, *60*(3), 245-260.

Dietz, J. (2006). Enterprise ontology – theory and methodology. Berlin and Heidelberg, Springer Verlag.

Duff, W., Craig, B., & Cherry, J. (2004). Historians use of archival sources: Promises and pitfalls of the digital age. *The Public Historian*, *26*(2), 7-22. doi:10.1525/tph.2004.26.2.7.

Duff, W. M., & Johnson, C. A. (2002). Accidentally found on purpose: Information-seeking behavior of historians in archives. *The Library Quarterly 72*(4), 472-496.

Ellis, D. (1993). Modeling the information-seeking patterns of academic researchers: A grounded theory approach. *The Library Quarterly*, *63*(4), 469-486.

Engeström, Y. (1987). *Learning by expanding: An activity theoretical approach to developmental research*. Helsinki: Orienta-Konsultit. Retrieved from http://communication.ucsd.edu/MCA/Paper/Engestrom/expanding/toc.htm

Engeström, Y. (2000). Activity theory as a framework for analyzing and redesigning work. *Ergonomics*, *43*(7), 960-974.

Foster, A. (2004). A nonlinear model of information-seeking behavior. *Journal of the American Society for Information Science and Technology*, *55*(3), 228-237.

Foster, A., & Ford, N. (2003). Serendipity and information seeking: an empirical study. *Journal of Documentation*, *59*(3), 321.

Giaretta, D. (2008, July 14). DCC approach to digital curation - under development. *DCC development website*. Retrieved from http://twiki.dcc.rl.ac.uk/bin/view/OLD/DCCApproachToCuration

Green, D., & Roy, M. (2008). Things to do while waiting for the future to happen: Building cyberinfrastructure for the liberal arts. *EDUCAUSE Review*, 35-48. doi: http://hdl.handle.net/10090/5694.

Hay, D. C. (1996). *Data models: conventions of thought*. New York, Dorset House.

Hemmig, W. S. (2008). The information-seeking behavior of visual artists: A literature review. *Journal of Documentation*, *64*(3), 343-362.

Kaptelinin, V., & Nardi, B. A. (2007). *Acting with technology: Activity theory and interaction design*. Cambridge, MA: MIT Press.

Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge Massachusetts and London, Harvard University Press.

Kuhlthau, C. C. (1991). Inside the search process: Information seeking from the user's perspective. *Journal of the American Society for Information Science*, *42*(5), 361-371.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton University Press.

Leont'ev, A. N. (2007, May 5). Activity, consciousness and personality. Retrieved from http://lchc.ucsd.edu/MCA/Paper/leontev/index.html

McCarthy, G. (2007). Finding a future for digital cultural heritage information resources using contextual information frameworks. In *Theorizing digital cultural heritage: A critical discourse* (pp. 245-260). Cambridge, MA: MIT Press.

Meho, L. I., & Tibbo, H. R. (2003). Modeling the information-seeking behavior of social scientists: Ellis' study revisited. *Journal of the American Society for Information Science and Technology*, *54*(6), 570-587.

Mostern, R. (2006). Historical gazetteers: an experiential perspective. Retrieved April 15, 2009, from http://www.ecai.org/projects/gazetteer/AHA-paper--short.pdf

Nardi, B. A. (1996). *Context and consciousness: activity theory and Human Computer Interaction*. Cambridge, MA: MIT Press.

Odum, L. (1998). The uses of archival materials by art historians. *MSLS Thesis. University of North Carolina at Chapel Hill*.

Palmer, C. L., & Neumann, L. J. (2002). The information work of interdisciplinary humanities scholars: Exploration and translation. *The Library Quarterly*, *72*(1), 85-117.

Palmer, C. L., Teffeau, L. C., & Pirmann, C. M. (2009). *Scholarly information practices in the online environment*. Dublin, OH: OCLC. Retrieved from http://www.oclc.org/research/activities/past/workagenda/researchinfo/support/practices.htm

Rogers, E. M. (1995). *Diffusion of innovations*. Free Press.

Searing, S. E. (1996). *Meeting the information needs of interdisciplinary scholars: Issues for administrators of large university libraries* (pp. 315-42).

Searle, R. J. (1997). *The Construction of Social Reality*. New York: Free Press.

Stone, S. (1982). Humanities scholars: information needs and uses. *Journal of Documentation*, *38*(4), 292-313.

Tibbo, H. R. (2003). Primarily history in America: how US historians search for primary materials at the dawn of the digital age. *American Archivist*, *66*(1), 9-50.

University of Minnesota Libraries. (2006). *A multi-dimensional framework for academic support: final report*. Minneapolis: University of Minnesota Libraries. Retrieved from http://purl.umn.edu/5540

Unsworth, J. (2000). Scholarly Primitives: What methods do humanities researchers have in common, and how might our tools reflect this? King's College, London. Retrieved August 4, 2009, from http://www3.isrl.illinois.edu/~unsworth/Kings.5-00/primitives.html