DHQ: Digital Humanities Quarterly
Preview
2023
Volume 17 Number 3

Making the Whole Greater than the Sum of its Parts: Taxonomy Development as a Site of Negotiation and Compromise in an Interdisciplinary Software Development Project

Jennifer C. Edmond <edmondj_at_tcd_dot_ie>, Trinity College Dublin https://orcid.org/0000-0001-9991-1637
Alejandro Benito Santos <abenito_at_usal_dot_es>, University of Salamanca https://orcid.org/0000-0001-5317-6390
Michelle Doran <m_dot_doran_at_ria_dot_ie>, Digital Repository of Ireland https://orcid.org/0000-0001-7850-6886
Roberto Therón <theron_at_usal_dot_es>, University of Salamanca https://orcid.org/0000-0001-6739-8875
Michał Kozak https://orcid.org/0000-0002-8715-9326
Cezary Mazurek <mazurek_at_man_dot_poznan_dot_pl>, Institute of Bioorganic Chemistry of the Polish Academy of Sciences
Eveline Wandl-Vogt <eveline_dot_wandl-vogt_at_oeaw_dot_ac_dot_at>, Austrian Academy of Sciences
Aleyda Rocha Sepulveda <aleyda_dot_rocha_at_oeaw_dot_ac_dot_at>, Austrian Academy of Sciences

Abstract

This paper describes the experience of a group of interdisciplinary researchers and research professionals involved in the PROgressive VIsual DEcision-making in Digital Humanities (PROVIDEDH) project, a four-year project funded within the CHIST-ERA call 2016 for the topic “Visual Analytics for Decision Making under Uncertainty — VADMU”. It contributes to the academic literature on how digital methods can enhance interdisciplinary cooperative work by exploring the collaboration involved in developing visualisations to lead decision-making in historical research in a specific interdisciplinary research setting. More specifically, we discuss how the cross-disciplinary design of a taxonomy of sources of uncertainty in Digital Humanities (DH), a “profoundly collaborative enterprise” built at the intersection of computer science and humanities research, became not just an instrument to organise data, but also a tool to negotiate and build compromises between different communities of practice.

1. Introduction and Overview

Increasingly, research policy-makers and funders are calling for inter- and transdisciplinary approaches to be applied in response to complex socio-scientific questions. Whilst the most exciting and ground-breaking innovations, to paraphrase Carlos Moedas, former European Commissioner for Research, Science and Innovation, may indeed be happening at the intersection of disciplines, this kind of cross-disciplinary collaboration poses many challenges, among which are the individual disciplinary values, concepts of proof and evidence, languages, technical cultures and methodologies employed by different domains (#bruce_etal2004; [Lowe, Whitman, and Phillipson 2009]; [Edmond et al. 2018]). It is perhaps unsurprising that overcoming the challenges associated with specialised terminology/vocabulary is often instrumental to the success of cooperative work involving individuals of different disciplinary backgrounds and expertise levels ([Skeels et al. 2008]; #bruce_etal2004). This may be, in particular, the case when the goal of the cooperation is to build a software platform, within and behind which mutual misunderstandings can become encoded, creating dissonances in structure and function, and potentially leading to difficulties for future users. This paper contributes to the academic literature on how a specific kind of technical artifact – in this case, a taxonomy of uncertainty – was used to enhance the functionality of a technical team and their co-developed platform itself created for the purpose of supporting interdisciplinary cooperative work. More specifically, we discuss how the cross-disciplinary 1 design of this taxonomy became not just an instrument to organise data, but also a tool to negotiate and build compromises between four different communities of practice, each of which came to the project with a very different epistemically 'native' approach to issues of uncertainty. In order to illustrate the many interlinked functions, the taxonomy was required to perform, we will briefly introduce the project, its aims, and the four diverse but complementary partners that came together to deliver it. We will then discuss the taxonomy itself, how it needed to be adapted from earlier, similar work, and some of the issues it allowed the project team to negotiate and discuss. In the final section, we will look at the application of the taxonomy and how it informs system function in a manner that is sensitive to the requirements of all of the disciplinary stakeholders in the project.

2. Background and Context of the Collaboration

The project in question is the PROgressive VIsual DEcision-Making in Digital Humanities (PROVIDEDH) project, which aimed to develop a web-based multimodal collaborative platform for the visual analysis of uncertainty in the Digital Humanities (DH) data analysis pipeline (discussed in further detail below). The consortium can be described as both complex and nonroutine [Strauss 1988], comprising a multinational team of four institutional partners and two overlapping sub-teams (composed of the two technical partners and two partners representing the arts and humanities). Further, the individual partner teams combine researchers and engineers of varying levels of expertise in specific branches of computer science and humanities disciplines as well as in open innovation and DH. 2

The PROVIDEDH project was originally conceptualised according to the imperatives of computer science as a discipline. The coordinator, based in the Visual Analytics and Information Visualisation Research Group (VisUsal) at the University of Salamanca (USAL), brought expertise in visual analytics, research methods, software engineering, and human-computer interaction (HCI). In addition to the project coordination activities, USAL led both the development of user interfaces and evaluation and the research for the definition and management of uncertainty in DH. The second technical partner, the Poznan Supercomputing and Networking Center (PSNC), brought extensive experience designing and building large-scale DH infrastructures, and led the development of the PROVIDEDH platform. Considering its intended user base, however, the technological expertise of USAL and PSNC alone would not have been sufficient to realise the PROVIDEDH platform. Realising the principle "that [Digital Humanities] tool builders must consider themselves as entering into a social contract with tool users" [Gibbs and Owens 2012], two further partners, Trinity College Dublin and the Austrian Academy of Science (in conjunction with the Ars Electronica Research Institute Knowledge for Humanity), were brought on to contribute from the perspective of Digital Humanities and Open Innovation Science respectively, and led research regarding user needs, user engagement, and exploitation and dissemination of results. 3

One of PROVIDEDH's first objectives was to define and describe the key types of uncertainty in DH research datasets. To date, various definitions, classifications and taxonomies of uncertainty have been proposed, but much of this work on the characterisation of uncertainty has been delivered within isolated domains. As Kouw et al. (2012) argue, typologies of this kind are valuable tools in the attempt to determine the extent to which uncertainties can in fact be explained. Perhaps more importantly from the perspective of PROVIDEDH is the authors' assertion that new forms of knowledge production that challenge our understanding of uncertainty "almost always occur at the boundaries between different disciplines." 4

3. Overcoming the Communicative Factor: Establishing a Baseline for Collaboration

As described above, the PROVIDEDH team is diverse with goals that could only be met via closely aligned collaboration. A recent report by the EU-funded SHAPE-ID project, which addressed the challenge of improving inter- and transdisciplinary cooperation between Arts, Humanities and Social Sciences (AHSS) and Science, Technology, Engineering, Mathematics, and Medicine (STEMM) disciplines, identified 25 factors that influence interdisciplinary success or failure, depending on the circumstances of a project [Vienni Baptista et al. 2020]. One of the first and primary challenges faced by the PROVIDEDH team relates to what Vienni and Baptista (among others) describe as the "Communicative" factor, which refers to the different "languages" employed by different disciplines ([Lowe, Whitman, 5 and Phillipson 2009]; [Boix Mansilla, Lamont, and Sato 2016]; [Vienni Baptista et al. 2020]). The absence of a common transdisciplinary language/terminology to discuss shared concepts such as uncertainty can substantially hinder interdisciplinary research and, conversely, processes to resolve this gap are exceptionally useful for illustrating how the development of shared intellectual artifacts can create a common ground for exchange.

Even among groups with apparently similar disciplinary backgrounds, conceptual frameworks may differ substantially ([Jeffrey 2003]; [Pennington 2008]). To cite one example, research conducted by the Knowledge Complexity (KPLEX) project found that inconsistent and contradictory statements in academic literature on such a foundational word as ‘data’ are manifold, even and indeed especially, within the field of computer science. In a series of interviews conducted with computer scientists, the project team found the trend “point[ing]towards an epistemic cultural bias towards viewing data, whatever it is, as broadly encompassing, and in terms of its function or utility in the research project, rather than a complex set of information objects that come with biases built in to them, and which might merit a certain amount of meta-reflection” [Edmond et al. 2022]. In essence, the discourse of data in computer science appears indicative of what Edward Hall would call a ‘high context’ culture in which the precise context, in which a word is used, can determine its intended meaning in a way that might be opaque to outsiders. Needless to say, if the possibility for variant understandings can be so high within a single discipline, the gaps between the humanities (a much lower context culture) and computer science surely add to the difficulty of creating a unified classification of certain phenomena. The discourse, or more accurately discourses, on key terminology such as data or uncertainty can have a serious delimiting effect on the potential of cross-disciplinary, cooperative projects: 6

”[O]ur traditional reliance on community ties to overcome the flaws in both our data and the terminology we use to speak of it do not translate well to larger scale interdisciplinary endeavours, to environments where the backgrounds or motivations of researchers/participants are not necessarily known or trusted or to environments where either the foundations of the research objects (such as is found in big data) or those of the algorithmic processing results (such as found in many AI applications) are not superficially legible to a human researcher.” [Edmond et al. 2018, p. 10]. 7

In this same report, the KPLEX project also set forth a number of recommendations for fostering fruitful interdisciplinary collaboration, including a commitment to the negotiation of key terminology and hierarchies [Edmond et al. 2018, p. 13]. As a first step then, a project must establish a baseline for collaboration. However, the requirements for this baseline, what can and cannot be taken for granted, are very much predicated on the starting position of those seeking to cooperate, which we outline briefly below. 8

3.1. The HCI Perspective

Although not specialised in the Digital Humanities, USAL brought previous experience working in this context into the PROVIDEDH project. From this experience, they were sensitised to how the increasing use of computational and visualisation artefacts in the humanities had been able to produce important results on both sides of the collaboration, but also to critical voices that had risen to point to the perils of producing visual representations that may prevent humanities researchers from performing correct critical interpretations of the analysed data [Drucker 2015]. The pernicious potential results of this effect have been at the centre of many debates among humanities and visualisation researchers and continue up to this day ([Coles 2017]; [Lamqaddam et al. 2018]; [Lamqaddam et al. 2021]). The root of this problem may be related to the abuse of black-box approaches that diminish the users’ capabilities to understand the inner workings of the algorithms at play and interpret results [Therón Sánchez et al. 2019]. Whereas the correct application of visualisation techniques is known to resolve some of these problems, there is still an urgent need for conceptual HCI and visualisation frameworks that are able to translate the needs of humanities researchers into the computational plane and use them to generate adequate user interfaces that can overcome the aforementioned problems. Along this line, uncertainty visualisation has been found necessary to effectively open the algorithmic black-box and effectively expose it to a human operator. However, the visualisation of uncertainty has been recently reconsidered in connection with how the persons actually perceive the visualisations trying to convey uncertainty, which unfortunately is not usually happening in the way the creator of the visualisation would expect [Padilla, Kay, and Hullman 2021], and much of the current work focuses on producing visualisations for evaluating decision-making in ad-hoc use-cases that can hardly be applied to a real humanities research context. To fill in this gap, the efforts presented 9

in this paper are oriented precisely towards building a theoretical visualisation framework that enables collaboration between many human and computational actors, in a temporally and spatially distributed manner.
The newly developed taxonomy fills a gap in the field of Humanities by addressing the issue of uncertainty from the perspective of Human-Computer Interaction. It lays the foundation for creating general guidelines and tools to help answer questions about when and how to communicate uncertainty in Digital Humanities workflows.

3.2. The Humanities Perspective

As one of the intended groups of end users of any tool or resource PROVIDEDH might build, it was critical that the needs and expectations of humanists would be central to the development of the project's platform. However, as a collective descriptor the term '(the) Digital Humanities' (DH) captures a broad range of research and research-related activities, sectors and stakeholders. This often obscures important differences between individual disciplines that bear on the ways they position themselves in relation to digital technology and with other disciplines. This perspective was represented in the project by Trinity College Dublin's Digital Humanities Centre, which brings together a team highly experienced in fostering the work of a range of traditional humanities disciplines, including history. This work focused on the experiences of historians and literary scholars with exposure to building and/or using digital research tools. The aim of this segmentation was first and foremost to establish the expectations of a specific intended user base, as only through such specificity could the questions of uncertainty be adequately broached. 10

Even within this context, there is great room for variation in interpretation and practices. As the user modelling work of the DH@TCD team showed, a historian may, on one level, question her ability to read a historic text or question the date attributed to a certain event within an earlier interpretation. She may also, however, question whether an event was understood at the time as it is generally described now, or indeed whether the records of a given event have become corrupted or biased over time. The novelty and complexity of DH research combined with the idiosyncratic challenges of working with cultural heritage data mean that uncertainty can hinder substantially in the use and reuse of digital cultural resources, which might make it more difficult for a user to probe the context or provenance of a given fact or interpretation. Throughout the DH analysis pipeline, many decisions have to be made which depend on managing uncertainty, pertaining to both the datasets and the models behind them. Traditional humanities tools for managing uncertainty, such as corroboration among multiple sources, or inclusion of footnotes or explications to note a point of uncertainty, are themselves built on the preservation of complexity and scrutability. This is an epistemic cultural bias that can lead to resistance to taxonomy-based approaches, as the reduction of parameters to fit a simplified model can be perceived as an opening to un- or misinformed conclusions. Yet without such tools, these disciplines cannot benefit from purpose-built ancillary tools in the form of interactive visualisations or novel user interfaces in the digital research environment. The successful development of a fit-for-purpose digital research environment – defined here as the adoption of the tool by the intended user base – is therefore highly dependent on the correct categorisation of the different types of heterogenous phenomena, including uncertainty, within DH. 11

3.3. The Open Innovation Perspective

The PROVIDEDH project and its taxonomy were not intended to have an impact only in historical research but also more widely as a way to understand the complex, tacit, or compound forms of uncertainty. Under the paradigm of Open Innovation, inflows and outflows of knowledge are positioned with the purpose to advance and expand possibilities of external through stakeholders. Thus, the boundaries between sourcing and the utilisation of an organisation's external environment become flexible with more opportunities for innovation. The application of inbound Open innovation in the Digital Humanities allows organisations to reframe knowledge behind activities that recombine, search and capture technologies in a prompt manner that on-board current and urgent issues. Open Innovation becomes a social process derived from key factors such as declining knowledge hegemony, mobility of workers and growing influence of start-ups in supporting information and communication technologies. In other words, Open innovation provides new channels to knowledge transfer and dissemination that ignite broader accessibility and diffusion of knowledge. The presence of Open Innovation practices creates new areas to stretch outreach that currently does not fit in existing subfields. As innovation is not a linear process, stakeholders feed into the development of work combining physical and digital engagements that centre people within new models that adapt to new bubbling markets. 12

Open Innovation goes across the boundaries of organisations touching into a wide set of fields, such as low-tech industries, small and medium-sized enterprises (SMEs), or not-for-profit organisations. Thus, the linkage of policies with innovation and science could help close gaps of uneven growth in productivity and prosperity by opening datasets along with public/private partnerships to enhance universities' relations with industries. Additionally, policies with an Open Innovation perspective embrace uncertainty not only by new funding schemes but by intersecting programmes for frontier science where complex problems are confronted. 13

The research reads, applies, and interacts with the apparatus that conforms system transitions in everyday life through a set of intersectional theories across a variety of points of design with the goal to demonstrate the complex interactions shaping and impacting our understanding of uncertainty within data and data collection practice. To answer these questions, the Austrian project team undertook a number of events and interventions to explore possible approaches that prioritise collaboration around uncertainty in data as a socialising experience challenging the current status quo. 14

The taxonomy developed in the project supports the Open Innovation paradigm by making it possible to explicitly express information about data uncertainty using characteristics relevant to the Digital Humanities and similar fields. The expression of the information is not only theoretical but can be applied in practice using the well-known and accepted TEI guidelines, as described in the following section. 15

3.4. The Standards Perspective (TEI)

To resolve the tensions raised by USAL in a manner able to meet the needs of both the professional historian and the citizen scientist, a particular development approach would be required. As with USAL, in recent years, researchers at PSNC had been involved as technical partners in a number of Digital Humanities projects. Through these projects in particular, they developed a sensitivity to the use of the TEI (Text Encoding Initiative) standard as an accepted mechanism for providing a common structure to digital editions and other digital artefacts created by humanistic researchers. The TEI standard enables simple text processing, corpus linguistic queries and other quantitative approaches to the texts. As such, understanding how the TEI might contribute to an accepted and ultimately sustainable expression of uncertainty seemed an exceptionally fertile approach. 16

Although the TEI specification includes several mechanisms to express uncertainty or precision (e.g., the <certainty> element, the <precision> element, or the @cert attribute), using them is not a common practice. In an attempt to establish a baseline understanding of precisely how well used they were, project team members used their presentation slot at the TEI 2019 conference [Kozak et al. 2021] to ask an audience of nearly 100 TEI users how many of them used this markup in their research work: only one person responded positively. 17

Clearly, the challenge is not just about having tags in the most widely accepted annotation standard to manage uncertainty in the humanities research process. It is about managing uncertainty within data and facilitating approaches to make it more explicit. To harness the affordances already inherent in the standard, however, the PROVIDEDH project focused its platform-building efforts on TEI-encoded texts, creating a collaborative platform *(https://providedh.ehum.psnc.pl)* that allows users to load their TEI files and analyse them in many ways. The platform also allows users to enrich the annotation layer of files with new entities or doubts about existing (annotated) entities and fragments. Furthermore, the annotation scenarios implemented in the platform allow the TEI uncertainty annotation specification to be tested and ultimately expanded by the project. 18

For the TEI guidelines, the taxonomy and the project itself served as a test case. In fact, the internal project requirements to express specific uncertainty characteristics, with the usage of TEI, initiated a number of discussions on how to improve the TEI guidelines themselves. Some of the discussions have already led to changes in the TEI guidelines (e.g., categories of uncertainty - *https://github.com/TEIC/TEI/issues/1934*) and some of them are ongoing (e.g., expanding possibilities of TEI tags - *https://github.com/TEIC/TEI/issues/2067*). 19

3.5. Assessment of the Collaborative Team

As the above profiles indicate, the four PROVIDEDH partners came to the project with aligned but still divergent goals 20 for the project and different understandings of uncertainty in the research process. This range of attitudes towards the fundamental concerns of the PROVIDEDH project (such as what we can tolerate in terms of uncertainty and what we need to participate in decision making, as shown in Figure 1 below) made for a challenging, if potentially very rich, interdisciplinary exploration of the limits of computer supported cooperative work.

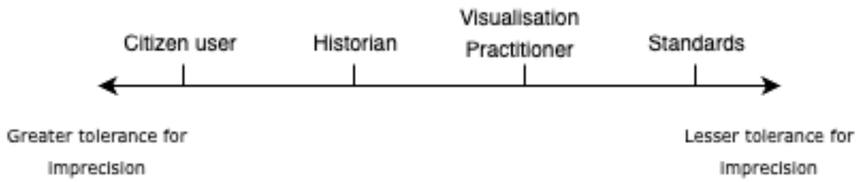

Figure 1. Ranked order of epistemic cultures in PROVIDEDH with regards to tolerance for imprecision.

These four perspectives, each representing the expertise and biases of one of the diverse partners in the PROVIDEDH project, have each enabled certain kinds of insight to be developed in the project. This was a strength of the collaborative team, but could also be a weakness, as gaps in the boundary languages [Bowker and Star 2000] between disciplines can often remain hidden for too long. In particular, we uncovered gaps not only in understanding, but also in epistemic cultures and values between the members of the team, which would need to be managed if the team was to achieve its aims in a way that could be validated equally by the whole multidisciplinary team. 21

Given the risk inherent in the variations exposed in Figure 1, the team realised it would be required to negotiate and place at the foundation of any system it would build a formally agreed mechanism by which to encode a shared interpretation of such key terminology, in particular the central issue of uncertainty. To do this, the partners co-created a formal taxonomy that seemed to hold the most potential, as it would both expose hidden disagreements and be able to form a backbone for the annotation and standardisation processes foreseen for the platform. For this, we were able to begin from significant work. 22

4. Developing the Taxonomy

In recent years, the graphical display of uncertainty has become an important trend in visualisation research ([Kay et al. 2016]; [Hullman et al. 2018]; [Kale et al. 2018]; [Hullman 2020]). For this, Digital Humanities represents an exciting field of experimentation [Benito-Santos and Therón Sánchez 2020]. In this regard, the first notable efforts to address the visual representation of uncertainty originated in the fields of geography and spatial data visualisation with proposals describing uncertainty taxonomies whose categories could be mapped to the different available visual channels ([MacEachren 1992]; [Fisher 1999]). For example, MacEachren’s employed the concept of “quality” to build his first uncertainty taxonomy, and accounted for the different manners in which uncertainty could be introduced into the data according to the stages of the typical data analysis pipeline. According to this, he described concepts such as accuracy – the exactness of data –, and precision – the degree of refinement with which a measurement is taken. Later in the text, MacEachren mapped these concepts to several visual variables (e.g., saturation), building on the famous previous work on semiology by the French cartographer Jacques Bertin [Bertin 1983]. In more recent work, MacEachren completed his previous work with two empirical studies focused on uncertainty visualisation in which a typology of uncertainty was matched against the spatial, temporal and attribute components of data [MacEachren 2012]. 23

	Open Innovation	Humanities	HCI	Standards
Definition of Uncertainty	Uncertainty can refer to a state of not knowing, not understanding, or not being aware of something; having incomplete or wrong information; not remembering certain information, or not knowing a certain outcome.	Uncertainty is a necessary condition for all humanistic research that can only be managed, never removed.	Uncertainty is a state of limited knowledge (imperfect or unknown information) where it is impossible to exactly describe the existing state or future outcome(s).	Uncertainty is a lack of complete true knowledge about something.
Sources of Uncertainty	A more or less complex process involving one or more actors that are working towards a common goal.	As a part of the research process, a mechanism by which to weigh possible interpretations and choose avenues to explore and attempt to corroborate.	The cognitive process resulting in the selection of a belief or a course of action among several possible alternative options.	A process in which the beings decide what to do in the near or far future in some aspects of their life

Table 1. Variant definitions of Uncertainty and Decision-Making among the PROVIDEDH project partners.

In a different approach, Fisher [Fisher 1999] proposed an uncertainty taxonomy elaborating the concepts of well- and poorly-defined objects, which are used to define the categories into which uncertainty can be classified. More concretely, Fisher's taxonomy presents a particular point of view within the more verbose taxonomy of ignorance developed by [Smithson 1989]. According to Fisher's views, uncertainty related to well-defined objects is caused by errors and is probabilistic in nature. This type of uncertainty is typically referred to as "aleatoric" or "irreducible". On the contrary, poorly defined objects relate to the concept of epistemic uncertainty, which can in fact be reduced. More recently, Simon and his colleagues adapted Fisher's taxonomy in their work on numerical assessment of risk and dependability [Simon, Weber, and Sallak 2018], which in turn served us a starting point to propose our own uncertainty taxonomy in the context of Digital Humanities research. Simon et al.'s adaptation of Fisher's taxonomy is reproduced in Figure 2 below: 24

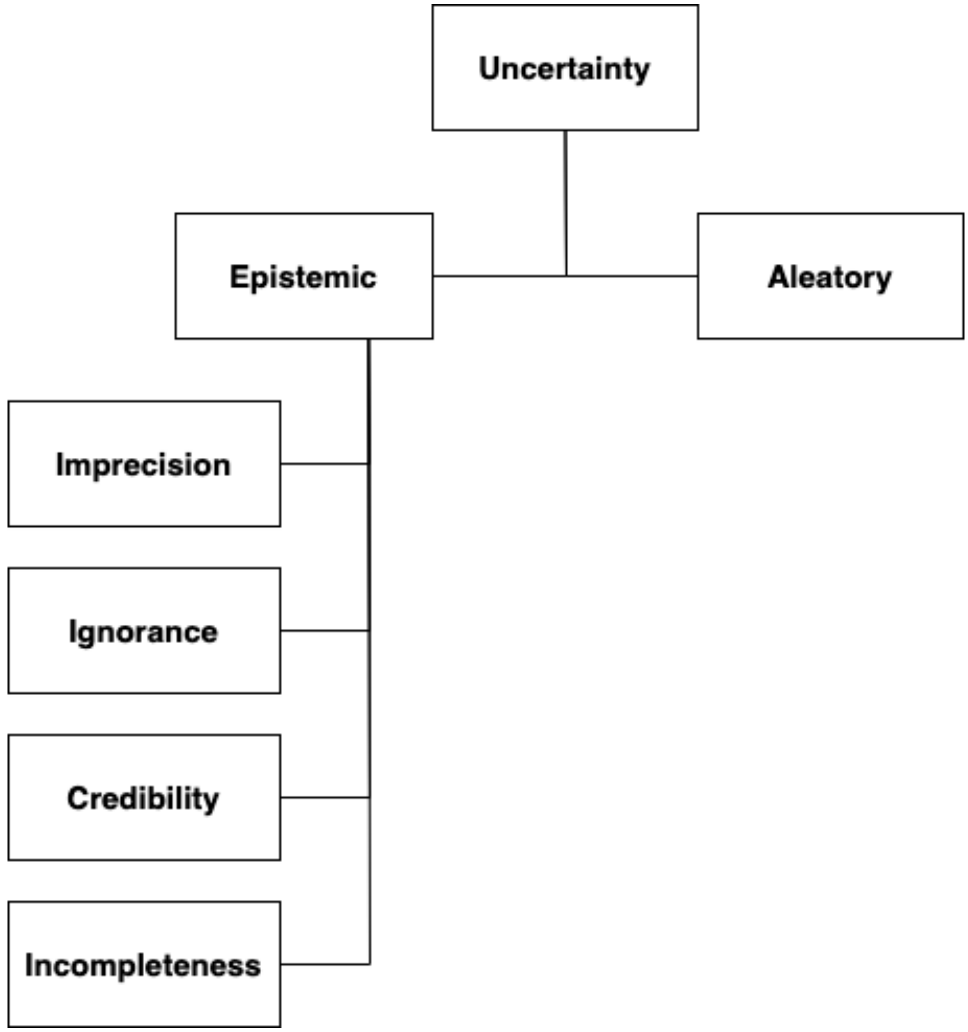

Figure 2. Simon et al.'s adaptation of Fisher's taxonomy of uncertainty.

The original uncertainty taxonomy by Simon et al. includes four more categories into which epistemic uncertainty might be classified: imprecision, ignorance, credibility and incompleteness. According to the authors, "imprecision corresponds to the inability to express the true value because the absence of experimental values does not allow the definition of a probability distribution or because it is difficult to obtain the exact value of a measure." The next category, ignorance (which can be partial or total), refers to "the inability to express knowledge on disjoint hypotheses." Incompleteness makes reference to "the fact that not all situations are covered." This is especially applicable to Digital Humanities research objects which may have been degraded or partially lost due to natural processes or the different transformations that may have been applied to the originals. Finally, credibility concerns "the weight that an agent can attach to its judgement," [Jarlbrink and Snickars 2017] and it is also of special importance to the aims of our project. For example, a researcher may or may not trust the observations made by other agents who have previously worked on the same data, according to her degree of grounded knowledge on the subject and personal preference and/or bias. 25

Although a strong starting point, none of the existing taxonomies was able to satisfy all of the partners in PROVIDEDH, as the balance between refinement and actionability was, in each case, never mutually recognisable. As Skeels and his colleagues noted in 2008, uncertainty is referred to inconsistently both within and among domains, they continue that "[w]ithin the amorphous concept of uncertainty there are many types of uncertainty that may warrant different visualisation techniques." [Skeels et al. 2008] The challenge is, therefore, not to arrive at a single definition that obscures differences of uncertainty, but to build dialogue between different research domains while recognising their differences, and ensuring that any definition of uncertainty the project as a whole would adopt could be applied and implemented by all partners in their own research contexts. This challenges the idea that it is possible to create a single view of uncertainty. "While any group's ontology is unlikely to match that of every individual within the group, the extent 26 of mismatch tends to increase with the scale of the group and the differences between the purpose of individual and group ontologies” [Wallack and Srinivasan 2009, p. 2]. The reality is that the development of a taxonomy, which must cater to the needs of a range of disciplines and stakeholders, may often lead to a result in which the specific requirements of not one subject area or stakeholder group are fully met, and yet which all can still use at some level of granularity.

Building on these insights, the group organised a series of shared workshops exploring the issue of conceptions of uncertainty and its relationship to decision-making in different contexts. These workshops commenced from existing taxonomies, such as Fisher’s (reproduced above), and sought to revise them so as to reach consensus, if not agreement, about how conceptions of uncertainty might be made concrete and actionable in a historical research system. The three workshops each approached a different aspect of the overall challenge, focusing first on the historian’s requirements for decision-making as compared to the system's needs for clarity and simplicity. The second continued this discussion with a wider and more diverse group of potential users, and the third looked at the same challenge from the perspective of open innovation. Robust discussions and negotiations about the taxonomy occurred at every one of these discussion points. For example, in the interaction designed to bridge what was perceived as a back-end processing ‘optimal’ with the perspective of the historian, nearly every term had to be renamed, glossed and refined, with, for example, the descriptor of ‘error’ being rejected as lacking sufficient precision. The domain experts also strongly advocated for the recognition of time-based element tracing the possible source of uncertainty from the original moment of a source’s creation through the many transformations over time (physical damage, cataloguing, transcribing, translating, digitisation, etc.) all the way through to the task of the researcher itself. Although later discussions continued to be active and important for the purpose of converging both understanding and usage patterns (in particular the meeting to discuss how to further extend the taxonomy’s use in the Open Innovation context), fewer changes were implemented over time. 27

Departing from the taxonomies proposed by [Fisher 1999] and Simon et al. [Simon, Weber, and Sallak 2018] that were introduced in the previous section, the project partners developed an HCI-inspired uncertainty taxonomy aimed at producing effective visualisations from humanities textual data and related annotations (see Figure 4). Findings in these workshops suggested that a matrix-based taxonomy would be of the most use to the project and the potential users of its outputs. This substantiated the earlier work conducted in the context of the KPLEX project (discussed above) which suggested that something like Peterson’s uncertainty matrix for simulation uncertainties in climate change [Petersen 2012] would provide the most effective approach for all sides. The key differentiator was the location or site of uncertainty. For example, Pang and his collaborators describe the different processes through which uncertainty can be introduced into data as: “the introduction of data uncertainty from models and measurements, derived uncertainty from transformation processes, and visualisation uncertainty from the visualisation process itself” [Pang, Wittenbrink, and Lodha 1997]. Petersen similarly distinguishes a number of locations of uncertainty: see his model below in Figure 3. 28

Location		Level			Nature	
		Statistical uncertainty	Scenario uncertainty	Recognised ignorance	Epistemic uncertainty	Variability uncertainty
Context	Natural, technological economic, social and political, representation					
Model	Model structure					
	Technical model					
Inputs	Driving forces					
	System data					
Parameters						
Model Outcomes						

Figure 3. Petersen's Uncertainty Matrix [Petersen 2012].

The matrix shown in Figure 4 was built as a result of the PROVIDEDH development process. 29

Location	Epistemic Uncertainty				Aleatoric Uncertainty
Researcher Task	**Ignorance** (lack of context)	**Non-credibility** (error or bias seen to be entering the system)	**Variation** (differing values within a contested category)	**Gaps** (outlier record where information is missing)	**Unavoidable** (due to positionality or natural phenomena)
Processing or Transformation					
Primary Source					

Figure 4. The PROVIDEDH Taxonomy.

Although it maintains the overall simplicity of having four categories of uncertainty (about which much negotiation of terminology was required) the final taxonomy also incorporates a number of other dimensions deemed important for the notation of uncertainty. First, on the y-axis, it incorporates a sensitivity to time and the possible actors that might have introduced uncertainty into a system. This is particularly important for historical research, as an uncertainty introduced by a more recent process (such as dirty OCR of historical texts) may be much easier to manage in a decision-making process than an original gap in what was seen or reported. Similarly, the matrix follows Fisher in applying the concepts of epistemic and aleatory uncertainty, with the former more likely to be a result of human lack of understanding, and the latter more likely to be one of imperfect sources and tools. Truly epistemic uncertainty is the hardest to address, but, as we can see in the taxonomy, there is only one of twelve possible states of uncertainty that is indeed fully epistemic, along with five others that are partially so. Knowing how uncertain a given uncertainty is, makes it much easier to build systems and annotation standards to accommodate it, but it also allows historians and citizen users to maintain their sense of uncertainty itself which is often a complex and potentially hybrid state. Through this instrument, we can see differences between those places that are likely to have wide agreement as to the nature of the uncertainty and what it would take to optimally resolve it (a gap in a primary source), and those areas where the questions are much more individualised, closer to the work of a single historian. 30

5. Implementing the Taxonomy

The final taxonomy was deployed as a fundamental element in the PROVIDEDH demonstrator platform for the 31
investigation of uncertainty in historical texts. Its primary functionality was to structure the manner in which users could indicate the nature and level of uncertain passages in the text and to visualise patterns in these annotations. This enabled a like-with-like comparison of annotations across users and documents. The use of the taxonomy in the system (for example in the visual annotator, see Figure 5, is based upon a conceptualisation of uncertainty that has two well-defined sources: 1) tool-induced and 2) human-induced, which can be respectively mapped to the two main categories devised by [Fisher 1999] that were introduced in previous sections of this paper.

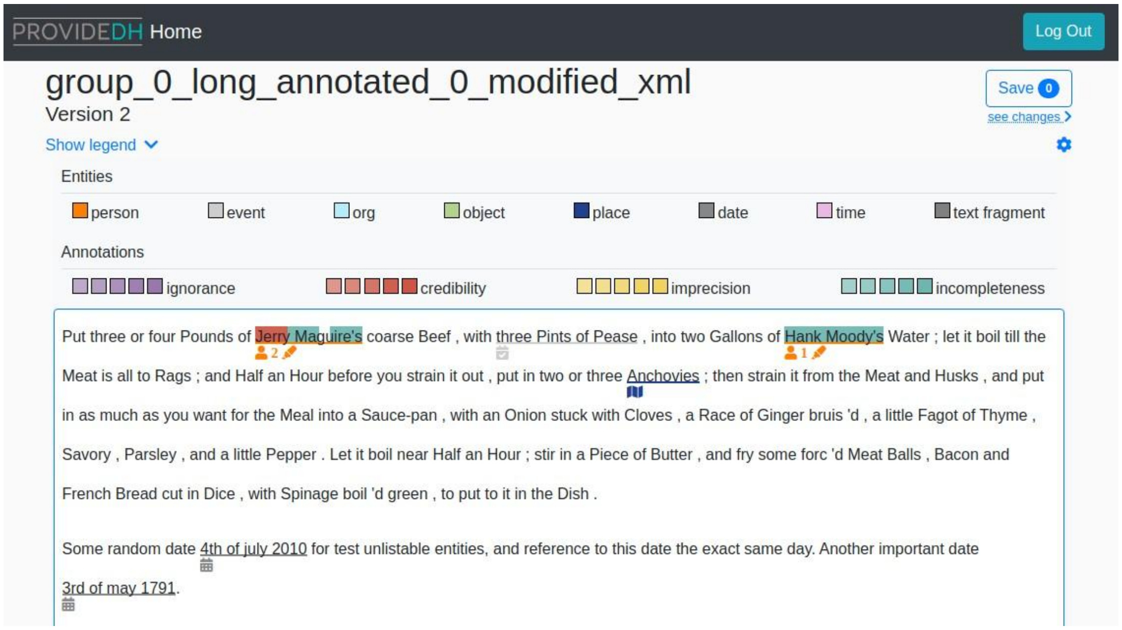

Figure 5. Screenshot of the visual annotator that was implemented in the project's platform. Here, users can annotate texts using at least four categories of epistemic uncertainty (ignorance, non-credibility, imprecision and incompleteness).

The first uncertainty category, tool (or machine)-induced (Figure 5, bottom), can be mapped to aleatoric uncertainty 32
(well-defined objects in Fisher's taxonomy) and results from the application of computational algorithms (or other historic tools) to the data. These often give their results with a variable degree of bounded uncertainty (e.g., topic models [Blei 2012] or word embeddings [Mikolov et al. 2013]). For this reason, this type of uncertainty is better represented as a continuous probability distribution.

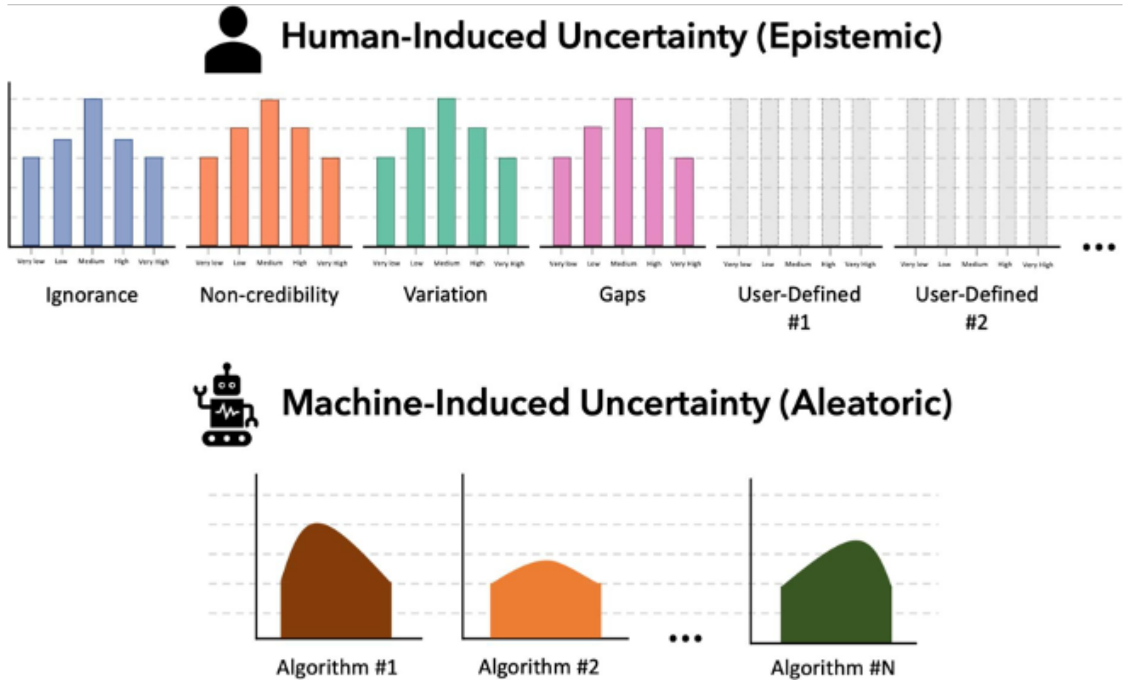

Figure 6. Proposed n+1 uncertainty taxonomy model. Top: human-Induced category with four predefined categories that map to the epistemic categories previously introduced by Simon and Fisher. To the right, user-defined categories can be added during the project's lifetime. Users can add more categories if required. Bottom: machine-induced uncertainty showing different runs of the same or different algorithms.

In addition, this representation allows a better understanding of speculative runs of a given algorithm and enhances the what-if analysis process. For example, a researcher could parametrise an algorithm with a fixed set of inputs and launch it a number of times, obtaining a range of mean values and deviations encoded in a document which, if correctly displayed, would allow her to get an idea of how the algorithm behaves. Analogously, the algorithm could be parametrised with a variable set of inputs created by the user running the computation or, as we will see next, by other past researchers. This operation mode would provide an answer to the question "What happens if I run the algorithm a million times using my assumptions?" or "What happens if I run the algorithm a million times using another person's assumptions?". As in the case of running the algorithm with the same parameters many times, the results of running the algorithm with different parameters many times could also be summarised in a continuous document allowing the said kind of what-if analysis. This would answer the question of "What happens if I use these other person's assumptions?" or "What happens if I decide not to trust another actor's assumptions?". We argue that the insights obtained from the described experiments are highly valuable for the user, specifically in the case of algorithms that are probabilistic in nature, such as topic models, and whose results – and interpretations – can vary greatly between different runs [Alexander and Gleicher 2016]. 33

The other category, human-induced uncertainty (bottom), arises from 1) direct interpretations of the raw data (which in turn may be based on others' previous interpretations and grounded expert knowledge of the user) and 2) interpretations of computational analyses performed on the data or 3) a combination of the two. This category is reported by users of our system in a 5-point Likert scale and thus it is best modelled as a discrete probability function. Moreover, we devised the dependency relationships between the categories in our taxonomy to be bidirectional and self-recurring, since, for example, the input parameters and data fed to the algorithms – and therefore their results – are derived from a user's previous interpretations of textual data and related machine- or human-generated annotations. In turn, these interpretations must necessarily be built upon previous insight obtained by the same or other users who apply computational techniques to the data, effectively forming a temporal belief network [Druzdzel and Simon 1993] that is fixated on the different versions of a dataset. In a first stage, we modelled human-induced uncertainty according to the uncertainty taxonomy of Figure 4 in what was called the 4+1 uncertainty model. Later, and following a series of evaluation sessions and workshops, we evolved the model to support a variable number of human-induced categories. This change was introduced after a series of user-driven experiments and evaluations in which we asked the 34 participants to annotate different types of historical texts (poetry and historical recipes). As we describe in our previous contribution, the uncertainty taxonomy was received positively by the Digital Humanities experts and lay users we engaged with during the design sessions that took part throughout the project [Benito-Santos et al. 2021]. From these experiences, we could check first-hand the latent need, as perceived by the participants, for a more nuanced and flexible way of communicating uncertainty in Digital Humanities projects, as well as the potential benefits that this could bring for collaborative work on Digital Humanities datasets. Specifically, we discovered that, sometimes, users could not pick one of the predefined categories when making annotations on the texts. Furthermore, we also detected that there almost never was a 100% consensus on the uncertainty type different users chose to annotate the same portion of text: e.g., whereas for some a linguistic variant of a word implied imprecision (e.g., they were more or less sure about a word's meaning, but the word varied from a standard form), for others it meant ignorance (e.g., they did not know what the word meant). These differences, we argue, are often subtle and very much depend on the mental state, ground knowledge on the subject and other traits of the particular user making the annotation. Beyond that, some users showed preference for labelling uncertainty in a particular annotation as a combination of different types. For example, a user could be more or less sure that a word found in the text represents a dialectal variation but at the same time not be sure of the word's meaning in a certain context. In this case, the user should be able to reflect her views using a combination of categories (e.g., imprecision + ignorance + 0/5 non-credibility + 0/5 gaps). This way of communicating uncertainty, we argue, allows a higher level of expressiveness and is able to capture the nuances of the user's mental state that needs to be captured and presented to other users working in the same dataset. To allow for even better communication of uncertainty between actors, we allow users to dynamically create new categories as required, which are established on a per project-basis in what is called the n+1 uncertainty model. This departs from the observed effect in our experiments that a group of users working on the same data will naturally establish their own uncertainty taxonomy that covers the needs and particularities of the specific tasks to be performed on the data. As such, they will use it as a communication tool to collaborate in a distributed manner in time and space by appending their annotations to the dataset according to these ad-hoc categories, effectively fulfilling one of the project's main objectives.

6. Conclusion

The function of the system ultimately built by the PROVIDEDH team, as a visualisation supported tool for historical research under conditions of uncertainty, was predicated upon the deployment of the taxonomy that underpins it. Interestingly, however, it was also the process of negotiating the taxonomy that allowed the team to function, overcoming the challenges inherent in interdisciplinary work and achieving the goals of their collaborative project. Lee described the effect of "boundary-negotiating artifacts" as twofold, having (1) an effect on the recipient, for whom information is more easily grasped, and 2) an effect on the artifact creator, who must strive to construct an artifact that is easily grasped [Lee 2007]. The results of the PROVIDEDH project can certainly attest to this effect and to the largely invisible, but essential, taxonomy and its development delivered. 35

Acknowledgements

This research was funded by the CHIST-ERA programme under the following national grant agreements PCIN-2017-064 (MINECO Spain), 2017/25/Z/ST6/03045 (National Science Centre, Poland), national grant agreement FWF (Project number I 3441-N33, Austrian Science Fund, Austria). 36

Alejandro Benito-Santos acknowledges support from the postdoctoral grant 'Margarita Salas', awarded by the Spanish Ministry of Universities. 37

Conflict of Interest

The authors declared that they have no conflict of interest. 38

Works Cited

Alexander and Gleicher 2016 Alexander, E., amd Gleicher, M. (2016) "Task-Driven Comparison of Topic Models", *IEEE Transactions on Visualization and Computer Graphics*, 22(1), pp. 320-329. {DOI: 10.1109/TVCG.2015.2467618}

Benito-Santos and Therón Sánchez 2020 Benito-Santos, A., and Therón Sánchez, R. (2020) “A Data-Driven Introduction to Authors, Readings, and Techniques in Visualization for the Digital Humanities”, *IEEE Computer Graphics and Applications*, 40(3), pp. 45-57. {DOI: 10.1109/MCG.2020.2973945}

Benito-Santos et al. 2021 Benito-Santos, A., Doran, M., Rocha, A., Wandl-Vogt, E., Edmond, J., and Therón, R. (2021) “Evaluating a Taxonomy of Textual Uncertainty for Collaborative Visualisation in the Digital Humanities”, *Information*, 12(11), p. 436. {Avaialble at: https://www.mdpi.com/2078-2489/12/11/436 or DOI: 10.3390/info12110436 (Accessed 21 October 2021).}

Bertin 1983 Bertin, J. (1983) *The semiology of graphics: Diagrams, networks, maps*. Translated by W. J. Berg. Madison: University of Wisconsin Press.

Blei 2012 Blei, D. M. (2012) “Probabilistic Topic Models”, *Commun. ACM*, 55(4), pp. 77–84. {Available at: http://doi.acm.org/10.1145/2133806.2133826 (Accessed 31 October 2018).}

Boix Mansilla, Lamont, and Sato 2016 Boix Mansilla, V., Lamont, M., and Sato, K. (2016) “Shared Cognitive–Emotional–Interactional Platforms: Markers and Conditions for Successful Interdisciplinary Collaborations”, *Science, Technology, and Human Values*, 41(4), pp. 571-612. {Available at: https://doi.org/10.1177/0162243915614103 (Accessed 08 February 2022).}

Bowker and Star 2000 Bowker, G. C., and Star, S. L. (2000) *Sorting Things Out: Classification and Its Consequences*. Cambridge: MIT Press.

Coles 2017 Coles, K. (2017) “Think Like a Machine (or not)”, *Proc. 2nd Workshop on Visualization for the Digital Humanities (VIS4DH)*, pp. 1-5.

Drucker 2015 Drucker, J. (2015) “Graphical Approaches to the Digital Humanities” in *A New Companion to Digital Humanities*. Hoboken: Wiley-Blackwell, pp. 238-250. {Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118680605.ch17 (Accessed 29 October 2018)}

Druzdzel and Simon 1993 Druzdzel, M. J., and Simon, H.A. (1993) “Causality in Bayesian Belief Networks” in Heckerman, D., and Mamdani, A. (eds.) *Uncertainty in Artificial Intelligence*. Burlington: Morgan Kaufmann, pp. 3-11. {Available at: https://www.sciencedirect.com/science/article/pii/B9781483214511500056 (Accessed 04 February 2022).}

Edmond et al. 2018 Edmond, J., Doran, M., Horsley, N., Huber, E., Kalnins, R., Lehman, J., Nugent-Folan, G., Priddy, M., and Stodulka, T. (2018) *Deliverable Title: D1.1 Final report on the exploitation, translation and reuse potential for project results*. {Available at: https://hal.archives-ouvertes.fr/hal-01842365 (Accessed 04 February 2022).}

Edmond et al. 2022 Edmond, J., Horsley, N., Lehmann, J., and Priddy, M. (2022) “Power through datafication” in *The Trouble With Big Data : How Datafication Displaces Cultural Practices*. {1st edn.} London: Bloomsbury Academic, pp. 135-157. {Available at: http://www.bloomsburycollections.com/book/the-trouble-with-big-data-how-datafication-displaces-cultural-practices/ch7-power-through-datafication/ (Accessed 13 December 2021).}

Fisher 1999 Fisher, P. (1999) “Models of uncertainty in spatial data”, *Geographical information systems*, vol. 1, pp. 191–205.

Gibbs and Owens 2012 Gibbs, F., and Owens, T. (2012) “Building Better Digital Humanities Tools: Toward broader audiences and user-centered designs”, *Digital Humanities Quarterly*, 006(2) {Available at: https://www.digitalhumanities.org/dhq/vol/6/2/000136/000136.html}.

Hullman 2020 Hullman, J. (2020) “Why Authors Don't Visualize Uncertainty”, *IEEE Transactions on Visualization and Computer Graphics*, 26(1), pp. 130-139. {DOI: 10.1109/TVCG.2019.2934287}

Hullman et al. 2018 Hullman, J., Qiao, X., Correll, M., Kale, A., and Kay, M. (2018) “In Pursuit of Error: A Survey of Uncertainty Visualization Evaluation”, *IEEE Transactions on Visualization and Computer Graphics*, p. 1. {DOI: 10.1109/TVCG.2018.2864889}

Jarlbrink and Snickars 2017 Jarlbrink, J., and Snickars, P. (2017) “Cultural heritage as digital noise: nineteenth century newspapers in the digital archive”, *Journal of Documentation*, 73(6), pp. 1228-1243. {Available at: https://doi.org/10.1108/JD-09-2016-0106 (Accessed 12 May 2023)).}

Jeffrey 2003 Jeffrey, P. (2003) “Smoothing the Waters: Observations on the Process of Cross-Disciplinary Research Collaboration”, *Social Studies of Science*, 33(4), pp. 539-562. {Available at: https://doi.org/10.1177/0306312703334003 (Accessed 08 February 2022).}

Kale et al. 2018 Kale, A., Nguyen, F., Kay, M., and Hullman, J. (2018) “Hypothetical Outcome Plots Help Untrained Observers Judge Trends in Ambiguous Data”, *IEEE Transactions on Visualization and Computer Graphics*, p. 1. {DOI:

10.1109/TVCG.2018.2864909}

Kay et al. 2016 Kay, M., Kola, T., Hullman, J. R., and Monson, S. A. (2016) “When (Ish) is My Bus?: User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems”, *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pp. 5092–5103, New York, NY, USA, 2016. New York: ACM. {Available at: http://doi.acm.org/10.1145/2858036.2858558 (Accessed 18 September 2018).}

Kozak et al. 2021 Kozak, M., Rodríguez, A., Benito-Santos, A., Therón, R., Doran, M., Dorn, A., Edmond, J., Mazurek, C., and Wandl-Vogt, E. (2021) “Analyzing and Visualizing Uncertain Knowledge: The Use of TEI Annotations in the PROVIDEDH Open Science Platform”, *Journal of the Text Encoding Initiative*, Issue 14. {Available at: https://journals.openedition.org/jtei/4239 (Accessed 06 April 2023).}

Lamqaddam et al. 2018 Lamqaddam, H., Brosens, K., Truyen, F., Beerens, J., Prekel, I., and Verbert, K. (2018) “When the Tech Kids are Running Too Fast: Data Visualisation Through the Lens of Art History Research”, *Proc. 3rd Workshop on Visualization for the Digital Humanities (VIS4DH)*.

Lamqaddam et al. 2021 Lamqaddam, H., Vandemoere, A., VandenAbeele, V. Brosens, K., and Verbert, K.“Introducing Layers of Meaning (LoM): A Framework to Reduce Semantic Distance of Visualization In Humanistic Research”, *IEEE Transactions on Visualization and Computer Graphics*, 27(2), pp. 1084-1094. {DOI: 10.1109/TVCG.2020.3030426}

Lee 2007 Lee, C. P. (2007) “Boundary Negotiating Artifacts: Unbinding the Routine of Boundary Objects and Embracing Chaos in Collaborative Work”, *Computer Supported Cooperative Work (CSCW)*, 16(3), pp. 307-339. {Available at: https://doi.org/10.1007/s10606-007-9044-5 (Accessed 12 May 2023).}

Lowe, Whitman, and Phillipson 2009 Lowe, P., Whitman, G., and Phillipson, J. (2009) “Ecology and the social sciences”, *Journal of Applied Ecology*, 46(2) pp. 297-305. {Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2664.2009.01621.x (Accessed 08 February 2022).}

MacEachren 1992 MacEachren, A. M. (1992) “Visualizing Uncertain Information”, *Cartographic Perspectives*, 0(13), pp. 10-19. {Available at: http://www.cartographicperspectives.org/index.php/journal/article/view/cp13-maceachren or 10.14714/CP13.1000 (Accessed: 14 February 2018)}

MacEachren 2012 MacEachren, A. M., Roth, R. E., O'Brien, J., Li, B., Swingley, D., and Gahegan, M. (2012) “Visual Semiotics Uncertainty Visualization: An Empirical Study”, *IEEE Transactions on Visualization and Computer Graphics*, 18(12), pp. 2496-2505. {DOI: 10.1109/TVCG.2012.279}

Mikolov et al. 2013 Mikolov, T., Sutskever, I., Chen, K., Scorrado, G., and Dean, J. (2013) “Distributed Representations of Words and Phrases and their Compositionality” in Burges, C.J.C., bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K. Q., (eds.) *Advances in Neural Information Processing Systems 26*. Red Hook: Curran Associates, Inc. pp. 3111–3119. {Avilable at: http://papers.nips.cc/paper/5021-distributed-representations-of-words-and-phrases-and-their-compositionality.pdf}

Padilla, Kay, and Hullman 2021 Padilla, L., Kay, M., Hullman, J. (2021) “Uncertainty Visualization” in *Wiley StatsRef: Statistics Reference Online*. Hoboken: John Wiley and Sons, Ltd, pp. 1-18. {Avilable at: https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118445112.stat08296 (Accessed 04 February 2022)}

Pang, Wittenbrink, and Lodha 1997 Pang, A. T., Wittenbrink, C. M., and Lodha, S. K. (1997) “Approaches to uncertainty visualization”, *The Visual Computer*, 13(8), pp. 370-390. {Available at: https://link.springer.com/article/10.1007/s003710050111 (Accessed: 14 February 2018).}

Pennington 2008 Pennington, D. D. (2008) “Cross-Disciplinary Collaboration and Learning”, *Ecology and Society*, 13(2). {Available at: https://www.jstor.org/stable/26267958 (Accessed 08 February 2022).}

Petersen 2012 Petersen, A.C. (2012) *Simulating Nature: A Philosophical Study of Computer-Simulation Uncertainties and Their Role in Climate Science and Policy Advice, Second Edition*. Boca Raton: CRC Press.

Simon, Weber, and Sallak 2018 Simon, C., Weber, P., and Sallak, M. (2018) *Data Uncertainty and Important Measures*. Hoboken: John Wiley and Sons.

Skeels et al. 2008 Skeels, M., Lee, B., Smith, G., and Robertson, G. (2008) “Revealing Uncertainty for Information Visualization”, *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 376–379, New York, NY, USA. New York: ACM. {Available at: http://doi.acm.org/10.1145/1385569.1385637 (Accessed 27 October 2018).}

Smithson 1989 Smithson, Michael. (1989) *Ignorance and uncertainty: emerging paradigms*. New York: Springer-Verlag.

Strauss 1988 Strauss, A. (1988) “The Articulation of Project Work: An Organizational Process”, *The Sociological Quarterly*, 29(2), pp. 163-178 {Available at: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1533-8525.1988.tb01249.x

(Accessed 03 May 2023).}

Therón Sánchez et al. 2019 Therón Sánchez, R., Benito-Santos, A., Santamaría Vicente, R., and Losada Gómez, A. (2019)“Towards an Uncertainty-Aware Visualization in the Digital Humanities”, *Informatics*, 6(3), p. 31. {Avilable at: https://www.mdpi.com/2227-9709/6/3/31 (Accessed 02 September 2019).}

Vienni Baptista et al. 2020 Vienni Baptista, B., Fletcher, I., Maryl, M., Wciślik, P., Buchner, A., Lyall, C., Spaapen, J., and Pohl, C.*Final Report on Understandings of Interdisciplinary and Transdisciplinary Research and Factors of Success and Failure*. {Available at: https://zenodo.org/record/3824839 (Accessed 03 February 2022).}

Wallack and Srinivasan 2009 Wallack, J. S., and Srinivasan, R. (2009) “Local-Global: Reconciling Mismatched Ontologies in Development Information Systems”, *2009 42nd Hawaii International Conference on System Sciences*, Waikoloa, HI, USA: IEEE, pp. 1-10. DOI: 10.1109/HICSS.2009.295.