This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

Communicating Uncertainty in Digital Humanities Visualization Research

Georgia Panagiotidou* KU Leuven & UCL Interaction Center

Houda Lamqaddam KU Leuven

Jeroen Poblome KU Leuven

Koenraad Brosens KU Leuven

Katrien Verbert KU Leuven

Andrew Vande Moere KU Leuven

Abstract—Due to their historical nature, humanistic data encompass multiple sources of uncertainty. While humanists are accustomed to handling such uncertainty with their established methods, they are cautious of visualizations that appear overly objective and fail to communicate this uncertainty. To design more trustworthy visualizations for humanistic research, therefore, a deeper understanding of its relation to uncertainty is needed. We systematically reviewed 126 publications from digital humanities literature that use visualization as part of their research process, and examined how uncertainty was handled and represented in their visualizations. Crossing these dimensions with the visualization type and use, we identified that uncertainty originated from multiple steps in the research process from the source artifacts to their datafication. We also noted how besides known uncertainty coping strategies, such as excluding data and evaluating its effects, humanists also *embraced* uncertainty as a separate dimension important to retain. By mapping how the visualizations encoded uncertainty, we identified four approaches that varied in terms of explicitness and customization. This work contributes with two empirical taxonomies of uncertainty and it's corresponding coping strategies, as well as with the foundation of a research agenda for uncertainty visualization in the digital humanities. Our findings further the synergy among humanists and visualization researchers, and ultimately contribute to the development of more trustworthy, uncertainty-aware visualizations.

Index Terms—digital humanities, data visualization, uncertainty, review

1 Introduction

Computational methods, including data visualization, are becoming common research tools in humanistic inquiry. Nevertheless, with all their potential to enable new research directions [48, 79], visualizations are also critiqued for their epistemological incompatibility to humanistic principles [38]. Emerging from this ongoing debate, the field of visualization for the digital humanities has been established [18], differentiating itself from positivist approaches to visualization and adopting its own, more critical [39] design principles. Interfaces created for humanistic research are made generous in their browsing [119], speculative and playful in their interaction features [55], semantically rich in their layouts [49] and critical in their points-of-view [39].

Nevertheless, data uncertainties originating from issues such as implicit errors [98], bias and precision [17] bring skepticism, even mistrust to humanists using visual analysis tools. Datasets in the humanities are often based on historical, incomplete sources and are the result of meticulously intense, manual and subjective collection processes [39, 124]. While humanists are accustomed and trained to critically navigate such uncertainty with their traditional methods, the switch to data-driven research is changing this practice. In order to represent uncertainty in meaningful ways therefore, we need to better understand current humanistic data practices and their coping approaches to uncertainty.

Building on existing taxonomies of uncertainty in digital humanities [106, 123], this work provides a bottom-up, empirical view of uncertainty visualization practice in humanistic research. Specifically, it examines a body of 126 publications from both digital humanities and visualization research in which it systematically documents the moments where authors discuss data uncertainty. This work then examines to what extent these uncertainties are communicated in their data visualizations and their otherwise handling.

*e-mail: g.panagiotidou@ucl.ac.uk

Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx

We use the term uncertainty similarly to Boukhelifa *et al.* [14], to denote moments of doubt or ambiguity in a data analysis process. Such a definition includes uncertainty that derives from all stages of analysis (e.g. data collection, modeling or other). Moreover, it captures indirect uncertainties that are the result of poor quality of the underlying sources of evidence such as partial historical sources, as well as direct uncertainties such as the confidence intervals of a statistical model [115].

We identified that most issues originate from missing, incomplete or conflicting data as well as the process of datafication. We uncovered how besides known strategies for coping with uncertainty, such as excluding data and evaluating its effects [14, 65, 80], humanists also embraced uncertainty as a separate dimension important to retain. We noted how comparatively few visualizations represented uncertainty and those that did, adopted four visualization approaches: established, dedicated, improvisational and experiential. These approaches differed on whether the representation was standard or custom and on whether it was explicit as glyphs or implicit as embedded into the visualization.

This work's contributions include: (1) a thorough understanding of how data and its uncertainty is visualized in the field of digital humanities; (2) a taxonomy of uncertainty origins that can be used as a lens to inform future research directions; (3) a taxonomy of coping strategies for handling uncertainty; and (4) a research agenda for future work in the field. This research therefore enables the development of uncertainty-aware visualizations that respect and acknowledge humanistic practice.

2 Related Work

As the focus of this work is on uncertainty visualization practice in humanistic research, we summarize key publications on uncertainty in data work, in the digital humanities and in visualization.

2.1 Uncertainty in Data Work

Data analysis work is a messy, approximate process [87,99,100]. In fact, data workers encounter multiple origins of uncertainty throughout their analysis process [14]. Data wrangling, when not properly documented, introduces problems of replicability to the analysis [67]. Decisions taken implicitly during data analysis can lead to unreliable conclusions,

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

in what has been referred to as the forking paths problem [65, 102]. Underlying implicit errors, i.e. non-recorded measurement errors inherent to datasets, may challenge the confidence domain experts have in data insights [90, 98]. Uncertainty can also originate at the visualization stage as visualization mirages, i.e. as charts that superficially convey a particular message misleading their readers regarding the underlying data [91]. Even the same data values may later be interpreted or trusted differently depending on discipline-specific 'intuitions' of those reading them [100].

Boukhelifa *et al.* [14] discovered that to cope with uncertainty, data workers use active and tacit strategies. Depending on domain practices, they actively *ignored* known uncertainties because of their irrelevance to the analysis, they tried to *understand* and become more aware of uncertainty dimensions, they deployed methods to *minimize* uncertainty and even tried to *exploit* it. Building on the forking paths problem and decision-making literature [80], Kale *et al.* [65] similarly document that analysts *acknowledge*, *reduce* or *suppress* uncertainty when making analytical decisions. For instance, they acknowledge it by exploring possible analysis paths, they reduce it by gathering additional information and suppress it by eliminating possibilities through intuition. We use these existing uncertainty coping strategies as starting points in our own analysis of uncertainty handling in the digital humanities.

2.2 Data Uncertainty in the Digital Humanities

As data and tasks vary across application domains, uncertainty is studied across contexts such as biomedical research [118], education [34] and daily decision-making [69]. In the digital humanities however, Windhager *et al.* [123] argue that humanistic concerns regarding uncertainty relate more to hermeneutics (their subjective meanings) rather than the well-defined metadata uncertainty often depicted in visualizations. They propose a data typology of uncertainty specific to cultural objects, arguing that all data values ranging from the original source object to its metadata can be imprecise, contested and missing [122, 124]. Similarly, Therón *et al.* [113] suggest a taxonomy from the analyst's perspective, explaining that uncertainty in the digital humanities can range from imprecision, ignorance and credibility to incompleteness. They also found that humanists reason about uncertainty in their research depending on its origin along the analysis process [8].

Various computational techniques have been proposed to manage uncertainty in humanistic data. Kräutli and Boyd Davis [75] note how the nuances of historic time cannot be represented as intuitively as language patterns that express temporal uncertainty and instead propose graphical probabilistic representations of temporal uncertainty. Martin-Rodilla *et al.* [89] developed a conceptual model for capturing ontological and epistemic vagueness as well as a series of certainty qualifiers for humanistic data. Theron *et al.* [113] suggested externalizing data provenance through progressive visual analysis methods so as to computationally take into account propagated uncertainties. Franke *et al.* [45] quantified confidence instead of uncertainty and visualized it alongside other parameters of the data. Nevertheless, no systematic, empirical examination studied uncertainty across digital humanities visualization research. This work therefore builds on existing studies and uses the above-mentioned taxonomies as starting points for analysis.

2.3 Visualizing Uncertainty

Over the past decades considerable research has examined the representation of uncertain data. Kinkeldey *et al.* [73] propose four dimensions for categorizing uncertainty depictions which we use throughout this paper for their ability to holistically describe the uncertainty visualization design space. They suggest uncertainty representation in visualization is *coincident* or *adjacent*, differentiating on whether the data and its uncertainty are shown in integrated or separate views. Uncertainty representation is *intrinsic* or *extrinsic* depending on whether the existing symbols or new glyphs are used to represent uncertainty. It is *static* or *dynamic* depending on whether interaction is required to show the uncertainty. Finally, uncertainty representation is *implicit* when it is embedded into the representation, instead of *explicitly* expressing it as separate or additional data [33]. For clarity, we note that while uncertainty itself is classified as direct and indirect depending on whether it originates from a fact or the underlying information that constitutes that fact [115], uncertainty representation is classified as explicit and implicit depending on how it is visualized.

Quantitative, probabilistic uncertainty is often displayed in dedicated *extrinsic* ways. Error bars, violin plots [32], quantile dotplots [44, 68], hypothetical outcome plots using animation [66] and samples of forecast tracks [81], all depict types of data distributions as separate parameters of the data. Such techniques support both experts and novices in decision-making tasks such as public transportation planning [68], hurricane preparation [81] or scientific research [97].

To give a sensation of the uncertainty, visualization may take more *intrinsic* approaches. Color [31], sketchiness [13] opacity, blur and fuzziness [85], have all been evaluated for their intuitive potential to express both quantitative and qualitative uncertainty. Still, their adoption and success varies. Sketchiness for instance, is considered equally suited to depict uncertainty as blur or opacity, but not in formal settings [13]. The user's task as well as other contextual parameters appear to be crucial for deciding on an uncertainty visualization approach [13, 85]. This is especially the case when evaluating uncertainty, since subjective descriptions of confidence in language can confound the results [59].

There are also an increasing number of *implicit* visualizations focusing on uncertainty when making data analytical decisions. As a way to represent alternative research paths, Liu *et al.* [82] investigated the end-to-end quantitative analysis process of publications and contributed with a series of graphs that visualized the analytical decisions of the authors . In another implicit visualization approach, interactive notebooks have been developed to allow for more manageable multiverse analysis [37] and for examining various 'states/hypotheses' within data analysis [117]. We capture all explicit/implicit as well as intrinsic/extrinsic representations of uncertainty in humanistic research so as to reach a holistic understanding of uncertainty in the humanities.

2.4 Surveying Visualization in the Humanities

This work differentiates itself from earlier surveys on visualizations for digital humanities in three ways: first, it examines visualization use for research purposes regardless of data type or humanities subdomain; second, it examines the use of visualization-based inquiry in the humanities as part of their broader research process [105], not only visualization designs; and third it provides knowledge related to uncertainty visualization and handling, a topic that has not been systematically reviewed so far. Nevertheless, where possible, it builds on existing surveys by using the same analysis categories and classifications, to ensure maximum comparability of research findings.

Windhager *et al.* [121] analyzed cultural collection visualizations and found that they present multiple levels of data granularity allowing users to browse collections in rich ways. Their survey classifies interfaces with respect to the data, visualization, user and task types, and adopt the Europeana classification for data types of cultural objects as images, audio, video, text, and 3D objects. We also adopt this data type classification to ensure comparability. For visual text analysis, Jänicke *et al.* [62, 63] showed how visualization supports among others, close and distant reading by providing additional information in the form of annotations or by highlighting relations among passages in corpus analysis. Our work overlaps with some of the publications and techniques they uncovered, yet extends the analysis to include all data types, not only text. Continuing their previous survey, Jänicke also collected additional publications that include visualizations from the two largest digital humanities journals [61]. They found three types of visualization use types in these articles: publications that take a novel visualization approach, theoretical publications as well as publications that apply existing visualizations, with the latter being a clear majority. We similarly capture the visualization use of our corpus to compare to these findings.

Benito-Santos and Therón Sánchez [9] deployed a data-driven approach and focused on a bibliographic analysis of the research community focusing on the collaboration networks among authors and the publication topics. Also focusing on collaborations, Ma and Li [84] reviewed the various inscriptions that digital humanities use, including

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

data visualizations. They found that inscription-use seems to be driven by the increase of interdisciplinary collaborations among humanists and STEM researchers. Our survey also approaches visualization as artifacts of science communication and does not discriminate between standard charts and custom interactive visualizations.

3 Methodology

Our full sampling and analysis methodology can be found as a step-by-step chart in the supplementary material.

3.1 Sampling

As we were analyzing publications from visualization and digital humanities research, which are typically indexed by different databases, we examined the most prominent venues for each separately. Accordingly, we scan the publications of the following seven venues: Digital Scholarship in the Humanities (DSH), Digital Humanities Quarterly (DSQ), IEEE Transactions on Visualization and Computer Graphics (TVCG), IEEE Computer Graphics and Applications (CG&A), EG Conference on Visualization (EuroVIS), IEEE Pacific Visualization Symposium (PacificVis) and ACM Conference on Human Factors in Computing Systems (CHI). A first systematic search of the above venues was conducted through the use of keywords. For the venues originating from the fields of visualization and HCI these keywords included `humanities`, `archaeology`, `art history`, `musicology`, `literary`, `linguistics`, `literature`, `law` and `history`, so as to capture the breadth of publications that relate to the humanities. For the venues originating from the digital humanities the keywords included `visuali*` and `visual analysis`. The timeline of the considered papers spans from 2016 to 2021. While this can be considered limited, our sample of 6 years provided a sufficiently large corpus. Moreover, as will be discussed in the following section, our analysis reached theme saturation already after analyzing approximately half of that corpus. Three additional criteria for inclusion were used:

1. The publication must describe empirical work, i.e. position and theoretical papers were excluded.
2. The publication must pertain to humanistic research questions. For instance, we excluded publications from visualization venues where the authors developed techniques that could be applicable for the humanities but did not present such a case study.
3. Each publication should have a minimum of two standard chart figures or one figure of a custom visualization. This criterion helped us filter out cases where the data visualization played only a secondary research role.

The publications collected from the keyword search were screened for inclusion by reading their titles and abstracts leading to 146 publications. Then, after closer examination of the full text, another 20 publications were excluded on the same criteria leading to a total of 126. Several publications were on the border of what would be considered visualization research for instance, software annotation tools that included timeline visualizations or 3D simulations and reconstructions that included abstract visualization elements. Since our research goals pertained to the understanding of data uncertainty practices in the visualizations of humanistic research, we were inclusive of these works.

3.2 Analysis Dimensions

For each publication we captured metadata including the venue of origin, the year, the humanities sub-domain that the research refers to and the research type. The research type was based on a preexisting taxonomy from TaDiRAH[1], a European initiative which has compiled a comprehensive list of digital humanities research activities. We then analyzed the publication corpus on four dimensions: (d1) the visualization parameters, (d2) the reported uncertainties encountered in the analysis process, (d3) the author's action to handle each reported uncertainty, and (d4) the representation of these uncertainties in the visualizations. As these dimensions relate to different types of information, we established an analysis method for each.

[1]`https://vocabs.dariah.eu/tadirah/`

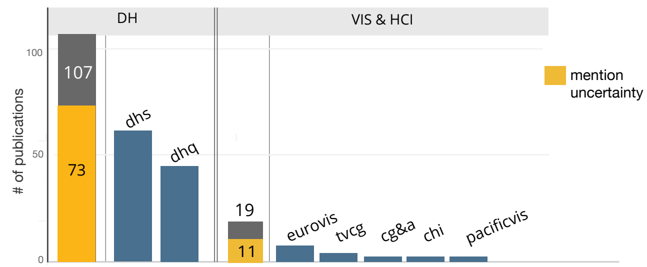

Fig. 1. Paper distribution among venues. In blue, the paper distribution among venues, in gray, the total numbers for the corresponding domain.

d1: Visualization parameters We setup an initial set of 16 criteria for this dimension, which included the type of original data source (e.g. text, image), type of abstracted data (e.g. table, network), the type of visualization, the visualization use and the software tools used to create it. Two authors, coder1 and coder2, then separately coded a random sample of 10 papers on these criteria. Encountering some misalignment, they established a refined definition of each criterion with which they separately coded an additional random sample of 10 papers. To confirm their alignment they run an inter-coder reliability check which returned a Cohen's kappa coefficient of 0.82 (= very good) across the 16 measures. The coders then discussed remaining misalignments and continued to code independently the remaining (N-10) papers.

d2: Uncertainty and d3: Coping dimensions To collect and analyze quotes on uncertainty and their handling of it, we used an iterative, mixed inductive and deductive thematic analysis augmented with two rounds of theme organization [16]. Specifically, the inductive aspects of d2 were informed by the two existing uncertainty taxonomies in digital humanities [9, 123] and the dimension d3 was informed by the two existing uncertainty coping strategies documented in visualization literature [14, 65], as described in the related work.

Starting the analysis, the two authors, coder1 and coder2 separately coded a random sample of 10 papers (the same sample as described above) on the two dimensions (d2,d3) with what each understood as representing uncertainty in the research process. Running into inconsistencies in their results during the first random sample, the two coders established a consistent definition of uncertainty in the publications. Specifically, uncertainty was captured in publication passages where the authors discussed data issues, advised caution or expressed doubt during their analysis process. This would exclude for instance general limitations that did not pertain to their specific analysis process. The two coders also established what constituted a unit of coding i.e. what part of the text would be tagged as a quote for uncertainty. Then, including a third author, and informed by existing practice in qualitative research of open coding free-form text [76], they established a method in which the text units for the thematic analysis were descriptive sentences that summarized the relevant passages of the publication text. This allowed codes to still be comparable even if the tagged text passages had some mismatch.

Having established the methodology for d2 and d3, coder1 extracted all the text units from the publications. Coder1 then grouped these text units into emerging themes until saturation was achieved (at approximately half the corpus). To ensure alignment and non-bias, coder1 created two code-books with these themes and their text units, one for each dimension. These were then discussed with all co-authors, including any text units that did not align. This process resulted in the re-organization of the code-book and clearer wording for the themes. These code-books form the basis of this paper's proposed taxonomies.

d4: Uncertainty visualization We exported the figures from the publication's PDF files using a custom Python script. Starting from the results of d2, we then systematically visually inspected each figure for

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

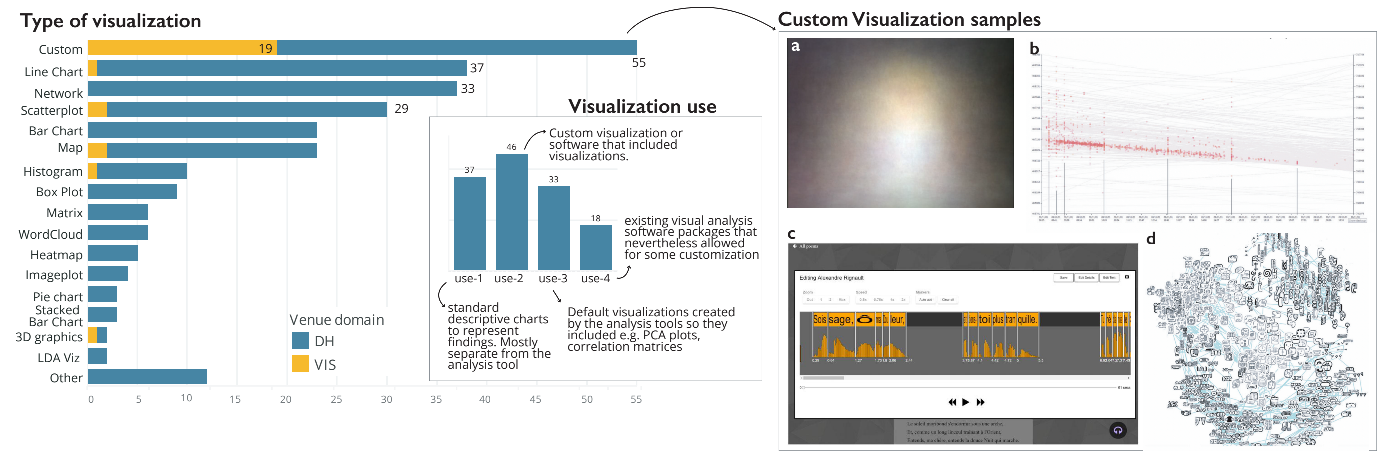

Fig. 2. Overview of the corpus. Left - A significant amount of publications used custom visualizations even when they originated from the digital humanities venues. The 'other' category includes the single instances of a Venn diagram, a treemap, a steam graph, a Shankey plot, glyphs, and a Gantt chart among others. Middle - Publications used their visualizations differently. Many (46) were exploratory visualizations meant to drive the analysis (use-2). Right - Custom visualizations from the digital humanities venues: (a) merging all frames from Disney movies to analyze the accumulative impact of watching a film [43]. (b) Although using the standard format of a slope graph, this 'Storygraph' uses a custom encoding in which the two y axes represent longitude and latitude accordingly with the x axis representing time. They then map events as they unfold in space and time [94]. (c) A custom chart that visualizes spoken word poetry intonation [4]. (d) A network of the relations among Mayan hieroglyphics [57].

its inclusion of uncertainty.

4 Corpus Overview

The grand majority (107) of the 126 publications came from the two digital humanities venues. Still, even within these venues, the humanistic disciplines varied greatly from history and literature, to game and performance studies. The corpus included publications that conducted content, stylistic, network and visual analysis among others. In the TaDiRAH classification, visual analysis is defined as the use of visualizations to examine humanistic data in new formats. It is thus inclusive of both exploratory and communication-oriented visualization.

4.1 Type of data

Humanistic source data go through multiple transformations to be able to be analyzed computationally. We therefore document this transformation process already from the original source data and map where uncertainty manifests. In our corpus, the grand majority of original source data was text (100), followed by images (20), and video (11). There are also 3D objects (6), such as archaeological artifacts, as well as audio (4), and digital-first data (1) like Wikipedia articles. Considering how the digital humanities field originated from literary studies and the overall textual dependency of humanities research [23], this bias towards textual data sources is expected.

The source data often needed to be digitized and potentially even 'datafied' before it could be visualized. Accordingly, using the data classification from [96], the digital format of the source data became text logs (73), tables (37), network trees (21) and fields (16). Some were also transformed into simulations (3) and ontologies (2) which fall outside the original classification. Interestingly, a subsample remained in their original form as images (9), videos (4) or audio (2)[2] and was analyzed as such using for instance ImagePlot. ImagePlot[3], is an image analysis plugin to ImageJ that helps analyze images spatially, based on parameters such as their color scheme.

4.2 Type of visualization

To examine how the type of visualization relates to the uncertainty representation, we map the types of visualization used in the publications. We also document the tools used to generate the visualization in an attempt to understand their impact on this relation. Accordingly, we documented a total of 269 visualizations, out of which 55 were custom. Custom visualizations either had custom visuals or they were custom linked dashboards of standard visualization charts. Such visualizations came from both DH (36/242) and VIS (19/26) venues. Line Charts and Networks are the second most popular types of visualization, which was expected for publications that conduct network analysis. Less expected perhaps is the relative popularity of Wordclouds (6) and ImagePlots (4) which may be reflective of the need to communicate results from content analysis or from examining visual materials such as film and game screenshots, book covers and posters.

The tool used to create the visualizations was not always explicitly mentioned. We believe that to be the case especially when the visualization was created with common tools such as Microsoft Excel. For standard visualizations that need little customization, open source tools like R packages (e.g. stylo and LDAvis), ImagePlot and Gephi are among the most popular. General purpose programming languages such as Java, JavaScript, Python and Unity were preferred for developing custom visualizations. There were overall few articles with map-based visualizations (21/126) but those that existed mentioned being developed using OpenStreetMaps or ArchGIS.

4.3 Visualization use

We hypothesized that uncertainty depictions in visualizations would differ depending on the visualization's purpose and their design as custom or standard charts. Accordingly, we tried to capture the various ways visualizations were used in the publications. We summarize these in four cases and show their frequencies in Figure 2. In the first case (use-1), the authors used standard descriptive visualizations in their publication to communicate their findings. For instance, to give an overview of the distribution of Google's art and culture digitized material, media scholars presented maps, time series and pie charts [74].

In the second case (use-2), humanists - often in collaboration with visualization experts - developed a custom visualization or software that included visualizations (e.g. [4–6]). The authors used visualization as a data analytical tool for exploratory purposes. Archaeologists for instance first analyzed the relations among Maya hieroglyphics and then created a custom force-network visualization of their relations for further exploration [57] (Figure 2-d). Use-2 is where visualization for digital humanities research (VIS4DH) [18] is often situated.

In the third case (use-3), the authors used visualizations as a means to disseminate model parameters and outputs. For the most part, these

[2]These counts do not sum up to the total publications count as the source data can be abstracted into many formats, for instance a text can become a network of characters as well as a table of its metadata.

[3]`https://github.com/culturevis/imageplot`

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

graphs were created by the analytical tools so they included for instance PCA plots, dendrograms for stylistic similarity and correlation matrices among others. This use is quite common as many publications describe an application of machine learning on humanistic research questions ([11,12,51]). For instance, conducting a combination of supervised and unsupervised sentiment analysis on Holocaust testimonials, historians used confusion matrices, word2vec plots and line charts to express process parameters [12].

A final use type (use-4) included existing visual analysis software packages that nevertheless allowed for exploratory analysis. This category is close to use-2 since even though the authors did not develop the software, they still conducted a visualization-driven analysis. The most commonly used software packages included Gephi, ImagePlot and LDAViz (both in R and Python).

Overall, we find that custom visualizations tools are nowadays quite common in DH (in 46 out of 126 publications). There thus seems to be progress from Jänicke's review of the field in 2016 [61] in which humanists were mostly using existing visualization software (use-1 and use-4) versus developing custom (use-2) visualizations.

5 Origins of Uncertainty

From all publications, quite a few (43/126) did not mention any explicit origin of uncertainty. That is not to say they do not mention limitations in their work, but that those limitations are not expressed directly. For instance, they may argue that a technique has been critiqued but they do not relate that with concrete parts of their research analysis as a potential source of uncertainty. Proportionally, publications from visualization venues reported on less data issues (9/19) than those from digital humanities venues (74/107) which could be attributed to the different themes of the journals and conferences.

Overall, there is a wide variation among origins of uncertainty, from quantifiable sources, such as statistical error, to qualitative implicit uncertainty, such as problems of classification and human subjectivity. We structure our results based on when in the research process uncertainty originated (see also Figure 3) in light of developments in visualizing research workflows as discussed in the related work.

5.1 Uncertainty within the Source Data

Missing data. 31 publications reported experiencing missing data sources or dimensions, which made their analysis conclusions uncertain. Even when some information existed, they stated they would need additional data or sources to be more confident. For instance, in comparison with the considerably more Euripidean plays, there were only six Aeschylean and seven Sophoclean plays still in existence, making their cross-analysis by classicists problematic [86].

Missing data were also the result of current conditions rather than past ones. For instance, authors lacked access to resources that would permit a more thorough or complete analysis due to the fact that the physical artefacts were either archived in inaccessible locations [46,50]) or even copyrighted [4, 77, 109]. Some publications also mention lacking computational power to run better or more complete versions of the data analysis [103]. For the purpose of our survey, in this category we also included *inhomogeneity* of datasets [71] which is the result of an imbalanced research effort on a topic.

Imprecise, ambiguous data. The second most common origin of uncertainty in humanistic data was related to its imprecision or ambiguity. In this case, the data dimensions came in ranges or were on a different granularity than what was required for analysis. Two notable examples included how dates were described as ranges on the scale of months or years [24], and how places are only algorithmic approximations of historical geographical locations [1, 2].

Conflicting data. In some situations, uncertainty originated from having multiple possible values for a single data point rather than no data at all - a fact which creates a conflict. In our corpus, linguists examining the primary sources of Dante Alighieri's writing, for instance, described how the historical annotations of his texts were often contradictory, attributing his writing to conflicting sources [7]. Conflicting data was also a core problem when comparing literary text translations

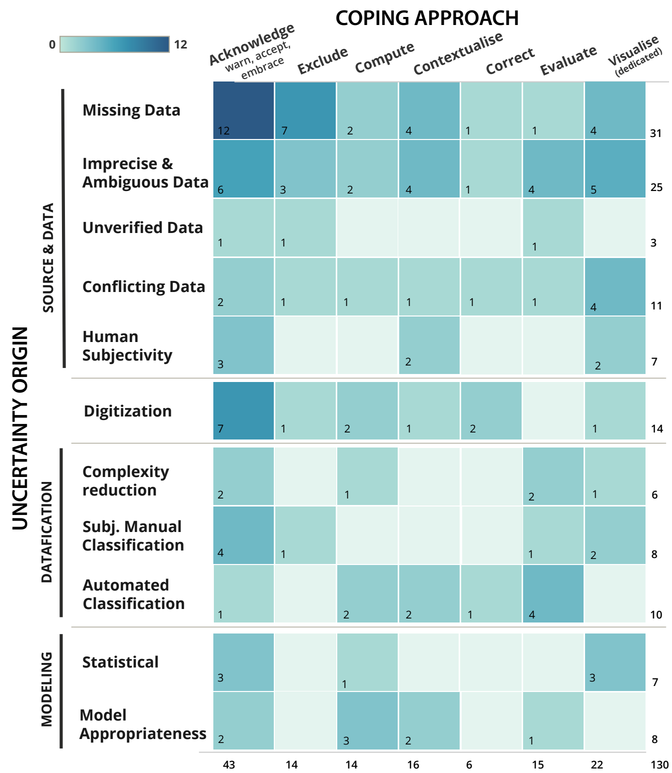

Fig. 3. Contrasting the uncertainty origins to their coping approach. The matrix is ordered by number of uncertainty origins.

as there was a need for alignment among paragraphs, sentences and words between the translation and the original [3, 10, 30, 64].

Unverified data. Although not as prevalent, some issues originated from having data values that were unverified, tentative or contested. For instance, historians trying to map Scottish medieval societal networks described that some dates and provenances mentioned in their sources were tentative and conclusions would need to be verified by traditional historical methods [60]. Similarly, other medievalists studying 16th century Spanish texts, mentioned how parts of their source data, namely the printing places of the texts may in fact be pirated and faked and thus need to be checked further before taken at 'face value' [22].

Human subjectivity. Interpretation plays a central role in the analysis of humanistic source data, however, we only collected seven instances where it is acknowledged as such directly. The linguists studying Dante's writing mentioned how different experts may interpret Dante's commentarists in subjective ways different to their own [7]. Similarly, archaeologists studying cuneiforms - the earliest writing form attested in history - acknowledged how interpreting these cuneiform syllabaries is a subjective process that introduced additional human error [50]. Other archaeologists studying pottery sherds vary in their identification ability during field surveying based on their experience, potentially biasing the final sampling results [98].

5.2 Digitization Errors

The process of digitization was a major source of errors. Most commonly the publications referred to errors from digitizing text with optical character recognition (OCR) [52, 77, 78, 104]. Still, such errors also originated when digitizing artworks and films since their color scheme might have been captured wrongly [43, 125], or when transcribing audio clips as they were spelled out differently depending on cultural factors [92]. Digitizing performance data, such as human dance poses, also gave errors as detection software was 'losing' keyframes [21]. Nevertheless, digitization errors were not just machine driven, as errors also originated because of human intervention for instance in crowd-sourced data such as subtitles or other volunteer transcriptions ([25, 92]).

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

5.3 Datafication Issues

The process of transforming societal and cultural phenomena into data for analysis is referred to as datafication [107]. Here we use it to refer to the process in which the digitized sources from humanities research become 'data' to be analyzed with visualization and other computational techniques. Datafying the original source data reduces its complexity. While this is the intention, it also brings along uncertainty regarding to what is inevitably excluded. In six publications, it was this complexity reduction that was directly quoted as the source of uncertainty. For instance, humanists studying library book patterns explained how the data sources they were using each had its own "*history, complexities, and caveats*" that could not be appreciated by the algorithmic approaches of analysis [108]. Besides this general attitude to reduction, we also identified two more specific sub-themes:

Subjective Manual Classification. Analyzing data computationally in the humanities often requires classifying sources into functional categories that may be artificial, or simply non-binary. For instance, media scholars felt forced to use film genres as a dimension in order to conduct a stylometric analysis, yet they still explained how film genres are *"fuzzy"* and "*evade easy categorization*" [56]. Similarly, media scholars examining film narrative sequences through computational means described how segmenting film scenes into close-ups, shoulder scenes and distance shots involves an amount of annotator subjectivity as "*film analytical concepts do not designate discrete entities"* and accordingly "*any computational distinction between a close-up and a shoulder close-up is purely arbitrary*" [5].

Automated Classification. Ten publications included computational methods to help with the datafication. These were also mentioned as a source of uncertainty, as the humanistic primary data are rarely in such a form to be classified computationally. For instance, trying to analyze intonation of poetry recordings, linguists explain that while the syllabi in the audio recordings were possible to discern through sound fluctuations, this process was still imprecise and manual editing was required [4]. In another example, media scholars argued that reducing historical film motifs in to counts of elements should not be considered as exclusive proxies of their importance [53].

While datafication is often critiqued for its reductionist approach, some embrace it as a new way of seeing. For instance, while flattening full animated films into a single frame reduces its complexity, it also "*magnifies a cinematic experience that is otherwise entirely unnoticeable: the pure, cumulative effect of duration on our eyes and brains without the distraction of narrative or image*" [43] (see Figure 2-a).

5.4 Modeling

Statistical Uncertainty. As many publications conducted statistical analysis, they mentioned uncertainty expressed as confidence intervals of the models. For instance, Hiippala *et al.* [54] expressed the confidence in their language identification of Instagram posts and visualized it as shading bands in their frequency over time. Similarly, Holobut *et al.* [56] visualized the sentiment analysis results of film dialogues and used boxplots to represent the distributions.

Model Appropriateness. Besides statistical uncertainty however, we also documented uncertainty which originated from using models and algorithms that were optimized for other settings. For instance, historians used models optimized for journalistic texts on historical narratives [40] or models optimized for short tweets (sentiment analysis) on longer journalistic texts [46]. In the latter, the uncertainty was handled by using only article headlines.

6 Representing Uncertainty in the Visualizations

We tried to be consciously inclusive in what constitutes a representation of uncertainty to not be biased towards representations we are accustomed to. Still, from the corpus of 83 publications that mention origins of data uncertainty, only 20 publications somehow depicted uncertainty information in their visualizations. We cautiously describe the main design approaches here but advise the reader to be aware of the relatively small sample number and the generalisability that it can produce.

These 20 publications used opacity, patterns, color as well as proximity, juxtaposition, animation and interaction flows to explicitly or implicitly depict uncertainty. These visualizations were disproportionately coming from visualization venues (7/20). Perhaps that is also because some visualization publications motivated their work on alleviating uncertainty as is discussed below. The type of uncertainties they try to depict mostly related to issues of the source data and specifically missing, ambiguous and conflicting data. Grouping these visualizations, we identified four design approaches that vary on whether they depicted uncertainty in a standard or custom way and on whether that uncertainty representation was implicit or explicit following the definition of Deitrick *et al.* [33].

Established. The first approach to visualization deals with quantifiable uncertainties and depicts them in what one would expect of uncertainty visualization techniques. Modeling errors and data distribution patterns are shown with error bars ([56, 72, 116]) and line cones [54]. Missing geolocated data are highlighted with a different background pattern in a choropleth map of library acquisition data [108] (Figure 4). These are explicit depictions of quantifiable uncertainty that many visualization software tools and packages also offer to authors.

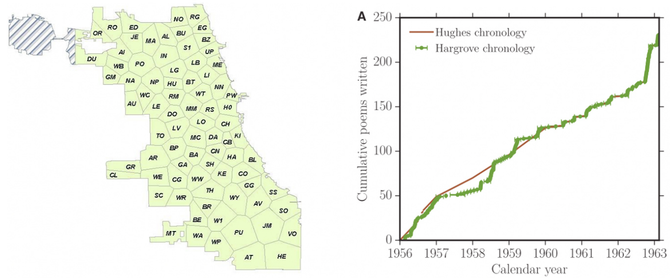

Fig. 4. Established approaches to uncertainty. Left - a map depicting library acquisition data. The authors mention a state with missing data and highlight it with a background pattern [108]. Right - Temporal uncertainty of Sylvia Plath's writing is depicted as two chronologies in red and in green. The green one includes ranges [116].

Improvisational. The second approach includes visualizations that use custom isolated 'tricks' to depict uncertainties and inconsistencies in the data. For instance, the missing geolocations of folktales were shown at another specific location on the map to keep them from being overlooked when browsing the database [95] (Figure 5-top). The continuously changing territory of the 'Papal states' in the 14th to 18th centuries was represented as an annotated single point in a static map as a "*symbolic location*" [101] (Figure 5-bottom). The temporal impreciseness among narrative events in literary texts was communicated by creating a custom chart that represents the overall temporal sequence as well as ranges to each start and end point [24]. These are explicit depictions of partially quantifiable uncertainty, that however visualization software tools do not offer by default so that the authors needed to appropriate their tools and data [27] to depict them.

Experiential. Another approach to uncertainty visualization intended to give an impression of uncertainty rather than quantify it. For instance, to allow for more ambiguous relationships among documents, the custom visualization software Pliny used spatial proximity [20]. Its authors mention that *"putting two items close together suggests some degree of connection without requiring that it be spelled out too specifically"* (see Figure 6-right). When building a reconstruction of the ancient city of Akgor, there were conflicting textual sources for a temple's color. Rather than prefer one over the other, the authors' solution was to allow the visualization to cycle among the colors of the different sources thus eventually depicting all the conflicting data [29]. In another instance, to depict OCR uncertainty the authors visualized the values of confidence given by the OCR algorithm directly 'on' the data by changing the opacity of the dedicated text segment [19] (Figure 6-left). These are all custom, implicit depictions of qualitative uncertainty, that use animation, proximity, opacity as visual cues to bypass quantification.

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

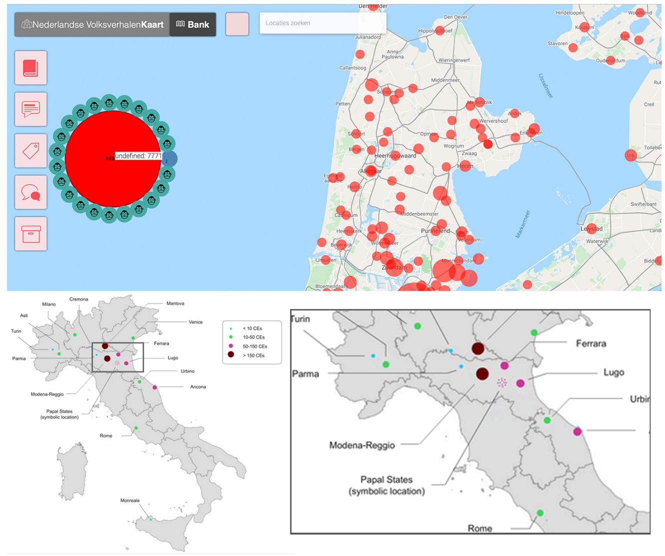

Fig. 5. Improvisational uncertainty approach. Top - Missing geolocations of Dutch folktales in the custom visualization were depicted in a random spot on the map. Screenshot from [95]'s tool. Bottom - Fluctuating locations over time are presented at a 'symbolic' map location [101].

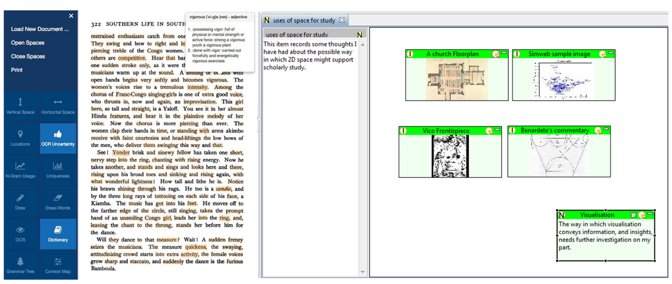

Fig. 6. Experiential approaches to uncertainty. Left - a digitized text is overlaid with color that changes opacity based on the confidence of the OCR algorithm [19]. Right - A custom software developed to support historical text and image interpretation [20], here ambiguous relationships among artifacts are depicted by their proximity.

Dedicated. The final approach includes custom visualizations whose main motivation is an uncertainty. Working with literary scholars, Stoffel *et al.* [111] created a visualization tool directly for disambiguating entities in book character networks (Figure 7-left). Alex *et al.* motivate their work, Palimpsest, a mixed human-automated text mining visualization tool that helps identify literary works set in Edinburgh, as a way to alleviate uncertainty appearing from computational classification [2]. Similarly, ElAssady *et al.* motivate their custom visualization for discourse analysis to the potential subjectivity and non reproducibillity of manually coding the data [42]. These are custom visualizations that have both explicit and implicit depictions of uncertainty.

Overall, there were both explicit (using glyphs and encodings) as well as implicit depictions of uncertainty (implying it through distance and interaction design). However, explicit depictions seemed most common, with implicit ones mostly found in custom visualizations.

7 Navigating Uncertainty in the Digital Humanities

Even if data uncertainties were not included in the visualizations, the authors often mentioned how they handled them.

Acknowledge. The authors described the issues in their publications without actively trying to address or alleviate it - they *acknowledged* it. Still, the motivation behind this acknowledgment changed as we identified three types of acknowledging: warning, accepting and embracing.

In most cases (25/47) the acknowledgment was presented in the form

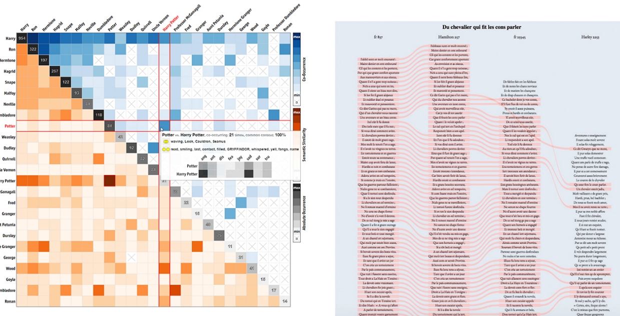

Fig. 7. Dedicated uncertainty visualizations. Left - A custom visualization that supports disambiguation in named entity recognition results [111]. Right - A dedicated visualization approach that supports working with conflicting data of translation texts [64].

of a *warning* or caveat in the research which readers were meant to take into account when evaluating the conclusions. This is especially the case with missing data sources. When *accepting*, the authors described the issues as an unavoidable part of the process. This is the case for instance with the human subjectivity that is embedded in the classification of film scenes as described above [5]. Referring to the OCR issues, some authors also *accepted* the data is sufficiently large to not have an overall impact on the final data analysis [52, 78].

Some authors *embraced* the data issues by arguing that the issues hold meaning and are information in and of themselves meant to be preserved and even communicated. Historians who analyzed the territorial papers, a set of 10,000 documents collecting US correspondence over six decades in the 18th century, encountered many spelling irregularities that hindered the analysis [78]. Rather than standardizing them, they actually embraced the irregularities as part of the data because "*even slight changes in spellings may prove a useful field for inquiry later*". The same is true for retaining multilingual words in other historians' analysis of holocaust testimonials who wanted to preserve the character of testimonies rather than exclude them, even though it would hinder the textual analysis [12].

Exclude. The authors *excluded* problematic points or sources from the sample. This was done either manually for instance, to balance out missing data that may skew the overall picture or even computationally through the use of heuristics to wean out data that were too imprecise to work with. Comparing datasets across regions, archaeologists described how some sites were missing parts of cuneiform sets which clearly skewed the data towards one direction [50]. Their solution was to try to balance this computationally by filtering out values that were unique to a site. Reflecting on the unverified source data of the 16th century Spanish texts, historians used a script to filter out combination of locations and documents that appeared but once in their datasets [22].

Compute. The authors tried to minimize [14] the effect of uncertainty by using algorithmic or computational means. For instance, performance scholars used averaging among keyframes to fill in the missing motion sensor data [21]. Music scholars assumed all network relations among artists appearing on the same album for lack of more granular information [47]. Environmental researchers examining historical fauna patterns from old journals, used fuzzy matching in an attempt to fill in missing data from OCR errors [109].

Evaluate. The authors included a human-in-the-loop to *evaluate* the relevance and extent of the uncertainty. For instance, previously mentioned issues of automated classification of poetry recordings into syllabi were handled by manual editing [4]. Overall, a popular technique appeared to be the mixed computational and manual approach where humanists started with an automated classification method to extract basic entities of interest and then allowed for the manual improvement of the results by human annotators (e.g. [2, 4, 5, 60, 93, 101]).

Contextualize. The authors added additional information regarding the uncertain points. This information was for instance, additional metadata about who has done the interpretation [7], information capturing

This article has been accepted for publication in IEEE Transactions on Visualization and Computer Graphics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TVCG.2022.3209436

the confidence of the OCR process [26], or even dedicated 'unknown' values [11]. In the latter, historians were trying to predict past gender relations in historical datasets. As what is considered a female versus male name changes over time, their solution was to contextualize this situation by adding 'unknown' values rather than force-fitting the current name association to the past. Moreover, humanists contextualized data by cross-referencing to other databases or existing ontologies. This was especially common for filling in missing data. For instance, historians analyzing censorship of Hebrew manuscripts used an existing list of Italian censors to complete missing information in their own data [101].

Correct. Finally, the authors allowed the editing of the data or system to correct errors even after deployment. This could be part of a crowd-sourcing endeavor as in the case of media scholars working with historical newspapers in Australia. There, wrongly assigned locations of paper clippings could be edited by users through their custom created tool [77]. Similarly, in an attempt to mitigate their partial use of only some of Dante's commentarists, linguists made their tool's underlying ontology expandable to more sources as they appear [7].

8 Discussion

We presented how humanists experience uncertainty in their research with visualizations and how they try to address it. We now discuss the implications of these findings for visualization design in the digital humanities as well as set an agenda for future work.

Datafication and conflicting data as origins of uncertainty. Our results are inline with Boukhelifa *et al.*'s taxonomy documenting imprecision, error and missing data [14] as well as Benito-Santos *et al.*'s taxonomy documenting ignorance, incompleteness and lack of credibility [8]. Still, we additionally captured less accounted for origins of uncertainty arising from conflicting data and issues in the datafication process. We identified how conflicting data originates from having too much information over a data point and found that it even triggered *dedicated* visualizations to navigate around it. We also observed how considerable uncertainty in digital humanities research originates from the datafication process. Drucker suggests that "*humanities documents [..] are not "data" [..] They have to be remediated to become "data" -[..] and in the process several issues come into play* [39, p. 245]. We demonstrated how classification and complexity reduction are two of the issues humanists face during this 'remediation process'. Critical data studies point out that classifications enforce power unequally, as they delineate what will remain visible and invisible within the data [15, 35, 83, 114]. This taxonomy demonstrates that what remains invisible within a classification is what seeds such uncertainty.

Visual and textual communication of uncertainty. Overall, from all the 126 reviewed publications, 83 mention uncertainty textually while only 20 visualized it. In a survey on why visualization authors do not include uncertainty in their visualizations, a multitude of reasons were mentioned, among which a fear that depictions of uncertainty would interfere with the main message of the visualization itself [58]. Nevertheless, as the core purpose of scientific publications is to transfer research findings, we expected uncertainty visualization to be more prevalent. We observed that to this end some publications went beyond, and appropriated existing tools to depict uncertainty. We thus argue that the mismatch of textual and visual expressions of uncertainty can also be partially attributed to a lack of standard ways to visualize uncertainty with existing visualization tools.

Navigating Uncertainty. The majority of uncertainty origins were acknowledged by the authors, with all other strategies except 'correct', having a similar occurrence pattern (Figure 3). Still, while not very common, 'correcting' as a strategy has not been explicitly highlighted. As documented in previous research [14, 65] uncertainty was often acknowledged in the text (43/130) by presenting it as a warning or an inevitable truth of the data. Nevertheless, we also identified how authors shifted the coping action beyond the original researcher by encouraging a human-in-the-loop evaluation of the results in addition to allowing a dynamic correction of the data even after a tool's deployment. While human-in-the-loop visual analytics are already theorized and deployed in practice, visualizations that permit dynamic correction of the underlying data are less so. Being able to correct values through the visualization interface rather than spreadsheets or databases is a long discussed feature in visualization research [67] that seems to resonate over multiple user groups [36]. In digital humanities research whose uncertainty often originates from issues in the original data sources, this flexibility seems crucial to ensure the trustworthy adoption of visualization systems.

Embracing uncertainty. This work demonstrated that in digital humanities research, uncertainty is also *embraced* as it holds meaning that needs to be retained. Visualization taxonomies of uncertainty include *ignore* or *acknowledge* as indicators of uncertainty that is not actively coped with [14, 65]. Nevertheless, humanists are trained on interpretivist paradigms which by definition accept uncertain, incomplete viewpoints to knowledge and information [38]. The uncertainty, ambiguity, and complexity thus are not faults, but can be productive commitments to their process [23, 41]. We therefore believe that visualization research should both support such uncertainty handling, as well as adopt more neutral vocabulary towards the currently negatively colored terms that describe uncertainty.

9 A research agenda for uncertainty in DH

Our findings enrich and extend existing uncertainty taxonomies in digital humanities visualization research with comprehensive empirical data of how and at what stages uncertainty is encountered in practice. With this refined taxonomy of uncertainty in digital humanities visualization, new directions for research can be established. Specifically, we propose three directions for a research agenda that relate to the opportunities for visualization research, opportunities for humanistic research as well as for their overall synergy.

First, this taxonomy functions as a list of open challenges for data visualization research around uncertainty. While for instance missing and imprecise data are already actively investigated for their impact [110] and design techniques [13], other types of uncertainty documented in humanistic research such as conflicting data or human subjectivity are less so. Future work can thus examine these in more depth as well as probe their transferability to other settings.

Second, such a taxonomy assists the meta-design of patterns that can empower humanists themselves to visualize data in ways that better support their own data practices. We documented how currently, visualization for the digital humanities often focuses on developing custom fit visualizations. Examining depictions of uncertainty in our corpus, we saw that humanists appropriated common tools to depict uncertainty in their own improvisational way. Appropriation, i.e. the adaption of technology in ways not envisaged by the designers [28], is a great resource for design as it informs designers on user requirements and even future innovation [27]. We argue that these visualization appropriations can be analyzed as starting points towards the creation of a design pattern 'handbook' for representing uncertainty in humanities research. We thus suggest a conceptual shift from focusing on developing new custom visualizations, towards creating new '*established*' patterns for uncertainty (as described in Section 6) that can be utilized with common visualization software like Excel and Gephi.

Third, this taxonomy can better support the dialogue among humanists and visualization researchers by bringing awareness to the multiple scales on which uncertainty enters humanistic research as well as by providing a common framework from which to springboard such a dialogue. The need for participatory ways to collaborate and engage domain experts along problem-driven visualization studies is established [70, 88]. Moreover, it is known that one of the challenges such interdisciplinary collaborations face is the different vocabularies among scientists [88, 120]. This taxonomy therefore supports collaboration efforts by creating a common vocabulary and thus acts as a boundary object [112] that helps connect concepts across communities. Future work can build on this taxonomy and create dedicated participatory sessions that help elicit and discuss uncertainty in grounded ways.

10 Limitations

Much like the publications of this review, this research has its own set of uncertainties. First, while it has been conducted collaboratively and reflectively, it may still suffer from the choice of the text excerpts of what is considered uncertainty. We believe that by establishing our definition of uncertainty and sharing the final codes we generated, this has been mitigated. Accordingly, it is possible that some uncertainties were not captured within the scope of our sample. Nevertheless, as the qualitative reviewing process has reached saturation already midway the sample size, we believe that even if some accounts were missed, these would be minimal in terms of the larger picture. We thus suggest the readers to not focus on the numerical counts of the uncertainties but their existence and overall signification. Some uncertainty coping strategies cannot be captured though the analysis of publications. We are aware for instance of the social dimensions of evaluating the trustworthiness of data [65, 98] which are not visible in our corpus. Finally, there is still differentiation to be accounted among digital humanists. For instance, the uncertainty experienced when applying machine learning to a corpus is not necessarily comparable to when developing a visual analysis tool to analyze artworks. Nevertheless, this variability is characteristic of the digital humanities as an interdisciplinary field that gathers multiple approaches under its umbrella [48].

11 Conclusion

We examined how uncertainty, i.e. moments of doubt or ambiguity in the data analysis process, was documented in 126 publications from digital humanities literature that use visualization. Analyzing this corpus, we proposed two taxonomies collecting a series of uncertainty origins over the digital humanities visualization research process as well as the strategies humanists used to handle them. We uncovered how comparatively few visualizations represented uncertainty origins and those that did adopted four different visualization approaches: *established*, *dedicated*, *improvisational* and *experiential*. With our findings we set a research agenda for the development of uncertainty-aware visualizations that acknowledge the existing humanistic practice.

References

[1] B. Adams. Chronotopic Information Interaction: Integrating Temporal and Spatial Structure for Historical Indexing and Interactive Search. *Digital Scholarship in the Humanities*, 36(3):525–541, 2021.

[2] B. Alex, C. Grover, J. Oberlander, T. Thomson, M. Anderson, J. Loxley, U. Hinrichs, and K. Zhou. Palimpsest: Improving Assisted Curation of Loco-specific Literature. *Digital Scholarship in the Humanities*, 32(suppl_1):i4–i16, 2017.

[3] M. S. Alharbi, T. Cheesman, and R. S. Laramee. TransVis: Integrated Distant and Close Reading of Othello Translations. *IEEE Transactions on Visualization and Computer Graphics*, 2020.

[4] C. Ardrey. Visualising Voice: Analysing Spoken Recordings of Nineteenth-century French Poetry. *Digital Scholarship in the Humanities*, 35(4):737–758, 2020.

[5] J.-H. Bakels, M. Grotkopp, T. Scherer, and J. Strati. Matching Computational Analysis and Human Experience: Performative Arts and the Digital Humanities. *Digital Humanities Quarterly*, 14, 2020.

[6] C. Bardiot. Theatre Analytics: Developing Software for Theatre Research. *DHQ: Digital Humanities Quarterly*, 14(3), 2020.

[7] V. Bartalesi, C. Meghini, D. Metilli, M. Tavoni, and P. Andriani. A Web Application for Exploring Primary Sources: The DanteSources Case Study. *Digital Scholarship in the Humanities*, 33(4):705–723, 2018.

[8] A. Benito-Santos, M. Doran, A. Rocha, E. Wandl-Vogt, J. Edmond, and R. Therón. Evaluating a Taxonomy of Textual Uncertainty for Collaborative Visualisation in the Digital Humanities. *Information*, 12(436), 2021.

[9] A. Benito-Santos and R. T. Sanchez. A Data-Driven Introduction to Authors, Readings, and Techniques in Visualization for the Digital Humanities. *IEEE Computer Graphics and Applications*, 40(3):45–57, 2020.

[10] Y. Bizzoni, M. Reboul, and A. Grosso. Diachronic tends in homeric translations. *DHQ: Digital Humanities Quarterly*, 11(2), 2017.

[11] T. Blanke. Predicting the Past. *Digital Humanities Quarterly*, 12(2):1–21, 2018.

[12] T. Blanke, M. Bryant, and M. Hedges. Understanding memories of the holocaust—a new approach to neural networks in the digital humanities. *Digital Scholarship in the Humanities*, 35(1):17–33, 2020.

[13] N. Boukhelifa, A. Bezerianos, T. Isenberg, and J. D. Fekete. Evaluating Sketchiness as a Visual Variable for the Depiction of Qualitative Uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2769–2778, 2012.

[14] N. Boukhelifa, M. E. Perrin, S. Huron, and J. Eagan. How Data Workers Cope with Uncertainty: A Task Characterisation Study. In *Conference on Human Factors in Computing Systems*, volume 2017-May, pages 3645–3656. ACM Press, 2017.

[15] G. C. Bowker and S. L. Star. *Sorting things out: Classification and its Consequences*. Inside technology. MIT press, Cambridge (Mass.), 2nd print. edition, 1999.

[16] R. E. Boyatzis. *Transforming Qualitative Information: Thematic Analysis and Code Development*. SAGE, London, UK, 1998.

[17] S. Boyd Davis, O. Vane, and F. Kräutli. Can I Believe what I see? Data Visualization and Trust in the Humanities. *Interdisciplinary Science Reviews*, 46(4):522–546, 2021.

[18] A. J. Bradley, M. El-Assady, K. Coles, E. Alexander, M. Chen, C. Collins, S. Janicke, and D. J. Wrisley. Visualization and the Digital Humanities: Moving Toward Stronger Collaborations. *IEEE Computer Graphics and Applications*, 38(6):26–38, 2018.

[19] A. J. Bradley, V. Sawal, S. Carpendale, and C. Collins. Textension: Digitally Augmenting Document Spaces in Analog Texts. *DHQ: Digital Humanities Quarterly*, 13(3), 2019.

[20] J. Bradley and M. Pasin. Fitting Personal Interpretation with the Semantic Web: Lessons Learned from Pliny. *DHQ: Digital Humanities Quarterly*, 11(1), 2017.

[21] P. Broadwell and T. R. Tangherlini. Comparative K-Pop Choreography Analysis through Deep-Learning Pose Estimation across a Large Video Corpus. *DHQ: Digital Humanities Quarterly*, 15(1), 2021.

[22] D. M. Brown, A. Soto-Corominas, and J. L. Suárez. The preliminaries Project: Geography, Networks, and Publication in the Spanish Golden Age. *Digital Scholarship in the Humanities*, 32(4):709–732, 2017.

[23] A. Burdick, J. Drucker, P. Lunenfeld, T. Presner, and J. Schnapp. *Digital_Humanities*. MIT Press, 2012.

[24] J. Burgers. Imagining the Continuously Present Past: Visualizing William Faulkner's Narratives and Digital Yoknapatawpha. *DHQ: Digital Humanities Quarterly*, 15(2), 2021.

[25] J. Byszuk. The Voices of Doctor Who–How Stylometry Can be Useful in Revealing New Information About TV Series. *DHQ: Digital Humanities Quarterly*, 14(4), 2020.

[26] J. Carrive, A. Beloued, P. Goetschel, S. Heiden, A. Laurent, P. Lisena, F. Mazuet, S. Meignier, B. Pincemin, G. Poels, et al. Transdisciplinary Analysis of a Corpus of French Newsreels: The ANTRACT Project. *Digital Humanities Quarterly*, 15(1), 2021.

[27] J. Carroll. Completing Design in Use: Closing the Appropriation Cycle. In *Proceedings of the 2004 European Conference on Information Systems (ECIS)*, volume 44, 2004.

[28] J. Carroll, S. Howard, F. Vetere, J. Peck, and J. Murphy. Identity, Power And Fragmentation in Cyberspace: Technology Appropriation by Young People Recommended Citation. In *ACIS 2001 Proceedings*, volume 6, 2001.

[29] T. Chandler and A. Clulow. Modeling Virtual Angkor: an Evolutionary Approach to a Single Urban Space. *IEEE Computer Graphics and Applications*, 40(3):9–16, 2020.

[30] T. Cheesman, K. Flanagan, S. Thiel, J. Rybicki, R. S. Laramee, J. Hope, and A. Roos. Multi-Retranslation Corpora: Visibility, Variation, Value, and Virtue. *Digital Scholarship in the Humanities*, 32(4):739–760, 2017.

[31] M. Correll and J. Heer. Value-Suppressing Uncertainty Palettes. In *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems (CHI)*, pages 1–11, 2018.

[32] M. Correll, S. Member, and M. G. Member. Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2142–2151, 2014.

[33] S. Deitrick. Evaluating Implicit Visualization of Uncertainty for Public Policy Decision Support. In *Proceedings AutoCarto 2012*, page 16, 2012.

[34] C. Demmans Epp and S. Bull. Uncertainty Representation in Visualizations of Learning Analytics for Learners: Current Approaches and Opportunities. *IEEE Transactions on Learning Technologies*, 8(3):242–260, 2015.

[35] C. D'Ignazio and L. F. Klein. *Data Feminism*. MIT Press, 2020.
[36] E. Dimara, H. Zhang, M. Tory, and S. Franconeri. The Unmet Data Visualization Needs of Decision Makers within Organizations. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, 2021.
[37] P. Dragicevic, Y. Jansen, A. Sarma, M. Kay, and F. Chevalier. Increasing the Transparency of Research Papers with Explorable Multiverse Analyses. *Proceedings of the 2019 Conference on Human Factors in Computing Systems (CHI)*, 2019.
[38] J. Drucker. Humanities Approaches to Graphical Display. *DHQ: Digital Humanities Quarterly*, 5(1), 2011.
[39] J. Drucker. Graphical Approaches to the Digital Humanities. In *A New Companion to Digital Humanities*. John Wiley & Sons, Ltd, 2015.
[40] S. Dunn and M. Schumacher. Explaining Events to Computers: Critical Quantification, Multiplicity and Narratives in Cultural Heritage. *DHQ: Digital Humanities Quarterly*, 10(3):000262, 2016.
[41] J. Edmond. Managing Uncertainty in the Humanities: Digital and Analogue Approaches. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM)*, 2018.
[42] M. El-Assady, V. Gold, C. Acevedo, C. Collins, and D. Keim. ConToVi: Multi-party Conversation Exploration Using Topic-space Views. *Computer Graphics Forum*, 35(3):431–440, 2016.
[43] K. L. Ferguson. Digital Surrealism: Visualizing Walt Disney Animation Studios. *DHQ: Digital Humanities Quarterly*, 11(1), 2017.
[44] M. Fernandes, L. Walls, S. Munson, J. Hullman, and M. Kay. Uncertainty Displays Using Quantile Dotplots or CDFs Improve Transit Decision-Making. In *Proceedings of the 2018 ACM Conference on Human Factors in Computing Systems (CHI)*, Montreal, Canada, 2018. ACM.
[45] M. Franke, R. Barczok, S. Koch, and D. Weltecke. Confidence as First-class Attribute in Digital Humanities Data. In *Proceedings of the Workshop on VIS4DH*, 2019.
[46] P. Ghasiya and K. Okamura. Understanding the Middle East through the eyes of Japan's Newspapers: A topic modelling and sentiment analysis approach. *Digital Scholarship in the Humanities*, 36(4):871–885, 2021.
[47] L. Gienapp, C. Kruckenberg, and M. Burghardt. Topological Properties of Music Collaboration Networks: The Case of Jazz and Hip Hop. *DHQ: Digital Humanities Quarterly*, 15(1), 2021.
[48] M. K. Gold., editor. *Debates in the Digital Humanities*. University of Minnesota Press, Minneapolis, MN, U.S.A, 2012.
[49] F. Gortana, F. von Tenspolde, D. Guhlmann, and M. Dörk. Off the Grid: Visualizing a Numismatic Collection as Dynamic Piles and Streams. *Open Library of Humanities*, 4(2):1–25, 2018.
[50] L. F. Hawkins. Computational Models for Analyzing Data Collected from Reconstructed Cuneiform Syllabaries. *DHQ: Digital Humanities Quarterly*, 12(1), 2018.
[51] W. Heeringa, F. Swarte, A. Schüppert, and C. Gooskens. Measuring Syntactical Variation in Germanic Texts. *Digital Scholarship in the Humanities*, 33(2):279–296, 2018.
[52] T. Heikkilä and T. Roos. Quantitative Methods for the Analysis of Medieval Calendars. *Digital Scholarship in the Humanities*, 33(4):766–787, 2018.
[53] E. Hielscher. The Phenomenon of Interwar City Symphonies: A Combined Methodology of Digital Tools and Traditional Film Analysis Methods to Study Visual Motifs and Structural Patterns of Experimental-Documentary City Films. *DHQ: Digital Humanities Quarterly*, 14(4), 2020.
[54] T. Hiippala, A. Hausmann, H. Tenkanen, and T. Toivonen. Exploring the Linguistic Landscape of Geotagged Social Media content in Urban Environments. *Digital Scholarship in the Humanities*, 34(2):290–309, 2019.
[55] U. Hinrichs, S. Forlini, and B. Moynihan. Speculative Practices: Utilizing InfoVis to Explore Untapped Literary Collections. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):429–438, 2016.
[56] A. Hołobut and J. Rybicki. The Stylometry of Film Dialogue: Pros and Pitfalls. *DHQ: Digital Humanities Quarterly*, 14(4), 2020.
[57] R. Hu, C. P. Gayol, J.-M. Odobez, and D. Gatica-Perez. Analyzing and visualizing ancient maya hieroglyphics using shape: From computer vision to digital humanities. *Digital Scholarship in the Humanities*, 32(suppl_2):ii179–ii194, 2017.
[58] J. Hullman. Why Authors Don't Visualize Uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):130–139, 2020.
[59] J. Hullman, X. Qiao, M. Correll, A. Kale, and M. Kay. In Pursuit of Error: A Survey of Uncertainty Visualization Evaluation. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):903–913, 2019.
[60] C. Jackson. Using Social Network Analysis to Reveal Unseen Relationships in Medieval Scotland. *Digital Scholarship in the Humanities*, 32(2):336–343, 2017.
[61] S. Jänicke. Valuable Research for Visualization and Digital Humanities: A Balancing Act. In *Proceedings of the Workshop on Vis4DH*, 2016.
[62] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann. On Close and Distant Reading in Digital Humanities : A Survey and Future Challenges. *Eurographics Conference on Visualization (EuroVis)*, pages 1–21, 2015.
[63] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann. Visual Text Analysis in Digital Humanities. *Computer Graphics Forum*, 36(6):226–250, 2017.
[64] S. Jänicke and D. Joseph Wrisley. Visualizing Mouvance: Toward a Visual Analysis of Variant Medieval Text Traditions. *Digital Scholarship in the Humanities*, 32(suppl_2):ii106–ii123, 2017.
[65] A. Kale, M. Kay, and J. Hullman. Decision-Making Under Uncertainty in Research Synthesis: Designing for the Garden of Forking Paths. In *Proceedings of the 2019 ACM Conference on Human Factors in Computing Systems (CHI)*. ACM, 2019.
[66] A. Kale, F. Nguyen, M. Kay, and J. Hullman. Hypothetical Outcome Plots Help Untrained Observers Judge Trends in Ambiguous Data. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):892–902, 2019.
[67] S. Kandel, J. Heer, C. Plaisant, J. Kennedy, F. Van Ham, N. H. Riche, C. Weaver, B. Lee, D. Brodbeck, and P. Buono. Research Directions in Data Wrangling: Visualizations and Transformations for Usable and Credible Data. *Information Visualization*, 10(4):271–288, 2011.
[68] M. Kay, T. Kola, J. R. Hullman, and S. A. Munson. When (ish) is My Bus? User-centered Visualizations of Uncertainty in Everyday, Mobile Predictive Systems. In *Proceedings of the 2016 ACM Conference on Human Factors in Computing Systems (CHI)*, pages 5092–5103, San Jose, CA, USA, 2016. ACM.
[69] M. Kay, D. Morris, M. Schraefel, and J. A. Kientz. There's No Such Thing as Gaining a Pound: Reconsidering the Bathroom Scale User Interface. *UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 401–410, 2013.
[70] E. Kerzner, S. Goodwin, J. Dykes, S. Jones, and M. Meyer. A Framework for Creative Visualization-Opportunities Workshops. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):748–758, 2018.
[71] R. Khulusi, J. Kusnick, C. Meinecke, C. Gillmann, J. Focht, and S. Jänicke. A Survey on Visualizations for Musical Data. *Computer Graphics Forum*, 39(6):82–110, 2020.
[72] H. Kim. Visualizing Textual Similarity of Shakespearean Suspect Texts: An Examination of the Henry VI Plays. *Digital Scholarship in the Humanities*, 36(4):919–933, 2021.
[73] C. Kinkeldey, A. M. MacEachren, and J. Schiewe. How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies. *Cartographic Journal*, 51(4):372–386, 2014.
[74] I. Kizhner, M. Terras, M. Rumyantsev, V. Khokhlova, E. Demeshkova, I. Rudov, and J. Afanasieva. Digital Cultural Colonialism: Measuring Bias in Aggregated Digitized Content Held in Google Arts and Culture. *Digital Scholarship in the Humanities*, 36(3):607–640, 2021.
[75] F. Kräutli and S. Boyd Davis. Known Unknowns: Representing Uncertianty in Historical Time. In *Electronic Visualisation and the Arts (EVA)*, pages 61–68, 2013.
[76] K. S. Kurasaki. Intercoder Reliability for Validating Conclusions Drawn from Open-Ended Interview Data. *Field Methods*, 12(3):179–194, 2000.
[77] S. Kutty, R. Nayak, P. Turnbull, R. Chernich, G. Kennedy, and K. Raymond. PaperMiner—a Real-time Spatiotemporal Visualization for Newspaper Articles. *Digital Scholarship in the Humanities*, 35(1):83–100, 2020.
[78] J. Ledolter and L. VanderVelde. A Case Study in Text Mining: Textual Analysis of the Territorial Papers. *Digital Scholarship in the Humanities*, 35(1):101–126, 2020.
[79] Lev Manovich. Data Science and Digital Art History. *International Journal for Digital Art History*, 1(1), 2015.
[80] R. Lipshitz and O. Strauss. Coping with Uncertainty: A Naturalistic Decision-Making Analysis. *Organizational Behavior and Human Decision Processes*, 69(2):149–163, 1997.
[81] L. Liu, L. Padilla, and S. H. Creem-regehr. Visualizing Uncertain Tropical

Cyclone Predictions using Representative Samples from Ensembles of Forecast Tracks. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):882–891, 2019.

[82] Y. Liu, T. Althoff, and J. Heer. Paths Explored, Paths Omitted, Paths Obscured: Decision Points & Selective Reporting in End-to-End Data Analysis. In *Proceedings of the 2020 Conference on Human Factors in Computing Systems (CHI)*, New York, NY, USA, 2020. ACM.

[83] Y. Loukissas. *All Data are Local: Thinking Critically in a Data-Driven Society*. MIT Press, 2019.

[84] R. Ma and K. Li. Digital Humanities as a Cross-Disciplinary Battleground: An Examination of Inscriptions in Journal Publications. *Journal of the Association for Information Science and Technology*, pages 1–16, 2021.

[85] A. M. Maceachren, R. E. Roth, J. O'Brien, B. Li, D. Swingley, and M. Gahegan. Visual Semiotics & Uncertainty Visualization: An Empirical Study. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2496–2505, 2012.

[86] N. Manousakis and E. Stamatatos. Devising Rhesus: A Strange 'Collaboration' Between Aeschylus and Euripides. *Digital Scholarship in the Humanities*, 33(2):347–361, 2018.

[87] Y. Mao, D. Wang, M. Muller, K. R. Varshney, I. Baldini, C. Dugan, and A. Mojsilovic. How Data Scientists Work Together with Domain Experts in Scientific Collaborations: To Find the Right Answer or to Ask the Right Question? *Proceedings of the ACM on Human-Computer Interaction*, 3(GROUP), 2019.

[88] G. E. Marai. Activity-Centered Domain Characterization for Problem-Driven Scientific Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):913–922, 2018.

[89] P. Martin-rodilla and C. Gonzalez-perez. Representing Imprecise and Uncertain Knowledge in Digital Humanities: A Theoretical Framework and ConML Implementation with a Real Case Study. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM)*, 2018.

[90] N. McCurdy, J. Gerdes, and M. Meyer. A Framework for Externalizing Implicit Error Using Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):925–935, 2019.

[91] A. McNutt, G. Kindlmann, and M. Correll. Surfacing Visualization Mirages. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI)*, pages 1–16, 2020.

[92] P. Meindertsma. Changes in Lyrical and Hit Diversity of Popular U.S. Songs 1956-2016. *DHQ: Digital Humanities Quarterly*, 13(4), 2019.

[93] G. Mihály Tóth. Recovering and Rendering Silenced Experiences of Genocides: Testimonial Fragments of the Holocaust. *Digital Scholarship in the Humanities*, 36(Supplement_1):i124–i136, 2021.

[94] B. Miller, A. Shrestha, and J. Olive. Visualizing Computational, Transversal Narratives from the World Trade Towers. *Digital Scholarship in the Humanities*, 32(2):344–354, 2017.

[95] I. Muiser, M. Theune, R. d. Jong, N. Smink, R. B. Trieschnigg, D. Hiemstra, and T. Meder. Supporting the Exploration of Online Cultural Heritage Collections: The Case of the Dutch Folktale Database. *DHQ: Digital Humanities Quarterly*, 11(4), 2017.

[96] T. Munzner. *Visualization analysis and design*. CRC press, 2014.

[97] L. M. K. Padilla, M. Powell, M. Kay, and J. Hullman. Uncertain About Uncertainty: How Qualitative Expressions of Forecaster Confidence Impact Decision-Making With Uncertainty Visualizations. *Frontiers in Psychology*, 11, 2021.

[98] G. Panagiotidou, R. Vandam, J. Poblome, and A. Vande Moere. Implicit Error, Uncertainty and Confidence in Visualization: an Archaeological Case Study. *IEEE Transactions on Visualization and Computer Graphics*, pages 1–1, 2021.

[99] S. Passi and S. J. Jackson. Data Vision: Learning to See Through Algorithmic Abstraction. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*, pages 2436–2447, New York, New York, USA, 2017. ACM Press.

[100] S. Passi and S. J. Jackson. Trust in Data Science: Collaboration, Translation, and Accountability in Corporate Data Science Projects. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–28, nov 2018.

[101] G. Prebor, M. Zhitomirsky-Geffet, and I. Miller. A Multi-dimensional Ontology-based Analysis of the Censorship of Hebrew Manuscripts. *DHQ: Digital Humanities Quarterly*, 14(1), 2020.

[102] X. Pu and M. Kay. The Garden of Forking Paths in Visualization: A Design Space for Reliable Exploratory Visual Analytics. In *Proceedings - 7th Biennial Workshop Evaluation and Beyond: Methodological Approaches for Visualization, BELIV 2018*, pages 37–45. IEEE, 2018.

[103] S. Rebora, J. B. Herrmann, G. Lauer, and M. Salgaro. Robert Musil, a War Journal, and Stylometry: Tackling the Issue of Short Texts in Authorship Attribution. *Digital Scholarship in the Humanities*, 34(3):582–605, 2019.

[104] G. Rotari, M. Jander, and J. Rybicki. The Grimm Brothers: A Stylometric Network Analysis. *Digital Scholarship in the Humanities*, 36(1):172–186, 2021.

[105] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim. The Role of Uncertainty, Awareness, and Trust in Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):240–249, 2016.

[106] R. T. Sánchez, A. B. Santos, R. S. Vicente, and A. L. Gómez. Towards an Uncertainty-aware Visualization in the Digital Humanities. *Informatics*, 6(3), 2019.

[107] M. T. Schafer and K. V. Es. *The Datafied Society: Studying Culture through Data*. Amsterdam University Press, 2016.

[108] J. Shanahan, R. Burke, and A. Lučić. Reading Chicago Reading: Quantitative Analysis of a Repeating Literary Program. *DHQ: Digital Humanities Quarterly*, 14(2), 2020.

[109] R. Smail, C. Donaldson, R. Govaerts, P. Rayson, and C. Stevens. Uncovering Environmental Change in the English Lake District: Using Computational Techniques to Trace the Presence and Documentation of Historical Flora. *Digital Scholarship in the Humanities*, 36(3):736–756, 2021.

[110] H. Song and D. A. Szafir. Where's my data? Evaluating Visualizations with Missing Data. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):914–924, 2019.

[111] F. Stoffel, W. Jentner, M. Behrisch, J. Fuchs, and D. Keim. Interactive Ambiguity Resolution of Named Entities in Fictional Literature. *Computer Graphics Forum*, 36(3):189–200, 2017.

[112] S. Susan, Leigh and G. James, R. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology. *Social Studies of Science*, 19(3):387–420, 1989.

[113] R. Therón, A. G. Losada, A. Benito, and R. Santamaría. Toward Supporting Decision-making under Uncertainty in Digital Humanities with Progressive Visualization. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*, pages 826–832. ACM, 2018.

[114] J. Thorp. *Living in Data: A Citizen's Guide to a Better Information Future*. MCD, 2021.

[115] A. M. Van Der Bles, S. Van Der Linden, A. L. Freeman, J. Mitchell, A. B. Galvao, L. Zaval, and D. J. Spiegelhalter. Communicating Uncertainty about Facts, Numbers and Science. *Royal Society Open Science*, 6(5), 2019.

[116] F. B. Wadsworth, J. Vasseur, and D. E. Damby. Evolution of Vocabulary in the Poetry of Sylvia Plath. *Digital Scholarship in the Humanities*, 32(3):660–671, 2017.

[117] N. Weinman, T. Barik, and S. M. Drucker. Fork it: Supporting Stateful Alternatives in Computational Notebooks. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI)*, 2021.

[118] D. Weiskopf. Uncertainty Visualization: Concepts, Methods, and Applications in Biological Data Visualization. *Frontiers in Bioinformatics*, 2(February):1–17, 2022.

[119] M. Whitelaw. Generous Interfaces for Digital Cultural Collections. *DHQ: Digital Humanities Quarterly*, 9(1), 2015.

[120] J. V. J. V. Wijk. Bridging the Gaps. *IEEE Computer Graphics and Applications*, 26(6):6–9, 2006.

[121] F. Windhager, P. Federico, G. Schreder, K. Glinka, M. Dork, S. Miksch, and E. Mayr. Visualization of Cultural Heritage Collection Data: State of the Art and Future Challenges. *IEEE Transactions on Visualization and Computer Graphics*, 25(6):2311–2330, 2019.

[122] F. Windhager, V. A. Filipov, S. Salisu, and E. Mayr. Visualizing Uncertainty in Cultural Heritage Collections. In *EuroRV3: EuroVis Workshop on Reproducibility, Verification, and Validation in Visualization*. The Eurographics Association, 2018.

[123] F. Windhager, S. Salisu, and E. Mayr. Exhibiting Uncertainty: Visualizing Data Quality Indicators for Cultural Collections. *Informatics*, 6(3), 2019.

[124] F. Windhager, S. Salisu, and E. Mayr. Uncertainty of What and for Whom - And Does Anyone Care? Propositions for Cultural Collection Visualization. *Proceedings of the Workshop on VIS4DH*, 2019.

[125] J. Wooseob. Media visualization of book cover images: Exploring differences among bestsellers in different countries. *DHQ: Digital Humanities Quarterly*, 11(3), 2017.