

Introducing Layers of Meaning (LoM): A Framework to Reduce Semantic Distance of Visualization In Humanistic Research

Houda Lamqaddam, Andrew Vande Moere, Vero Vanden Abeele, Koenraad Brosens and Katrien Verbert

Abstract—Information visualization (infovis) is a powerful tool for exploring rich datasets. Within humanistic research, rich qualitative data and domain culture make traditional infovis approaches appear reductive and disconnected, leading to low adoption. In this paper, we use a multi-step approach to scrutinize the relationship between infovis and the humanities and suggest new directions for it. We first look into infovis from the humanistic perspective by exploring the humanistic literature around infovis. We validate and expand those findings through a co-design workshop with humanist and infovis experts. Then, we translate our findings into guidelines for designers and conduct a design critique exercise to explore their effect on the perception of humanist researchers. Based on these steps, we introduce Layers of Meaning, a framework to reduce the semantic distance between humanist researchers and visualizations of their research material, by grounding infovis tools in time and space, physicality, terminology, nuance, and provenance.

Index Terms—Infovis, Humanities, Digital Humanities

1 INTRODUCTION

"Google 'visualizations'! [...] It's a rich panoply of ugly, conceptually bankrupt, and morally and ethically and intellectually useless visualizations for the humanities." [15]

It was in these words that humanist and visual theory scholar Drucker [15] provoked the audience of the 2016 Info+ conference¹, expressing her perception on modern visualization tools for the humanities. While it may seem surprising considering the broad spectrum of tools introduced by the infovis community, this claim echoes several voices within humanistic fields. Seen from a humanistic point of view, visualization can be found to have 'the air of an alien element' [23]. More concretely, researchers have critiqued infovis tools for, among other reasons, conducting to data reduction [44], centering concerns from foreign disciplines [42], and containing misleading rhetoric [5].

Research about adoption and user trust of digital tools and visualization in DH fields [20, 71], combined with humanist critique of current approaches [44], uses [25], and perceptions [21] of infovis, lends support to Drucker's skepticism. Today, visualization remains largely under-utilized as a humanistic research tool [20]. Moreover, the relationship between infovis and Digital Humanities (DH) is often thought as a one-directional, service oriented help from visualization to humanists [2]. Historically, novel visualization tools for the humanities context have suffered from two important issues: (1) the lack of input from the humanists who will in the end be the tool's main users [71], and (2) different expectations from the infovis and humanities communities which means novel infovis contributions risk being too technical for the DH venues, or too simple for VIS venues [31].

However, visualization research about humanistic domains has grown significantly in the recent year. Specialized venues such as the VIS4DH workshop (2016-present) have contributed to the creation of a shared space for humanists and infovis researchers to contribute and discuss tools, and methodology within specific fields [39, 48, 68], or about the DH at large. The growth of this community, in turn, points

to the emergence of a new research field at the intersection of visualization and humanistic concerns. At the same time, researchers have dug into the specificities of humanistic data, tasks, and epistemology. They have advanced the visualization field by building models, conducting surveys, and proposing novel modes of collaboration between the two fields. Among them, Windhager et al. [73] surveyed a broad variety of datasets, and proposed mature efforts of data models within the field. Others have examined the collaborative research process between infovis and the humanities. For instance, Hinrichs et al. [26] propose the role of visualization as a research thinking process between the two domains, where the prototypes built act as an important knowledge creation process, and benefits both parties involved. More recently, Schetinger et al. [55] have argued for a re-purposing of the users-data-tasks triangle used in Visual Analytics to better accommodate DH concerns. In parallel, visualization researchers have also invested in designing tools, and applying user-centered methodology in order to create tools that best support humanists. Tools like Poemage [46], for visualizing sonic topology within poetry, or StoryFlow [41], that visualizes dynamic story lines, show effective visualizations built around humanistic data and tasks. Several researchers also explicitly included user-centered approaches in their methodology. In recent years, we can cite the work of Valdivia et al., who report collaborating with historians to develop use-cases and test their PAOH interface [65]. Pfluger et al. used their own background in art history to develop a visualization system, before evaluating it with field experts [51].

As an application field, the humanities form a unique neighbor to visualization research. Humanities researchers build on established validated research methods and complex theories of knowledge and critique. They also bring novel, dynamic, and rich data to the visualization table. On the other hand, visualization of data can provide a radical new way of looking at humanistic material, therefore expanding the insights humanists can build. Through methods such as 'distant reading', introduced by Manovich [45], infovis enables users to look at their material through a new lens and gain previously unattainable perspective. Yet there seems to be a disconnect between these two worlds, prohibiting adoption of infovis tools by humanists [5, 15, 71]. How, then, can we better understand the disconnect between humanistic research data, tasks and culture and traditional visualization approaches? Can involving humanists in the research process help identify missing themes for visualization for the humanities? What elements are critical in making better visualizations for the humanities?

To answer these questions, we first started by exploring the humanistic literature on the topic of information visualization. We examined the characteristics of humanistic 'data', the critical perspective on infovis

All authors are affiliated with KU Leuven, Belgium. E-mail: Firstname.Lastname@kuleuven.be

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.xxxxxx

¹<http://informationplusconference.com/>

from a humanistic standpoint, and the sources and obstacles to user trust. We describe these findings in Section 2. Based on our findings from the literature, we also recognized the importance of building interdisciplinary collaboration in order to identify challenges and possible directions. We therefore held a collaborative workshop based on the Creative-Visualization-Opportunity Workshop structure devised by [35]. In this workshop, we gathered researchers in HCI/VIS and in the humanities with the goal to explicit needs, possible directions, and recurring themes. These findings, along with the methodology behind the workshop, are described in Section 3. We used the findings from Section 2 and 3 to create an initial set of guidelines addressing the challenges of designing infovis tools for DH research. In order to evaluate the accuracy and impact of our guidelines, we designed our third and final step: a design critique exercise and semi-structured interview with 12 humanist researchers. This study allowed us to scrutinize and build upon our previous conclusions. Section 4 describes the methodology and the results from this step. For each section, we synthesize the major findings from the corresponding step, as identified by the first author, and validated by two of the co-authors. These findings are then used as basis for each of the guidelines of the framework we contribute.

Based on these steps and findings, we introduce the concept of semantic distance to the information visualization context. This notion is used in the field of computational linguistics to refer to the relatedness between words and concepts [4]. We apply it to the context of infovis to refer to the degrees of separation between humanistic concepts and their representation within visualization tools. This distance is semantic because it is based on a gap in perceived meaning between users' knowledge and the shape in which this knowledge becomes translated through visualization tools. It can be short: for instance, for visualizations of purely quantitative datasets, or ecological visualizations of cartographic data onto maps. But it can also represent a wide gap that impacts the users' perspective on the tool, for example when rich audiovisual data points are exclusively reduced to their metadata and represented quantitatively only. We present our final findings in the form of Layers of Meaning (LoM): a framework to reduce the semantic distance in visualization for humanistic research. With this framework, we intend to center the critical aspect of *meaning* and connection with material in humanistic research practice. We propose five directions to reconstruct this connection by injecting meaning into different layers of infovis tools, thus connecting it semantically to original characteristics of humanities material and research practice. These directions are built on the findings extracted from each step. They consist of grounding in time and space, physicality, terminology, nuance, and provenance. We argue that by implementing Layers of Meaning, designers can improve understanding, interpretation, and overall experience of infovis tools to users as it allows them to make use of humanistic skills and conventions to read visualizations of data. This framework can be used by researchers and practitioners to design infovis tools specific to humanities datasets and researchers.

2 LITERATURE REVIEW

What explains the negative perception of visualization - and digital tools in general - in the humanities? To address this question, we explore four parallel paths: the characteristics of data itself within the humanities, user trust as a critical component to adoption, issues stemming from interdisciplinarity, and finally, current debates within infovis.

2.1 A culture of 'data' in humanities research

The category of 'humanities' includes a wide range of disciplines, including literature, philosophy, rhetoric, classics, history, study of art, music, and design [5]. These disciplines all present specific practices, cultures, and histories, making the task of defining the data characteristics for DH, virtually impossible. Interestingly, one has to look outside of traditional humanistic literature to find definitions and categorizations of data within humanistic research. This can be explained mainly by the inadequacy of 'data' as a concept to describe the reality of humanist researchers' material. While in infovis, data is most often a natural building block of research, this reality does not directly translate to humanist approaches. To illustrate this point, Posner asks to

imagine an external person referring to one's family photo album as a dataset [52]. This analogy helps to exemplify three axes of discomfort: (1) the reduction of an object of immaterial worth to its potential value once mined/summarized/computed, (2) the invitation to imagine said photo album as a datasheet knowing that most of its value will not translate to such a format and (3) the realization that this discomfort is not obvious to this external person, when it is so critical to the photo album owner. In the same talk, Posner also explains:

When you call something data, you imply that it exists in discrete, fungible units; that it is computationally tractable; that its meaningful qualities can be enumerated in a finite list; that someone else performing the same operations on the same data will come up with the same results. This is not how humanists think of the material they work with [52].

Research in DH gives us a better look into concrete characteristics of humanistic material that has already been translated into manipulable data. Schöch identifies two types of such datasets: structured, clean, explicit data (smart data) and voluminous, varied data (big data) [56]. Other research looks at components of humanistic material: on one hand annotations, connections, reviews, and comments deriving from the research process; on the other hand, the concepts, persons, works, materials themselves [6]. This definition is the closest to the data/metadata classification that is common in data science, although characteristics of both the data and metadata at hand are still of a wide range. Seemingly, this data does not fit into one set of categories or qualifiers. It can be qualitative, quantitative, or any other type that does not fit into this binary division. It can be collected, but also created, structured inherently, or unstructured, ongoing by nature, analog, non-discrete [64]. The language, concepts, and entities it consists of can themselves belong to semiotic systems meaning that their dimensions cannot be physically measured, or that their interpretation is heavily tied to their context [18, 56]. Moreover, these datasets are typically incomplete [23], contrary to other disciplines where the institution performing the research is often the one producing the datasets, as humanists work with data from different sources, that were not generated with their research goals in mind [23]. These characteristics, along with the ambiguity, uncertainty and contradiction that define humanistic data tradition, are now sought to be embraced [14] rather than shunned.

At the same time, humanists have reflected on the added value of information visualization for their research practice. What is lost in information and value inherent to transforming rich works of art to easily manipulable variables [23]? Manovich [45], who proposed the definition of information visualization as data reduction, insists that infovis robs humanistic materials from their nature and unique attributes. "Those who study the humanities", he states, "always focused on analyzing and interpreting details of the cultural texts, be they poems, paintings, music compositions, architecture or, more recently, computer games, generative artworks and interactive environments. This is one of the key differences between humanities and sciences - at least, as they were practised until now." [44].

2.2 Usability, trust, and adoption

Research on infovis tools in the humanistic context is still at an early stage. To the best of our knowledge, no work has yet looked at the specific adoption rate of infovis tools by humanists. Therefore, we look into literature around the adoption of digital tools in general by humanist researchers.

Based on literature around use of digital tools in humanities domains [5, 71], we find that major obstacles of adoption include concerns around data quality and usability issues. Warwick et al. [71] cite data quality as a critical element for user adoption. Most users will abandon a digital tool when in doubt about the quality of its sources [71]. Assessment of validity and *provenance* of sources is an essential element to humanistic research [66]. For these, scholars rely on a "complex repertoire of information skills" [71] that do not translate well to digital environments. An example of these information skills given by Warwick is the ability to deduce the usefulness of a book based on elements such as its design, its publisher or even its *size*. Warwick suggests that

users will be more likely to trust digital resources that are clear about purpose and identity of creators, process of gathering and selecting material, as well as the rationales behind technical decisions [71]. Concerning usability, research suggests that faced with usability constraints users are more likely to be deterred by a digital resource than by a physical one [71]. Recent work suggests that indeed basic usability issues prevent true acceptance and adoption of the tools [61], and that lack of integration of stakeholders in the development process is a major reason for that. Finally, from a visualization perspective, humanistic material brings about a set of open challenges. Janicke [32] reports geospatial and temporal uncertainty, visualizing transpositions of text, innovative techniques for close reading and workflow reconstruction as major challenges for humanistic datasets. The topic of uncertainty in humanistic datasets is further developed by Edmond [17] who characterizes it as inherent, partial, due to errors or bias. Therefore, mistrust can be caused by anything from incomplete information to tools looking dated, containing broken links, or simply a ‘sense of something wrong’ [17].

2.3 Tales of interdisciplinary collaboration

Collaboration between humanistic domains and visualization research faces a set of issues inherent to interdisciplinary research. For problem-driven visualization - i.e. where designers aim at solving the real-life problem of a group of users [57] - an encounter with the target users’ field and practices is natural. It also brings about a set of gaps and pitfalls. Van Wijk [69] suggests two main gaps that can occur during interdisciplinary infovis cooperation: a knowledge gap (if collaborators do not have the basic understanding of each others’ concepts and needs) and an interest gap (if the usefulness, and therefore the field interest in a novel approach is not gauged the same way by each collaborators’ field). McKirby and Meyer [36] identified axes for these pitfalls along the lines of language (using a common language rather than the breadth of both collaborators’), outlook on the other discipline (thinking one’s to be objectively better, or more precise), not defining the actual visualization need of a project, ensuring you stay an expert in your field, and being mindful of the level of interdependence of the collaboration.

In the recent years, multiple methodologies have been introduced to effectively address these pitfalls. For example, Simon et al. [58] have proposed the liaison role as a solution to interdisciplinary collaboration gaps. The liaison is a person with expertise in both visualization and target domain who will act as a buffer and mediator between collaborators from the two fields. Others have described user-centered design as a critical way (‘the royal road’ according to Van Wijk [69]) to mitigate the reported gaps. Koh et al. [37] introduced participatory workshops that aim to increase domain experts’ awareness of visualization, and improve engagement. Recently, this concept is further developed by Kerzner et al. [35], who developed a framework for conducting workshops aiming at extracting ‘opportunities for visualization software’ from domain experts. These user-centered approaches have been reported to reduce the time needed for problem characterization [35]. Results also suggest that they have improved domain experts engagement and agency within infovis projects [37]. While the aforementioned methodologies are not specifically thought to address concerns of humanities researchers, their benefits highlight the value of user-centered approaches in addressing interdisciplinary gaps. They also position our work within a wide effort to better address user needs across fields of expertise, knowledge and language difference.

2.4 Ongoing debates within infovis

Within the infovis community, debates are taking place that directly echo current humanist critiques. Researchers are questioning the relevance of accepted practices, thus giving rise to new models and directions for the future. In this section, we examine two of these topics: the effect of style and embellishments, and the need for a critical infovis.

Style and embellishments: The debate around visual style in infovis is highly relevant to visualization for the humanistic fields. Humanist literature describes many of the contemporary visualization approaches as having roots in the mathematical language of statistical models [5] and empirical sciences [42]. Humanists have instead called for borrowing from the visual communication design tradition [42], as there

is debate around the adequacy of the mathematical language to represent the rich physicality and nuance of humanist material. In recent years, several visualization scholars have questioned the universality of Tufte’s ink-to-data ratio [63] by examining the effect of stylistic choices and embellishments on information visualizations. Tufte’s push for a reduction of any non-erasable ink within a data presentation, along with the popularity of minimalism in web and design spheres [62] have had a wide impact on the design of visualizations [30], where decorative elements described as ‘chartjunk’ are sought to be virtually eliminated. This approach however, can be at odds with the need to connect with artefacts’ aesthetic qualities, as described by Manovich [44]. Without necessarily carrying data, decorative elements can also contribute to set an immersive environment, or help ground a visual element in a physical practice. Doulikaridou [12] comments on the effect of the addition of CSS borders as *frames* around artwork in a browsing tool for instance, as they link users to ‘the practice of a historically meaningful and deeply familiar process’, therefore positively impacting their experience. Likewise, multiple studies have examined the effects of non-efficient elements on user experience, finding benefits on user interpretation, accuracy and recollection [29, 70]. Other research has delved into the evocative power of aesthetic elements. Through revisiting Bertin’s Semiology of Graphics, Jégou explores the effect of aesthetics on cartography based on principles of semiology, art history and design [33]. He argues that aesthetics are downplayed in Bertin’s approach, as the positivistic, rigorous science aspects of maps are put in opposition to colors, for instance, which are presented as emotional, subjective, almost mystical. Jégou also evokes composition and page layout as aesthetic attributes that ought to be explored deeper for their positive effect on perception [33].

Towards a critical infovis: Humanist literature describes a long-lasting tradition of criticality towards infovis practice. This tradition stems from approaches of *critical theory* aimed at questioning norms and revealing power structures, in the goal of reaching emancipation [19]. In our context, this critique shows a *conceptual friction* between the technical tradition of infovis, and humanistic practice. First, the *approach* to data itself has been scrutinized, as researchers have commented on the dangers and assumptions behind data [27], datafication [67], and information visualization positivist practice [53]. According to researchers, the veracity communicated through information visualization comes not from empirical conclusions or interpretations alone, but from the authority that numbers and charts might inherently convey [25]. The questions of which numbers were selected, through which method, by whom, and for what purpose, as well as which representation was chosen, which data was discarded, and which visual encodings were assigned, often remain unexpressed for simplification reasons. For Burdick [5], the issue with visualizations is that they convey the rhetoric of clarity and precision, even though they are themselves constructed interpretations. Drucker [21] argues that current visualization is *anathema* to humanist practice, as it is based on a reification of information. Therefore, humanists have consistently called for representation of rich textured information in non-reductionist ways, allowing interpretable dimensions, and explicating the hidden assumptions behind the capturing and representation of data. Designers, on the other hand, have often not taken the specificities of humanistic research into consideration while designing for DH, and have sometimes translated lack of engagement with data visualization to signify lack of interest, or knowledge [71]. These field characteristics, challenges and directions raise the need for reflection, within the visualization community, around ways to better address visualization for humanities research. In recent years, several voices have emerged to enrich infovis practice with notions of criticality. Researchers - many of them with interest in humanistic practice - have called for a critical infovis [11], an ethical infovis [8], and discussed the potentials of data feminism [9] and data humanism [43]. More recently, Myer and Dykes proposed a novel perspective on design studies and contributions in infovis as subjective and socially constructed [47]. Within other application fields, other models have been introduced as a push to better address uncertainty visualization [28], propose measures to conserve interpretability and subjectivity [7], or imagine new models of interaction [10, 72].

There is a need for an approach in infovis that better supports humanist researchers in their practice, while addressing questions of uncertainty, visual complexity, contradiction and interpretability. We therefore set out to explore whether we can find new ways to think information visualization for the DH, by including stakeholders in the process, and addressing critiques stemming from humanist research.

2.5 Main Findings

F1: Humanistic ‘data’ has essential material and immaterial aspects. The humanistic approach to *data*, itself a debated concept, is fundamentally different to the one familiar to technical fields. Humanist researchers are acquainted to the analysis of *materials* (artefacts, persons, and pieces) that have distinct material and immaterial aspects. These aspects may often not be reduced to quantities or metadata.

F2: Sources and data quality are central to user confidence and trust. Data provenance is a central aspect of datasets for humanistic practice, with crucial impact on adoption and user trust. It also empowers users by improving *transparency* and *traceability*, enabling them to make their own informed decisions about the quality of data sources.

F3: The visual language of infovis can create conceptual friction. Humanist research describes a philosophical resistance to infovis practices that evolved from statistical visual language and practice. Attempting to create an environment of neutrality can be perceived as reductive and rhetorically misleading. The statistical visual language can also reduce the possibilities of interpretation and nuance within humanistic knowledge.

F4: User-centered approaches have the potential to bridge interdisciplinary gaps, and improve adoption and usability. Several of the described adoption issues can be linked to the lack of involvement of humanist researchers in the design and evaluation process. In order to build tools that privilege usability, adequacy, and support of humanists tasks and domain specificities, user-centered approaches are necessary. Existing user-centered methodologies designed for infovis research can be used for this purpose.

3 CREATIVE VISUALIZATION OPPORTUNITIES (CVO) WORKSHOP

To address the aforementioned challenges, we set out to identify the themes and potential directions for infovis in humanistic research. To address F4, we chose to set up a methodology involving participants from DH and infovis. The aim of this workshop was to examine the reliability of the findings from section 2, identify recurrent themes in humanists’ needs and solutions, and encourage interdisciplinary collaboration. In this section, we describe the process and results from a CVO workshop [35] gathering both end-users and designers.

The participants belonged to two main backgrounds: researchers in humanistic domains (two professors, two PhD students researching art history and philosophy) and experts in HCI and infovis (two professors, one post-doc). In the remainder of the paper, these will be respectively referred as infovis research (VISr) and humanities research (HUMr) participants. They are individually coded as VISr_[1-3] representing VISr participants, and HUMr_[4-7] for HUMr participants.

To begin the workshop, participants were divided in two mixed groups (of 3 and 4 participants). Each group contained participants from both backgrounds. The two groups stayed separate for all remainder activities, only coming together for collective discussion and synthesis between activities. Participants were primed for collaboration and collegiality through a round of Analogy Introductions [35] and the setting of workshop rules. Afterwards, a description of a dataset studied by the HUMr participants was presented to everyone. This dataset consisted of historical artists and their communities, including relationships between persons, events linked to time and space, art pieces, and digitized images of archives. It therefore included tabular, relational, spatial and temporal data, as well as rich visual material. By including different types of data and material, we aimed at representing the variety and richness of humanities datasets. The workshop was structured around three main activities, adapted from [35]:

Wishful Thinking (A1): We present data and domain scenario to participants and ask “What would you like to know? What would you

like to do? What would you like to see?” [35] from HUMr participants, and “what could you extract?” to VISr participants.

Visual Analogies (A2): We show examples of visualisations and ask participants for liked/disliked aspects, and relevance to their domain.

Visual Scenarios (A3): We ask participants to draw a prototype based on the scenarios driven from A1, and the characteristics from A2.

Through these activities, we aimed to extract user data exploration needs and potential (A1), explore preferred and appropriate visualization options for user scenarios (A2), then build on these findings (A3) to identify themes for infovis that best supports humanistic research. The protocol and activities are described in full in supplemental material.

Each of the three activities ended with a collective discussion and identification of general themes. A general discussion was held at the end about the global results and themes from the entire workshop. The main author took note of the gathered themes as testaments of the collectively identified directions. We validated these themes with two of the co-authors after the workshop and grouped them into the themes described below.

3.1 Workshop Results

Connection and intimacy with the data (T1) Several of the described needs were indicative of high level perceptual aims. In the first category, HUMr participants mentioned goals relating to building connection and intimacy with the dataset. This included points like immersion in studied persons communities, as well as time and space environments. Participants mentioned wanting to achieve a feeling of “hearing and smelling” these communities (HUMr_6), and reconstructing art worlds (HUMr_4, HUMr_6).

3D techniques and metaphors (T2): HUMr participants judged visualizations containing metaphors and 3D elements positively. They commented on their “interactive and intuitive” (HUMr_4) nature, and appreciated the metaphor of a 3D timeline to a “walk through time” (HUMr_6). Metaphors reusing codes from the domain field (for instance audio wave visuals for a visualization of music genres) were appreciated by participants from both groups. Participants thought these could give specialised audiences a better understanding of the data through a format they were used to reading (VISr_1, HUMr_7).

Visible artefacts (T3): Visualizations with visible artefacts, and visualizations that centered in on the artist in some way, received a large interest from the HUMr group.

Time and space as backdrop (T4): HUMr-made sketches were characterised by the use of time and space as backdrops for the visualization. For instance, HUMr_4 used a data-comic style organisation as each panel shows a state of the network in time (Fig. 1, a), while HUMr_5 organised theirs in a multi-resolution context+focus set of views, based around an interactive timeline (Fig. 1, b).

Approximations and re-assessments (T5): Participants described wanting to acquire new insight on the data in order to challenge and confirm or reassess traditionally held beliefs. Aims included ideas such as using data-driven methods to draw portraits of “typical” artists or communities (HUMr_4), extracting a more realistic - or more “inclusive” - description of an art community (HUMr_6), or finding out “forgotten” artists that may have had local impact but were not remembered by historians (HUMr_6).

Predictive power (T6): The element of prediction based on mathematical models created interest when put forward by the VISr group. While some participants mentioned mistrust in “the computer telling experts what to do, or what is the truth” (HUMr_5, HUMr_6), supplementary information or possible predictions being suggested by the system at the request of users were seen as enriching and empowering. Therefore, a new possible direction emerged, as HUMr participants imagined the possibilities of “playing with data” (HUMr_6), explained as carrying thought experiments, and seeing potential results of changing small variables on the future of a person or a community.

3.2 Workshop Discussion

Based on the themes described above, as well as the final closing discussion, we discuss the resulting themes of this workshop, as identified by both domain and infovis experts.

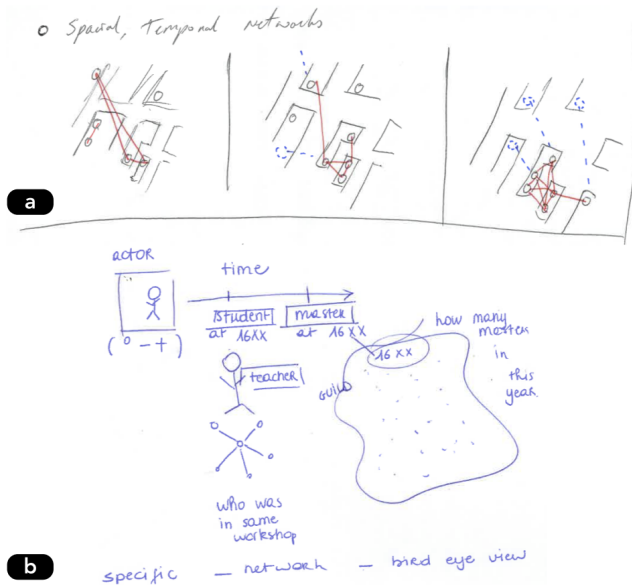


Fig. 1. (a) Data-comic style sketch drawn by HUMr_4. Artists are represented as nodes linked to their geographical location, and relationships between them are represented by edges. The spatial background and the succession of temporal states allow for two extra dimensions to be added to the node-link diagram. (b) Coordinated multiple views drawn by HUMr_5. The focus is on one artist shown at the top left of the view. Other views show timeline of events of said artist, network of collaborators, and overview of entire guild.

3.2.1 Metaphors, Immersion and Intimacy

From results T1, T2, T3 and T4, we can first see that metaphors, images, olfactory and auditory perception being mentioned, point to directions that are centered on user experiences to create a connection with the original material and its context. Physical artefacts are characterized by layers of meaning that can be exploited for their digital representation. These semantic characteristics are often lost during the translation of material to on-screen representation. From these first themes, we derive the need to make use of the features of these physical attributes, as well as contextual elements of time and space, to recreate a sense of connection with the data. This theme also echoes Manovich's call for direct visualization. A semantically consistent representation of the content data can be achieved by retaining and reproducing characteristics of the original material. This view is supported by other voices in the field [15], as researchers have deplored the semantic dryness of data visualization when compared to the richness of iconography, sense-related information, and cultural elements of other visual media. In the infovis literature, several tools show examples of metaphors in practice [22, 40, 50, 54]. Metaphors were also a consistent subject throughout the workshop (T2), as they broaden the possibility of data representation outside of beaten paths of edge-node diagrams and grids. Similarly, while the downsides of 3D representations are well described in the literature (e.g. occlusion [3, 24], potentially disorienting navigation [3], added cognitive load [38]) these were appreciated by HUMr participants for their immersive attributes. In the context of our workshop, possibilities to recreate this immersiveness were discussed. Since part of the dataset concerned tapestry producers and artists, the image of weaving visualizations was discussed as a potential direction, making use of the techniques of drawing, painting and weaving to communicate information to researchers, all the while retaining the tacit dimensions and conceptual universe of the studied material. Indeed, features of pictorial art remain largely unexploited within visualization for humanities [15]. Specifically combined with conventions belonging to the targetted humanistic domain, this opens a new breadth of possibility for visual design for infovis.

3.2.2 What-ifs and thought-experiments

While HUMr participants had low interest in fully predictive or highly processed data representations, they were enthusiastic about data-driven models that could process large datasets and then only feeding results in the form of contextualized elements that still support users' mental models: e.g. a portrait of a 'typical' actor, community, or a potential alternative life path for an artwork or person. Some of these findings have also been reported in the visual analytics literature [59]. However, previous work has described current limitations to this approach. Indeed, the size of the studied dataset has to be large enough to allow for sensible data-driven predictions. In his discussion of data usage in the humanities, Schöch [56] gives an example in literary analysis:

For the British nineteenth century novel, for example, the calculations go as follows: an estimated 20-30,000 novels were published in the nineteenth century; of these, only around 6,000 are estimated to be still existent in the holdings of libraries and private collections worldwide. Of these 6,000 novels only about half, that is 3,500 novels, have been digitized in full text mode and with sufficient metadata. [56]

The final size of the dataset in Shoch's example is effectively equivalent to 20 percent of the initial production [56] [34]. However, the quality of results of prediction models is highly dependent on the size of the corpus. Smaller data samples lead to unreliable non-generalisable output. For this approach to be applicable to humanist data, we would need to consider larger data corpora, as well as specifically address a reported mistrust of data mining within the humanities [13].

3.3 Reflecting on the CVO workshop format

In the previous section, we describe the aims and the results from the workshop. This section discusses the differences between user backgrounds that we have seen during it.

Complementary contributions: VISr and HUMr participants contributed different but complementary aspects to the workshop in terms of nature and goal. The elements put forward by HUMr in A1 gave the workshop a direction in terms of requirements and challenges (T1, T2). Predictive power, the major contribution from VISr in A1, complemented that by providing a possible way to answer some of those challenges with the use of data-driven methods (T3). Outside of this example, we found that participants from both backgrounds were rigorous in defining and explaining their decisions, based on their own expertise. They were also showing encouragement to pursue novel directions, when these aligned with their understanding of their field.

Different sought characteristics: the different complementary perspectives by different user subgroups was visible during A2 as well, as the Visual Analogies reviews differed largely between the two groups. VISr participants gave better reviews to visualizations showing more dimensions, even when that was at the expense of instant legibility, while HUMr preferred visualizations that were easily legible, focused on aesthetics, and included concrete links to the represented information (T2, T4). Visually complex ones were described as being "scary for us" (HUMr_4), or requiring a "high level of visual literacy" (HUMr_6) by HUMr participants. This corroborates previous patterns reported in the literature around digital humanities. In a 2012 paper, Gibbs et al. report users preferred easy-to-use tools over polished ones [20].

Finally, the mock-ups showed visualizations inspiring themselves from the contributions of each other, as well as from the results of previous activities. We found the separation of tasks and expectations by user group to be conducive of collaboration and cross-pollination.

3.4 Main Findings

F5: Users aim to experience connection with digitized material. We described in F1 the material and immaterial aspects of artefacts in DH research. We find that humanists' intimacy with their research material is expected to translate to digital representations, through immersiveness, time-space contextualizing, and direct visualization.

F6: Visual conventions familiar to humanists can enrich representation of their material. Approaches found in pictorial and visual

arts are underutilized in infovis. Combined with conventions from humanists' specific domains, they can be used to create immersion and fit in with users' conceptual universe. This finding also relates to F3, as humanists prefer a visual language built around their research practice.

F7: VIS and humanist experts within a CVO workshop act as both barrier and stimulator for one another. Participants from each background took turns in halting trains of thought from the other group, and encouraging others, based on their domain knowledge. Participants from both field backgrounds therefore acted as *liaisons*: safekeeping their expertise, while enthusiastically exploring that of the other group.

F8: Semantic distance is a central issue in infovis tools for humanistic research. We identify a crucial gap between humanists' experience of their material and research practice on one side, and the representation of said material in infovis tools on the other. This gap can best be described as the loss of *meaning* that occurs when rich qualitative material is stripped of its characteristics and translated to dots and lines on a chart. We refer to this gap as *semantic distance*. This distance translates into a difficulty for users to build connection to their researched material, thereby affecting adoption.

4 DESIGN CRITIQUE EXERCISE

Based on the findings in sections 2 and 3, we identified semantic distance between original material and data representations as a critical aspect to enhance infovis practice. We therefore synthesized the findings described above into 3 main axes that we iterated over with all authors of this paper. We then translated these axes into guidelines to investigate their impact. In this section, we describe these guidelines, then outline our protocol for assessing their impact on humanistic users' experience with infovis tools. Finally, we present the results from this study and the main findings it reveals.

4.1 Guidelines

We extracted three axes for reducing semantic distance in infovis for humanistic research. We iterated over these axes through three rounds of reviewing by each of the authors. The results are described below.

(1) *Visual framing (F1, F5)* Infovis tools should contextualize data representations by visually linking to original objects of study. Through sidebars, linked views, or inline media, visual layout can introduce contextual elements tying the data to the original artefacts, for instance to include works of art, actors, or time/space as anchoring elements. This way, the distance between users' experience with physical objects and the abstract representation of data elements is reduced, and connection to material can be recreated.

(2) *Encoding channels (F1, F3, F6)* Some data channels (area, length shape) are commonly used to communicate precise quantitative values. Nuanced interpretable variables can be shown through interpretable encodings. Shading and tone for instance can assist in representing uncertainty or disagreement.

(3) *Data transparency (F2)* Largely discussed in the literature, source validity is an essential element to humanistic involvement with studied material. However, dataset sources are not always fully described within visualization tools. Adding a link to the data source or an image of the archival material allows users to make use of their own judgement about the data reliability and their trust in it.

4.2 Phase A: Design

Recruitment. We recruited five PhD students involved in visualization research to design prototypes based on provided datasets (more about the datasets below). These participants were recruited from the research group of the main author, and they were not familiar with the theme of this study or its goal. We will refer to these as Designers[1-5] to differentiate them from the participants of Phase B, described below.

Data. We proposed three datasets that exemplify tree main types of data [49]: tabular, relational and spatial. We chose datasets within humanities domains: A collection of artists exposed at the Tate museum [60], historical paintings of the Mediterranean region, [1], and a corpus of graphic novels [16], that we transformed into a relational dataset by adding mock links between data points. All the subsets collected were trimmed to contain the same number of nodes.

Tasks. Each designer was instructed to design prototypes for visualizations for the task of initial exploration of a corpus by domain experts. They were free to use tools they would typically use for prototyping (i.e. drawing software such as Figma or Illustrator and graph-making tools such as Tableau and Gephi). They were also provided with additional media related to the dataset (e.g. book covers, photos of artwork) that they could optionally use. Interactivity was not required, as end users would see a static version of the visualizations. It could, however, be implied (the prototype could show a state after interaction for example). Each designer was instructed to spend a maximum of 1h to design each visualization. At the end of this first step, we gathered participants' initial designs, and briefed them for step 2 of the design phase. For this step, we sent a brief consisting of the three initial guidelines described in section 4.1. We then asked the designers to adapt their designs taking into account these guidelines. The full instructions given to the designers in this step can be found in the supplemental material.

Outcome. This process resulted in a pair of submissions from each designer. The set of initial submissions will be our 'without-guidelines' condition, and the modified ones will be our 'with-guidelines' condition. After filtering out the lower-quality submissions (i.e. where at least one of the pair presented readability issues), we were left with 6 pairs of prototypes: two pairs for each dataset type.

4.3 Phase B: Interviews

Recruitment. We recruited researchers in various humanities fields for a study designed to gather their insights and preferences around our two sets of visualizations. All participants were researchers affiliated with the authors' university. They were chosen through a random selection within the university's Arts and Humanities department members. From the 40 persons contacted, 12 agreed to participate (6 of them women, and 6 men). They were compensated with a 10 euro gift-card to a retail webshop. As can be seen in Table 1, participants belonged to a wide range of humanities domains, and consisted of seven PhD students, two professors, two post-docs and one researcher.

	Field	Position
1	Contemporary political French philosophy	Postdoc
2	Applied linguistics	Professor
3	Ancient history: Economy in Ancient Egypt	PhD
4	Arabic and islamic studies	Professor
5	Corpus linguistics	PhD
6	Musicology	PhD
7	Political philosophy	Researcher
8	AH*, cultural economics and economic sociology	PhD
9	AH*	PhD
10	Medieval music and medieval liturgy	Postdoc
11	English Literature: Narrative theory	PhD
12	Philosophy of art	PhD

Table 1. Field of practice and position of each of the 12 participants to Phase B of the design critique exercise. AH* stands for Art History.

Tasks. We printed the submissions collected in Phase A and sorted them in pairs showing the same dataset and authored by the same designer. We randomized the order (without-guidelines first or with-guidelines first) to mitigate bias. We iterated through the entire list pair by pair in random order, and asked participants to rate each visualization on a scale of 0 to 5, based on preference. Then, we interviewed participants using the following open-ended questions:

- (1) Can you elaborate on the ranking you gave?
- (2) Is there a version you trust more than the other?
- (3) Is there a version you would like to use more than the other?

Afterwards, we selected the pair(s) with the biggest difference in ranking for an in-depth interview.

4.4 Data analysis

The main author took notes during the sessions, collecting the answers to each question separately. We recorded the values given by participants for quantitative questions, and took notes of the verbal comments.

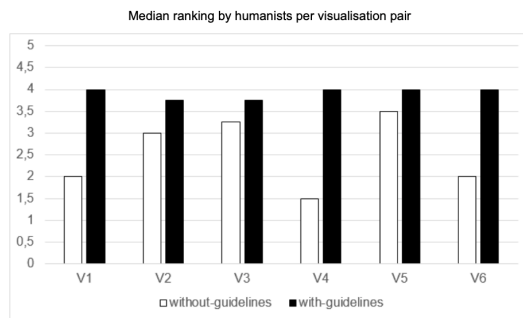


Fig. 2. For all visualization pairs, the design with-guidelines received a higher median ranking from humanists than the one without-guidelines.

We then performed a thematic analysis based on the results to the qualitative questions. We started by a familiarisation phase to identify recurrent characteristics and built a list of categories. We then iterated over them twice in order to refine the themes.

4.5 Results

4.5.1 Phase A: Designs

We found that guidelines were not all implemented by all designers, who reported different levels of ease with each one of them. The first element, visual framing was added through elements such as timelines and spatial distributions. Interestingly, a few of the initial submissions had already included images of the original material, however, none of them included sources or data provenance. Designers reported finding it easy to include sources in the visualization after being primed to do. Overall, the most difficult guideline to implement was reportedly Guideline 2, around encoding channels, with only one submission introducing elements of uncertainty that existed in the provided dataset (see figures. P1_T and P1_T_g of the supplementary material). This result corroborates the literature on the matter, as research suggests the lack of reporting on uncertainty as a field norm, even as visualization designers recognize its importance [28]. The full collection of submissions can be found in the supplementary material.

4.5.2 Phase B: Ranking

Submissions in the with-guidelines condition scored higher (average rating: 3.91/5) than the ones without guidelines (on average: 2.5/5). Figure 2 shows median rating by humanists per visualization pair.

4.5.3 Phase B: Trust and Preference

On average, the visualizations perceived as more trustworthy belonged to the with-guidelines condition in 66.7% of cases, versus the without-guidelines condition in 6.9%. In the remaining 26% of cases, participants reported finding no difference in trustworthiness. Presence of a dataset source was the most cited reason for trustworthiness (19 occurrences). The next most cited reason (6 occurrences) was the perception that a version showed more information, including uncertainty when it occurred in the dataset, as the prototypes were perceived as more 'honest' [P2, P3]. Finally, other justifications included the perceived beauty of the tool, as well as other elements such as colors, quality of legend and text, and perceived 'professional' style.

Regarding intent of use, participants preferred visualizations in the with-guidelines condition in 77.8% of cases, versus 15.3% in the without-guidelines condition. In 6.9% of cases, participants reporting having no preference. The most cited justifications for participants preference were ease of use and additional information. Others invoked elements included chosen symbolism and direct access to material. Figure 3 shows the breakdown of trust and intent of use per visualization and per condition.

4.5.4 Phase B: Themes

Direct access to material: Overall, we found that description of data provenance was the single most critical element in defining preference

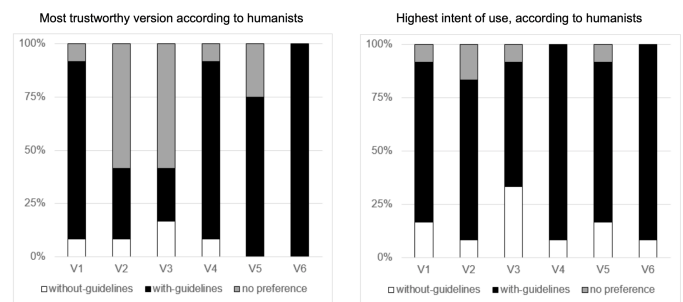


Fig. 3. Humanists ranked the designs with-guidelines as more trustworthy (left). They also indicated higher intent of use than for the without-guidelines designs (right).

of participants towards one condition or another. P11 appreciated "the fact that elements [of the prototype] are clickable, and take you to the original information". P12 described that if they did not have a way to go back to the original information, the tool can be useful "but is missing something critical that would allow me to go judge for myself. Direct access to data is the most important thing." Participants also criticized the lack of direct inclusion of artefacts in visualizations. "Since it is about graphic novels, we want to see the covers directly"(P2).

Meaning of visual symbols: Participants were sensitive to meanings and connotations of visual choices taken by the designers. For instance, the prototype shown in Figure 4 was criticized for the choice of the cross symbol to represent artists' deaths. P1 described being "annoyed" at the choice of an icon so closely tied with Christian symbolism, which was not found adequate for the region portrayed. Another prototype showing a map of the world, was also negatively perceived due to the presence of visual elements that obscured a part of the globe. According to P3, "This discards half the world, it's not cool. If there's a south-African artist for instance, you can't see them, it just doesn't work."

Aesthetics and style: The perceived beauty of visualizations was described by many participants as an important factor. This aspect appeared in justifications of preference, intent of use, and even - as mentioned in the previous section - trust. However, this reached a limit as the visualizations were perceived as 'too commercial', when they seemed more polished, and were therefore associated with commercial websites. P3 commented that a prototype was "more visually pleasing, but it reminds me more of a commercial website than something you would use for research". This, in turn, affected the trust of participants negatively. Another participant expressed: "this version is less pleasant, because it doesn't have any images, and it's supposed to show painters. Although it seems more scientific". This tradeoff was noted by P8 who noted that a representation seemed more scientific "only because it is more boring. I feel like in academia, boring is considered better. People take it more seriously if you *don't* have a nice visual".

4.6 Main Findings

F9: Data provenance and transparency have a direct impact on user trust. We found the visibility of data provenance to be the single most impactful variable on users' trust of an infovis tool. It gives users control by allowing them to make use of their expertise to define their trust in the tool. Transparency also enhances trust in the tool designers, who were judged to be more 'honest' when divulging sources and reporting uncertainty.

F10: Humanists approach concepts within infovis tools and their meaning with criticality. Humanists are sensitive to the meaning and connotation of concepts used within an infovis tool. These include concepts that are tangential to the main task, such as icon choice.

F11: Ugly tools are untrustworthy; Commercial tools are inadequate. Aesthetic quality and style of infovis tools has impact on trust and intent of use. However, overly polished tools are associated with commercial, rather than scientific use.

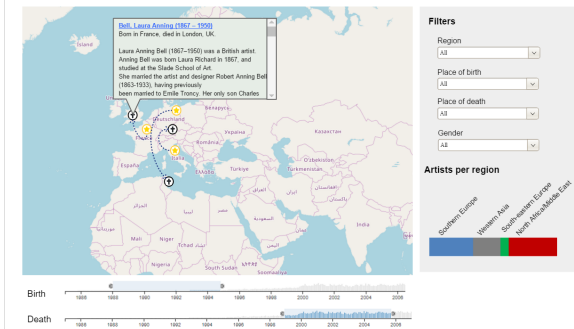


Fig. 4. Design submitted by D4. Humanists critiqued the choice of the cross symbol to denote death, when this symbol is not universal, nor necessarily relevant for the displayed region.

5 LAYERS OF MEANING: A FRAMEWORK TO REDUCE SEMANTIC DISTANCE FOR INFOVIS IN DH

In this section, we describe our main contribution in the form of an encompassing framework for visualization in humanistic research.

Based on the main findings from sections 2, 3 and 4, we propose *Layers of Meaning*, a framework aimed at enriching the semantic expressiveness of infovis tools for humanities research.

Looking at the findings that we defined in the previous sections, we find that semantic distance of infovis can be reduced by grounding it via two main categories of axes: **research material**, and in **research practice**. Reducing the distance to research material can be done through grounding visualizations in elements that contextualize and connect to this material, namely: time and space, and physicality. On the other hand, reducing the distance to research practice can be achieved through grounding in terminology and nuance. Finally, the two categories come together through grounding in provenance. This section explains each of these axis, and provides recommendations on how to implement them. Figure 5 provides a synthesis of this section by showing how the two main categories relate to the five axes of the framework. Each element in the diagram shows an axis of the framework, describing its meaning, and recommendations for implementation.

5.1 Reducing distance to research material:

An important part of building humanistic knowledge revolves around the analysis of material - be it artefacts, persons, or works. In order to provide context for this material we recommend reducing the semantic distance to the artefacts by grounding visualizations in **time and space** and in **physicality**.

5.1.1 Grounding in time and space (see: F5)

Description: Humanistic data often does not consist of individual discrete elements and is sought to be analyzed in relationship to its context. We mean by ‘Grounding in time and space’ the framing of the visualization in general, or data points in particular, in contextual elements of time and/or space, with the goal of providing background to data points’ existence and relationships.

Addresses: Lack of background needed for material analysis

Recommendation: Grounding in time and space can be achieved by inline elements such as labels and tags that link data points to their space and time context. If needed, this grounding can also be done through larger visual elements such as map backgrounds or timelines structures.

5.1.2 Grounding in physicality (see: F1)

Description: We have seen that stripping physical historical or archival materials down to their metadata, reduces the information they contain, and the extent with which researchers can interpret them. Visualization

tools should ground the visualization in physicality by accompanying views with elements of the physical material.

Addresses: Difficulty of identification and connection with material

Recommendation: Grounding in physicality can be done through keeping the material’s texture, visual characteristics, as well as including images or renderings. Including original material also reduces the gap between physical material and its digital representations, allowing to recreate connection and intimacy in users.

5.2 Reducing semantic distance to research practice:

Beyond connection to research material, infovis tools should also aim at closing the conceptual gap towards humanistic research practice. We propose that this can be done through grounding visualizations in **terminology** and in **nuance**.

5.2.1 Grounding in terminology (see: F10)

Description: Humanistic research is rich with conceptual definitions, analyses, and terminology. Humanistic concepts already exist within well-established structures and ontologies. This information architecture will not necessarily be found to be the most appropriate for the visual representation. Still, the use of original concepts and structure within the visualization supports reading by situating the information inside the views into the mental models of the users.

Addresses: Disconnect between user language and conceptual dictionary within digital tool

Recommendation: Designers should include categorisations based on humanist users’ conceptual dictionaries. This can be done by using the same terminology as in the original data, and implementing a similar conceptual structure via the chosen information architecture. Through this approach, the distance between users’ knowledge and the data structure suggested by the visualization is reduced.

5.2.2 Grounding in nuance (see: F3)

Description: Humanistic knowledge is characterized by nuance, interpretation and ongoing debate. Visualization designers should accept and acknowledge the discomfort of uncertainties, contradictions and disagreements within datasets. Understanding these nuances is necessary to better apprehend the field and its culture. It also prevents designers from inadvertently reducing complex nuanced arguments and knowledge into simple binary values. Although it can be at odds with the precise quantification approach needed for developing infovis tools, interpretation of this nuance is an essential tool in the humanistic research toolbox. Building visualizations that leverage humanists expertise of analysis and interpretation is a way we can ground the data in the nuances of humanistic research practice.

Addresses: One-dimensional representations of complex, debated or uncertain facts

Recommendation: Visualizations need to make use of dimensions that express nuance, intensity, and concurrence in order to encode aspects such as nuance, uncertainty, and disagreement. Perspective, shading, tone, position in the view (center or margins) are all channels that were identified in the literature as adapted to communicate these additional dimensions.

5.3 Bridging between research material and practice:

Finally, we recognize that the recurrent emergence of provenance as a finding shows its crucial aspect in the association between infovis and humanist material. Moreover, we find that it has the ability to crystallize the connection between research material and research practice. Provenance bridges the path from artefacts and knowledge - be they historical or contemporary - to their current digitized representation. We can therefore see it as an articulation of this link between material and practice, between history and present, between physical and digital.

5.3.1 Grounding in provenance (see: F2, F6, F9)

Description: Data provenance is one of the most essential characteristics of material in the humanities. The lack, or vagueness of which has a direct impact on usability and adoption of visualization tools. Including provenance information allows to provide authenticity to

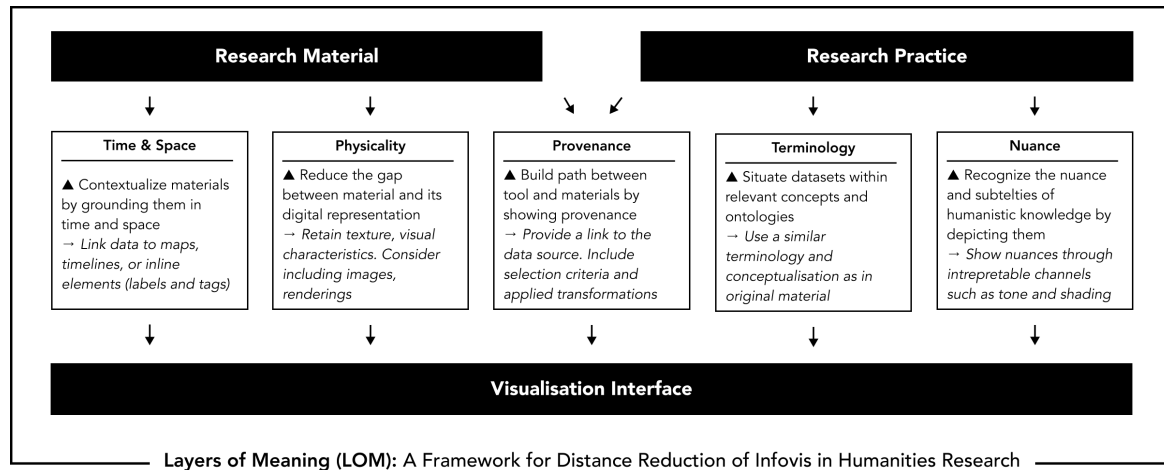


Fig. 5. Diagram showing the five axes of the Layers of Meaning (LoM) framework, divided along two categories.

the data presented. It also empowers users to make use of their own expertise to gauge the credibility of the data presented.

Addresses: Impact in usability if data quality is doubted.

Recommendation: Designers should provide paths to the origin of the represented data in a transparent manner. The source information needs to include not only link to original dataset, but also conflicts within it, decisions taken to remove elements, or additional meta-data.

5.4 Limitations and future works

On the methodology. The elected methodology allowed what we see as an important exploration of the reported discomfort towards infovis by the humanities as an application field. However, we acknowledge two main limitations to it: the sample sizes of the workshop and design critique exercise, and the lack of a longitudinal study design. In the future, our findings will be stronger once confronted with a wider audience of humanists, using a long-term study design that allows users to not only see a visualization prototype, but include it in their research protocol in order to properly assess it. Therefore, we hope our findings will be strengthened and improved through additional user research.

On the relationship among axes, and relevance for users. Regarding the contributed framework, we want to stress that not all axes carry the same relevance for all tasks and users. Neither do they aim at suggesting a common set of features or layout in all infovis tools for the humanities. The goal is rather to address the different challenges reported by humanities researchers, and visually complement other features of infovis tools. We therefore recommend that designers make use of their expertise, combined with user-centered research to interpret these axes depending on the context. We also plan to clarify the overlap and relationships among these axes better in future works.

On the aesthetic effect of the framework axes. Finally, we are aware that some of the axes in our framework will inevitably result in visualization tools needing to host additional visual elements. Additional framing, textures, and process rationales may be at odds with visually minimalist approaches to infovis that mainly aim to keep cognitive load low. However, the literature around both the perception of the neutral chart language, and the evocative powers of embellishments in visualizations suggest benefits to this approach. By grounding the material in rich visual elements, connecting it to its physical existence, we argue that one can accomplish two complementary goals. First, we are able to utilise the vast body of under-utilized conventions and tools of visual languages predating infovis, and in doing so, improve user engagement, recollection and comprehension with the adapted choice of aesthetic elements. Second, we address the discomfort reported by many humanist researchers around the tradition of a scientifically positivist visual language in infovis [53] that ‘do not reflect many of the core concerns and conditions inherent in humanities research’ [42].

Most critically, we argue that information visualization with richer singular visual elements sheds light on the presence of human intervention during the process. This, in turn, can shift the perception to making it appear as the exercise in data gathering, interpretation and transformation that it effectively is.

6 CONCLUSION

In this work we describe how we pursued a multi-step user-centered approach to explore the challenges in infovis for DH. We describe findings from humanist and infovis literature, the process and results from a co-design CVO workshop, and the insight from humanists through a design critique exercise. We found that the workshop was helpful in inciting contributions from interdisciplinary participants. The collaboration between participants from both backgrounds led to the setting of specific use-cases, needs and scenarios on one hand, and on the other, allowed the broadening of possibilities of techniques, tools and visualization approaches to answer these needs. It also enabled us to extract elements of importance from user and designer perspectives, and to create a space of discussion about these trade-offs and how to nuance them. From the analysis of the results and the critical approaches we found in the literature, we implemented initial guidelines and iterated over them through a design critique exercise aimed at understanding user trust, preferences, and experience with visualization tools. Based on these elements, we propose Layers of Meaning, an overarching framework to reduce semantic distance in infovis for humanistic research. Thinking of visualization as layers of opportunities to *inject meaning* and *reduce the distance* to researched material can lead to new directions in designing for humanistic research that allow for rich interpretable visual representations, grounding in time, space and physicality, and concrete links to the studied material. Evidently, this work is only a first step in our attempt to analyse the collaboration between visualization and the humanities, and to discuss future directions for this emerging field. With it, we aim to contribute in the effort to conceive better infovis tools for the humanities, that do justice to the depth of the material, and the expertise of the researchers. With our powers combined, the future of information visualization for the humanities could be the rich, pleasant and functional landscape that its users expect.

ACKNOWLEDGMENTS

The authors would like to extend deep gratitude to Georgia Panagiotidou for her participation to the study design, to Imane Remmal for her reviewing efforts, and to the participants in the workshop and design critique exercise for their labor and time.

This research is part of Project Cornelia, a project funded by the University of Leuven (KU Leuven) and the Flemish Fund for Scientific Research-Belgium (FWO-Vlaanderen).

REFERENCES

- [1] Geocoded art - database of geotagged artwork. url-<http://www.geocodedart.com/>. Accessed: 2020-06-02.
- [2] A. J. Bradley, M. El-Assady, K. Coles, E. Alexander, M. Chen, C. Collins, S. Jänicke, and D. J. Wrisley. Visualization and the digital humanities. *IEEE computer graphics and applications*, 38(6):26–38, 2018.
- [3] R. Brath. 3d infovis is here to stay: Deal with it. In *3DVis (3DVis), 2014 IEEE VIS International Workshop on*, pp. 25–31. IEEE, 2014.
- [4] A. Budanitsky and G. Hirst. Semantic distance in wordnet: An experimental, application-oriented evaluation of five measures. In *Workshop on WordNet and other lexical resources*, vol. 2, pp. 2–2, 2001.
- [5] A. Burdick, J. Drucker, P. Lunenfeld, T. Presner, and J. Schnapp. *Digital Humanities*. MIT Press, 2012.
- [6] T. Burrows. Sharing humanities data for e-research: conceptual and technical issues. *Sustainable data from digital research: Humanities perspectives on digital scholarship*, 2011.
- [7] S. Carpendale, A. Thudt, C. Perin, and W. Willett. Subjectivity in personal storytelling with visualization. *Information Design Journal*, 23(1):48–64, 2017.
- [8] M. Correll. Ethical dimensions of visualization research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, p. 188. ACM, 2019.
- [9] C. D’Ignazio and L. F. Klein. *Data feminism*. MIT Press, 2020.
- [10] M. Dörk, S. Carpendale, and C. Williamson. The information flaneur: A fresh look at information seeking. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1215–1224, 2011.
- [11] M. Dörk, P. Feng, C. Collins, and S. Carpendale. Critical infovis: exploring the politics of visualization. In *CHI’13 Extended Abstracts on Human Factors in Computing Systems*, pp. 2189–2198. 2013.
- [12] E. Doukaridou. Reframing art history. *International Journal for Digital Art History*, (1), 2015.
- [13] J. Drucker. Is there a “digital” art history? *Visual Resources*, 29(1-2):5–13, 2013.
- [14] J. Drucker. Performative materiality and theoretical approaches to interface. *DHQ: Digital Humanities Quarterly*, 7(1), 2013.
- [15] J. Drucker. Dimensions of visualization for the humanities. 2016.
- [16] A. Dunst, R. Hartel, and J. Laubrock. The graphic narrative corpus (gnc): design, annotation, and analysis for the digital humanities. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, vol. 3, pp. 15–20. IEEE, 2017.
- [17] J. Edmond. Managing uncertainty in the humanities: Digital and analogue approaches. In *Proceedings of the Sixth International Conference on Technological Ecosystems for Enhancing Multiculturality*, pp. 840–844, 2018.
- [18] F. J. García-Peñalvo. Digital humanities data processing. *Journal of Information Technology Research*, 2016.
- [19] R. Geuss et al. *The idea of a critical theory: Habermas and the Frankfurt School*. Cambridge University Press, 1981.
- [20] F. Gibbs and T. Owens. Building better digital humanities tools. *DH Quarterly*, 6(2), 2012.
- [21] M. K. Gold. *Debates in the digital humanities*. U of Minnesota Press, 2012.
- [22] 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), 2018.
- [23] E. Graham. Introduction: Data visualisation and the humanities, 2017.
- [24] I. Herman, G. Melançon, and M. S. Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on visualization and computer graphics*, 6(1):24–43, 2000.
- [25] R. L. Hill. The political potential of numbers: data visualisation in the abortion debate. *Women, Gender & Research*, 26(1):83–96, 2017.
- [26] U. Hinrichs and S. Forlini. In defence of sandcastles: research thinking through visualization in dh. *Digital Humanities 2017*, 2017.
- [27] A. L. Hoffmann. Data violence and how bad engineering choices can damage society. *Medium*. Retrieved June, 19:2019, 2018.
- [28] J. Hullman. Why authors don’t visualize uncertainty. *IEEE transactions on visualization and computer graphics*, 26(1):130–139, 2019.
- [29] J. Hullman, E. Adar, and P. Shah. Benefitting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2213–2222, 2011.
- [30] O. Inbar, N. Tractinsky, and J. Meyer. Minimalism in information visualization: attitudes towards maximizing the data-ink ratio. In *ECCE*, vol. 7, pp. 185–188, 2007.
- [31] S. Jänicke. Valuable research for visualization and digital humanities: A balancing act. In *Workshop on Visualization for the Digital Humanities, IEEE VIS*, 2016.
- [32] S. Jänicke, G. Franzini, M. F. Cheema, and G. Scheuermann. On close and distant reading in digital humanities: A survey and future challenges. In *EuroVis (STARs)*, pp. 83–103, 2015.
- [33] L. Jégou. Expanding the sémiologie graphique for contemporary cartography, some ideas from visual semiotics, art history and design. *Cartography and Geographic Information Science*, 46(2):182–188, 2019.
- [34] M. L. Jockers. *Macroanalysis: Digital methods and literary history*. University of Illinois Press, 2013.
- [35] 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, 2019.
- [36] R. M. Kirby and M. Meyer. Visualization collaborations: What works and why. *IEEE computer graphics and applications*, 33(6):82–88, 2013.
- [37] L. C. Koh, A. Slingsby, J. Dykes, and T. S. Kam. Developing and applying a user-centered model for the design and implementation of information visualization tools. In *2011 15th International Conference on Information Visualisation*, pp. 90–95. IEEE, 2011.
- [38] M. Kyritsis, S. R. Gulliver, S. Morar, and R. Stevens. Issues and benefits of using 3d interfaces: visual and verbal tasks. In *Proceedings of the Fifth International Conference on Management of Emergent Digital EcoSystems*, pp. 241–245. ACM, 2013.
- [39] H. Lamqaddam, K. Brosens, F. Truyen, J. Beerens, I. de Prekel, and K. Verbert. When the tech kids are running too fast: Data visualisation through the lens of art history research. In *3rd Workshop on Visualization for the Digital Humanities at IEEEVIS*.
- [40] K. Latif and R. Mayer. Sky-metaphor visualisation for self-organising maps. In *Proc. of the 7th International Conference on Knowledge Management*, 2007.
- [41] S. Liu, Y. Wu, E. Wei, M. Liu, and Y. Liu. Storyflow: Tracking the evolution of stories. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2436–2445, 2013.
- [42] J. Lorber-Kasunic and K. Sweetapple. Visualising texts: a design practice approach to humanities data. In *Digital Research in the Humanities and Arts Conference*. DRHA, 2015.
- [43] G. Lupi. Data humanism: the revolutionary future of data visualization. *Print Magazine*, 30, 2017.
- [44] L. Manovich. What is visualisation? *Visual Studies*, 26(1):36–49, 2011.
- [45] L. Manovich. *Museum without walls, art history without names: visualization methods for Humanities and Media Studies*. Software Studies Initiative, 2012.
- [46] N. McCurdy, J. Lein, K. Coles, and M. Meyer. Poemage: Visualizing the sonic topology of a poem. *IEEE transactions on visualization and computer graphics*, 22(1):439–448, 2015.
- [47] M. Meyer and J. Dykes. Criteria for rigor in visualization design study. *IEEE transactions on visualization and computer graphics*, 26(1):87–97, 2019.
- [48] M. Miller, H. Schäfer, M. Kraus, M. Leman, D. Keim, and M. El-Assady. Framing visual musicology through methodology transfer. *arXiv preprint arXiv:1908.10411*, 2019.
- [49] T. Munzner. *Visualization analysis and design*. CRC press, 2014.
- [50] K. V. Nesbitt. Getting to more abstract places using the metro map metaphor. In *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on*, pp. 488–493. IEEE, 2004.
- [51] H. Pfluger, D. Thom, A. C. Schutz, D. Bohde, and T. Ertl. Vechart: Visually enhanced comparison of historic art using an automated line-based synchronization technique. *IEEE Transactions on Visualization and Computer Graphics*, 2019.
- [52] M. Posner. Miriam posner’s blog. <http://miriamposner.com/blog/humanities-data-a-necessary-contradiction/>, Jun 2015.
- [53] G. Resch. *Denaturalizing Information Visualization*. PhD thesis, University of Toronto (Canada), 2019.
- [54] A. Sallaberry, Y.-c. Fu, H.-C. Ho, and K.-L. Ma. Contact trees: Network visualization beyond nodes and edges. *PLoS one*, 11(1):e0146368, 2016.
- [55] V. Schetinger, K. Raminger, V. Filipov, N. Soursos, S. Zapke, and S. Miksch. Bridging the gap between visual analytics and digital humanities: Beyond the data-users-tasks design triangle.
- [56] C. Schöch. Big? Smart? Clean? Messy? Data in the Humanities. *Journal of Digital Humanities*, 2(3):2–13, Dec. 2013.
- [57] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visual-*

- ization and computer graphics, 18(12):2431–2440, 2012.
- [58] S. Simon, S. Mittelstädt, D. A. Keim, and M. Sedlmair. Bridging the gap of domain and visualization experts with a liaison. In *Accepted at the Eurographics Conference on Visualization (EuroVis 2015, Short Paper)*, vol. 2015. The Eurographics Association, 2015.
- [59] C. D. Stolper, A. Perer, and D. Gotz. Progressive visual analytics: User-driven visual exploration of in-progress analytics. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1653–1662, 2014.
- [60] Tategallery. [tategallery/collection](http://tategallery.com/collection), Feb 2018.
- [61] K. Thoden, J. Stiller, N. Bulatovic, H.-L. Meiners, and N. Boukhelifa. User-centered design practices in digital humanities—experiences from dariah and cendari. *ABI Technik*, 37(1):2–11, 2017.
- [62] L. Thorlacius. Aesthetics and function in web design. In *CADE2004 Web Proceedings of Computers in Art and Design Education Conference. Conference proceedings: <http://asp.cbs.dk/cade>*, 2004.
- [63] E. R. Tufte. *The visual display of quantitative information*, vol. 2. Graphics press Cheshire, CT, 2001.
- [64] G. Uboldi and G. Caviglia. Information visualizations and interfaces in the humanities. In *New Challenges for Data Design*, pp. 207–218. Springer, 2015.
- [65] P. Valdivia, P. Buono, C. Plaisant, N. Dufournaud, and J.-D. Fekete. Analyzing dynamic hypergraphs with parallel aggregated ordered hypergraph visualization. *IEEE Transactions on Visualization and Computer Graphics*, 2019.
- [66] J. M. van der Zwaan, M. van Meersbergen, A. Fokkens, S. Ter Braake, I. Leemans, E. Kuijpers, P. Vossen, and I. Maks. Storyteller: Visualizing perspectives in digital humanities projects. In *International Workshop on Computational History and Data-Driven Humanities*, pp. 78–90. Springer, 2016.
- [67] J. Van Dijck. Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology. *Surveillance & society*, 12(2):197–208, 2014.
- [68] P. van Wierst, S. Hofstede, Y. Oortwijn, T. Castermans, R. Koopman, S. Wang, M. A. Westenberg, and A. Betti. Bolvis: visualization for text-based research in philosophy. In *3rd Workshop on Visualization for the Digital Humanities*, 2018.
- [69] J. J. Van Wijk. Bridging the gaps. *IEEE Computer Graphics and Applications*, 26(6):6–9, 2006.
- [70] A. Vande Moere, M. Tomitsch, C. Wimmer, B. Christoph, and T. Grechenig. Evaluating the effect of style in information visualization. *IEEE transactions on visualization and computer graphics*, 18(12):2739–2748, 2012.
- [71] C. Warwick. Studying users in digital humanities. *Digital humanities in practice*, 17(5):1–21, 2012.
- [72] M. Whitelaw. Generous interfaces for digital cultural collections. *Digital Humanities Quarterly*, 9, 2015.
- [73] F. Windhager, P. Federico, G. Schreder, K. Glinka, M. Dörk, 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, 2018.