

# In search of an interpretative environment for digital traces: the building of Arvest


**Clarisse  
Bardiot** 

Université  
Rennes 2,  
Rennes, France

**Jacob Hart** 

Université  
Rennes 2,  
Rennes, France

**David Rouquet**

  
Tetras Libre, Saint  
Martin d'hères,  
France

<https://doi.org/10.1515/JDH-2024-0013?locatt=label:JDHFULL>

Published online the 5th May 2025

Experience the full interactive  
multi-layered reading experience

[-> Read here](#)

Send yourself a reminder  
to read this article later

[-> Send E-Mail](#)

**Keywords:** Multimodal, Annotation, Traces, Historiography, Development

**T**he rise of born-digital materials and the digitization of existing archives is fundamentally changing the nature of sources that humanities researchers work with. As sources transform to multimodal digital traces from which we look to extract data it becomes necessary to also reflect upon the digital environments within which we work. Building upon a decade's work of digital humanities research and software development, we present Arvest: a tool for

interpretative manipulation of digital traces based on IIF (International Image Interoperability Framework). After a comprehensive state of the art and establishing the theoretical framework that drives development, we shall present the technical specifications of Arvest, and how we intend it to simultaneously propose novel solutions to issues pertaining to working with digital traces and seamlessly integrate with existing tools. We shall also present a case

study of research based on a multimodal corpus from the Horizon2020 COESO Project which

drove Arvest's initial development and illustrates some of its core functionalities.

1



© Clarisse Bardiot - Jacob Hart - David Rouquet. Published by De Gruyter in cooperation with the University of Luxembourg Centre for Contemporary and Digital History. This is an Open Access article distributed under the terms of the [Creative Commons Attribution License CC-BY](#)

## 2 Introduction

3 In the humanities, driven by the rise of born-digital materials and the digitization of existing archives, traces are becoming digital -making them a new type of primary source for historians. Digital technology fundamentally changes their ontology: becoming digital, traces have become fragmented, reduced to units of information, data. Digitizing a trace means discretization, making it suitable for calculation and manipulation. That is the fundamental distinction between the analogue and the digital trace. Data from traces are no longer traces. They are “*capta*” (Drucker 2011). Historiography in a digital context is not a revival of quantitative history in the strict sense of the term. It calls first and foremost for practices, and in this practice, the narrative is framed. It is a question of practice that goes beyond simply exposing traces: we expose traces to manipulation, to modelling (McCarty 2019) (Flanders, Jannidis 2019), and theories and concepts to operationalization (Moretti 2013). Hence, a paradigm shift is taking place, inviting us to make history differently. Because traces are transformed, they invite a renewal of the critique of the sources we thought we knew and make research questions evolve, leading to other epistemologies. The digital turn revitalizes source criticism, methodological thinking, and narratives, moving beyond traditional methods of close reading to computationally assisted reading and contributing to the general reflection on digital history.

4 In this context, the digital traces of contemporary events are characterized by their number, multimodality, and fragility. As early as 2003, Roy Rosenzweig warned us: “*Historians, in fact, may be facing a fundamental paradigm shift from a culture of scarcity to a culture of abundance.*” (Rosenzweig 2003). The unprecedented increase in digital artefacts and data presents specific difficulties highlighted by several actors in digital history (Manning 2013, Graham, Milligan, Weingart 2016, Rygiel 2017, Milligan 2022). Consider for example the digital archives of a performing arts company, which could contain thousands of files from a single production. These files, varying from images and videos to text and software, encapsulate elements of the performance, its creation, and its reception. The challenges in distinguishing between these different types of data highlight the complexity of working with digital archives. While some traces are easily associated with public reception or final performances, others, such as a rehearsal audio recording, defy simple categorization (is it merely a testimony of a rehearsal, a technical test, or a file used in performance?) and often require direct consultation with the creators for accurate interpretation. Despite the voluminous data available, and while they are of a greater variety and concern phenomena that were previously ‘untraceable’, many aspects remain under-documented. Not all traces can be collected from all collaborators. Certain stages of collaboration leave behind scant traces, such as a few photographs from a workshop or sporadic notes from offstage discussions. Certain parts of digital activity itself rarely result in preserved traces: those occurring on social networks, or during conversations on videoconferencing platforms or applications such as Slack. Lastly, digital traces are fragile: it is estimated today that the

average lifespan and readability of a digital file is five years. In a few years, there is no guarantee that we will still be able to open and read certain files, especially those made in proprietary software. The loss of traces is a real and inevitable phenomenon: in the past few decades, the number of artists who have already lost not only traces but works due to digital erosion is far too high (see for example in media art (*Noordegraaf, Saba, Le Maître, Hediger 2013*) and in music (*Bachimont, Blanchette, Gerzso, Swetland, Lescurieux, Morizet-Mahoudeaux, Donin, Teasley 2003*)).

- 5 Despite the complex characteristics of digital traces, they nonetheless offer the promise of renewed studies. To continue in the context of performing arts studies, digital traces are a paradigm shift not only for the history of theatre but also for creative process and aesthetic analysis. To fulfil these promises, we need to build specific tools and environments in order to handle and interpret digital traces in all of their subtleties and complexity. The development of such an environment reveals five primary challenges that must be overcome: to move away from a 'silo' approach and allow for the analysis of traces in a multimodal environment; to allow for both manual and automatic annotation of traces, including complex documents like video recordings; to allow for navigation between close reading and distant reading perspectives; to avoid losing context by linking data extraction and the original trace; to reintroduce narrative into visualizations in order to reconcile temporal and spatial representations.
- 6 These concerns have guided the development of Arvest, a multimodal environment for linking data and their context, annotating digital traces, and deriving narrative from data. While the starting point for this research and software is the analysis of corpora documenting the creation processes of performances, Arvest is aimed at many disciplines in the humanities and social sciences, especially historians. The goal of this tool is indeed to reconcile traditional hermeneutic approaches with methods from digital history. The application is still under development: as of the writing of this article, a [first full release](#) is available since March 2025. In this paper, we will first revisit the various lab-objects and co-design stages that led to its design. In a second part, we will detail the main epistemological and hermeneutic challenges as well as the technical choices and functionalities implemented. Finally, we will present a case study conducted with Arvest before concluding on future prospects within the framework of the [ERC-funded STAGE project](#).

## 7 **Developing lab-objects**

- 8 One of the major characteristics of digital traces is their experiential nature. To overcome the specific difficulties of digital traces, it is necessary to experiment with and on digital traces. Without action, some elements would not be observable, some objects would not be readable. That is why we rely not only on the methodologies and tools of digital humanities but also on action research by developing our own instruments. Arvest is the continuation of two previous projects, *Rekall* and [MemoRekall](#), both spaces for questioning themes, methodological tools, and environments in which we create lab-objects. To use Bruno Bachimont's terms, the lab-object allows us to "*avoid having to choose between an object restricted enough to allow its study and broad enough to observe the effects of its complexity. An artistic production taken as a lab-object presents a locality that neither dilutes nor dissolves its complexity, which remains observable for comprehensive study.*" The lab-object also allows for multidisciplinary work, and in the case of *Rekall* and *MemoRekall*, a co-design approach as well as an opening towards citizen science.

## 9 **Rekall**

10 As mentioned above, digital traces of works are numerous, and reading each trace is nearly impossible. In the context of performing arts studies, the question thus becomes: what information, what knowledge about the performances is it possible to extract from these traces? Which data should be prioritized to reconstruct a creative process? Contrary to the practice of creating a specific interface for each show, Clarisse Bardiot has led a reflection on the design of software that can be applied to as many works as possible (*Bardiot 2021*, *Bardiot 2020*). She undertook this project in 2007 as part of the [DOCAM program](#) on "Documentation and Conservation of Media Arts" led by the Daniel Langlois Foundation in Montreal, and then worked in collaboration with [Buzzing Light](#) and Thierry Coduys for the development and design of a prototype for Rekall, starting in 2012.

11 Two main objectives guided the design of Rekall: to help artists document their works so they can be reproduced and thus counteract the obsolescence of digital technologies; to assist researchers in studying the genetics of works in a context of born digital documents and big data. Yet, it is the same materials (traces of the creative process and the works) that are collected, analysed, and visualized, according to different modalities and timelines. For artists, collection takes place during the creative process itself, progressively through the work stages; it is a living archive, in perpetual motion. The researcher intervenes once the process is completed, once the documents have somewhat stabilized and frozen, and they become traces of what took place. The goal is not to freeze the traces but, on the contrary, to foster a living documentation, which can be augmented, refined (based on contributions from artists, their teams, or researchers and curators) and also interpreted differently depending on research questions that are translated into visualization modes. The accumulation of traces is a key element for the effective operation of Rekall. Indeed, by analysing these traces and correlating them with each other, Rekall gradually reveals characteristics specific to a given work.

12 Usual practices of corpus processing separate documents by file types and work in sealed 'silos' (texts/images/sounds). Such treatment prevents us from obtaining an overall vision of the collected traces and examining many phenomena, such as the evolution of an idea through different traces, from an image to a text, from a text to a spreadsheet, from a spreadsheet to a video recording. To address this essential issue, Rekall was designed from the start (*Bardiot 2020*) as a multimodal environment that allows an infinite number of traces of different natures to be aggregated in the same interface. Rekall brings together texts, images, sounds, videos, programs, etc. in the same space. Trace analysis is based on the metadata present in each file that allows the extraction of crucial information (author, date of creation, place of creation, keyword, etc.). This information is then used by various data analysis and visualization algorithms to reveal behaviours and practices. By playing with axes and different filters that offer many different perspectives on a given creation, Rekall's functionalities allow for graphical analysis of the work through its traces. These different views allow navigation between the micro and the macro, between close reading and distant reading, between diachronic representation and synchronic representation. In other words, being able to represent a process without erasing its complexity. For a detailed description of the prototype and its interface, we refer to (*Bardiot 2021*).

## 13 **MemoRekall**

14 Initially, Rekall was developed to manage documents within a multimodal framework, enabling the interconnection of various documents. Among these documents, video already

held particular importance, with a very first attempt to build an interface in order to add video annotations and to allow comparisons. Video has become an indispensable tool for a variety of research in the humanities. For example, in the field of intangible cultural heritage, videos of cultural practices enable a meticulous study of gestures, expressions, and social interactions (Hou, Kenderdine, Picca, Egloff, Adamou 2022). In sociology, recordings of focus groups or everyday life situations provide rich data on behaviours and social dynamics (Krueger, Casey 2015). Similarly, in linguistics, videos of verbal exchanges are essential for studying the nuances of non-verbal and paraverbal communication, such as body language, and facial expressions. In performing arts history, video recordings help to better understand the creative process (Waterhouse, Jenett, Hager, Coniglio 2021) and build aesthetic analysis of performances (Aronson-Lehavi, Skop, Via Dorembus 2021). Videos indeed provide invaluable data, however they often need to be accompanied by expert interpretation and contextual information to be fully effective and explicit in conveying their intended insights. While video traces have spread across numerous fields and are used for a variety of purposes, they are rarely self-sufficient: to be explicit, they require commentary.

- 15 Taking the importance of video documentation into account, the video annotation functionality in Rekall was spun-off into MemoRekall: a separate, user-friendly web application designed for annotating videos and linking them to other documents or web pages. These annotations and links form a "capsule," a new type of document that can be embedded in websites. Two complementary annotation strategies are allowed: inter-documentary (linking the video to a broader set of documents) and intra-documentary (annotating within the video itself). In the process of making the audio-visual document explicit by adding annotations and links to other documents, the video is "redocumentarised" (Zacklad 2007). In this way, MemoRekall makes it possible to reappropriate documentary corpora. From a single video, different documentary strategies can be developed, depending on the specific challenges of each capsule author: in the case of performing arts studies, artists need to document their works so that they can be easily distributed, adapted or reenacted; cultural institutions wish to share creative processes with their audiences, as well as elements of analysis on their websites; teachers need to present their students with different multimedia resources to accompany their courses on the history of the performing arts, or as part of educational activities linked to performances seen in class (please see an example capsule below, and for others visit [www.memorekall.com](http://www.memorekall.com)).

## 17 **A co-design approach grounded in citizen science**

- 18 Such projects as Rekall and MemoRekall often fall somewhere on a spectrum between two approaches: on the one hand, the development of tools within the context of a specific research project that look to interrogate a specific corpus and that are designed to investigate specific epistemological questions; on the other, the development of more generic working environments which have a certain user profile in mind but do not base themselves around a specific use-case during development. In the former case, the tools that emerge can be characteristically project-specific, and can only be further applied in a restrained number of contexts; in the latter, tools runs the risk of allowing for functionalities and workflows that can appear useful in concept, but which crumble when applied to actual case-studies. This polarity illustrates the tensions that can arise when researchers and programmers meet to work in collaboration on a project.

- 19 This spectrum is a caricature – projects will inevitably find themselves somewhere along this gauge without reaching either of its extremities – but it is a useful idea to bear in mind when planning and executing projects that see academic, computer science and citizen profiles coming together to collaborate and create “*interdisciplinary trading zones*” (Oiva, Pawlicka-Deger 2020). There are numerous examples in recent years where this balance would appear to have been achieved: for example the ERC-funded FluCoMa project (Tremblay, Green, Roma, Bradbury, Moore, Hart, Harker 2022) led by Pierre Alexandre Tremblay at the University of Huddersfield saw computer scientists working closely with creative coders to bring a toolkit of machine learning algorithms and sound corpus manipulation tools to a number of creative coding environments; another is Taylor Arnold and Lauren Tilton’s Distant Viewing Lab (Arnold, Tilton 2023) at the University of Richmond which produced a python package (Arnold, Tilton 2019) for distant viewing in the digital humanities across a number of research projects with the intention of making their tool accessible to non techno-fluent researchers through a high-level command line interface.
- 20 Following a design thinking methodology (Cross 2011), Rekall and MemoRekall have emerged from a development process centred on collaboration and co-design. This approach aligns with the principles of citizen science, a practice where scientific research is conducted, in whole or in part, by amateur or nonprofessional scientists. Workshops have been carried-out in different contexts that were structured specifically to engage with users’ needs following the UK Design Council’s Double Diamond Framework (Groeger, Schweitzer 2014): workshops were divided into exploration and development phases, where users first discover the tool and define their needs, then these needs are developed and delivered back to the users at which point the process can start again. Working in this way allows us to understand how the tools we have developed are understood by the users, meet interface and functionality needs that may have been opaque to the developers, and ground the tools in real use.
- 21 By involving a diverse group of stakeholders, including non-experts, in the research process, Rekall and MemoRekall not only benefit from a wide range of perspectives and expertise but also foster a sense of ownership and engagement among participants. For example, over the years, MemoRekall has been developed in the context of various citizen-science research projects, such as the Erasmus+ funded COSMIC project (Bardiot, Menassel 2022) and the Horizon2020 funded COESO project (Bardiot, Hildebrand, Ferrando, Graffione, Blin, Marranca 2022). In each instance, user feedback (professional circus teachers, choreographers, students from high school to MA) was meticulously considered and incorporated into the tool, addressing aspects ranging from interface design and user experience to specific functionalities and workflows. In the context of the COSMIC project, the application was tailored to the needs of circus education. This adaptation process involved close collaboration with circus educators and practitioners to identify and integrate features that support the complex and dynamic nature of circus training. Specific functionalities (slow motion, visual annotation) were developed to facilitate the detailed analysis of physical movements and the documentation of training routines. This collaborative work has yielded its fruits, and MemoRekall is now widely used by researchers, artists, cultural professionals, teachers and schoolchildren.
- 22 **From MemoRekall to Arvest**
- 23 In the context of the COESO project, the co-design approach led us to reconsider the video centric approach of MemoRekall and more broadly its development in the context of FAIR principles. During the work sessions on documenting the collaboration between a

choreographer and dancer (Cosetta Graffione) and a philosopher (Stefania Ferrando), we concluded with Irénée Blin (Laban notator on the project) that video could no longer be the central document for annotation work. Instead, we needed to consider a network of multimodal documents that could annotate each other, as well as a means of changing perspective by opening documents that can be used as new entry points into the annotation network. Based on this observation, a researcher (Clarisse Bardiot), a team of developers (Tetras Libre), a Laban notator (Irénée Blin) and a post-doc (Jacob Hart) worked together to improve the MemoRekall application and provide a prototype for a new version of the software: the development of a multimodal annotation editor to link a video recording to its Laban notation and vice-versa.

- 24 This new conceptual approach and these new functionalities required a complete reimplementa-tion of the tool. In order to base this reimplementa-tion on open standards, the [W3C Web Annotation Guidelines](#) provided an extensible data model and protocol for multimodal annotation across the World Wide Web. In addition, these W3C recommendations are endorsed by the [IIIF \(International Image Interoperability Framework\)](#) initiative, which is receiving increasing attention from museums and other cultural institutions distributing digital content. For example, [Europeana](#), which is a data source for the annotations initially targeted in the COESO project, uses IIIF to distribute its content.
- 25 Therefore, in order to meet COESO's objectives, we decided to implement Memorekall-IIIF, a prototype for multimodal annotation that capitalizes on Memorekall's experience and communities, but is also based on IIIF's open standards. The tool is based on the IIIF [Mirador viewer](#) and can be extended via themes and plugins to suit specific community needs. The prototype, a COESO-based example and a sandbox are presented here: <https://arvest.app/en/community/projects/coeso>. This prototype was the basis for what has now become Arvest. Its functionalities and the choice of the IIIF are described in the second part of this paper.

## 26 **Evaluating Video Annotation Tools: A Comparative Study**

- 27 One of the first objectives of Arvest is the building of an effective video annotation tool for digital humanists based on the IIIF protocol and MemoRekall functionalities. In our study, we analysed 51 video annotation tools focusing on multiple aspects such as supported storage services, annotation types, cooperation features, sharing capabilities, interoperability, extensibility, usage mode, business model, and governance.
- 30 Most of the video annotation tools analysed support YouTube as the primary storage service. Tools like MemoRekall and Hihaho also support other services such as Vimeo. Additionally, some software like Advene and Omeka allow for local storage, providing extra flexibility for users who prefer to work offline.
- 31 Textual annotations are the most common among the tools studied. However, some software goes beyond by offering advanced annotations such as inserting images, URLs, and interactive elements. For instance, Timeliness allows timed annotations linked to images, Google maps, text, and URLs. Edpuzzle stands out by enabling the addition of quizzes, text, and voice recordings, making it a comprehensive educational tool for creating interactive videos.
- 32 Collaboration is an important feature for many tools. MemoRekall and VideoAnt enable multiple users to work on the same project, facilitating team collaboration and co-creation

of annotated content. However, other tools do not offer specific cooperation features, which can limit their use in collaborative contexts.

- 33 Most tools offer sharing options via direct links or HTML embedding. VideoAnt, for example, allows projects to be shared via links, embedded in HTML pages, and exported in various formats. YiNote is notable for its export capabilities to multiple formats such as JSON, PDF, and markdown, making it easier to share and integrate annotations across different platforms.
- 34 Interoperability is an area where many tools have limitations. A few tools, such as YiNote and Mindstamp, allow export to standardized formats or integration with other platforms. Extensibility through plugins or APIs is not widely available but is present in some tools like Verse and Hihaho, offering users the ability to customize and enrich the core functionalities according to their specific needs.
- 35 The tools also vary in terms of usage mode. Some are strictly online, such as FrameTrail and Loom, while others, like autoEdit and Entity Mapper, can be used locally.
- 36 The majority of tools are free to use, although some require account creation. Others offer premium features or are entirely paid services, like Mindstamp and Dspace.
- 37 The tools are developed by a variety of entities, ranging from academic institutions, such as MemoRekall (Université Rennes 2) and VideoAnt (University of Minnesota), to professional development teams and open-source communities. This diversity in governance often influences development priorities, support quality, and the sustainability of the tools.

## 38 **Arvest, an environment to interpret digital traces**

- 39 The survey on video annotation tools, although not exhaustive, reveals that no existing solutions fully meet all the requirements necessary for humanities research. On a practical level, these requirements include: storage services such as YouTube, Vimeo and PeerTube, as well as public and private video facilities; a wide range of annotation possibilities (aside the document or as overlays, textual and visual annotations, and the ability to link to web pages and documents); collaboration features; sharing options on the web and import/export in various formats (CSV, JSON, XML); interoperability and, more generally, adherence to FAIR principles; extensibility through plugins or APIs.
- 40 On a more general level, the passage from MemoRekall to Arvest shows how difficult digital traces are to interpret and thus the necessity of an *“interpretative interface”*, which Johanna Drucker describes as supporting *“acts of interpretation rather than simply returning selected results from a pre-existing data set.”* It *“will orchestrate [...] the shift from conceptions of interface as things and entities to that of an event-space of interpretative activity”*. She concludes: *“More attention to acts of producing and less emphasis on product, the creation of an interface that is meant to expose and support the activity of interpretation, rather than to display finished forms, would be a good starting place”* (Drucker 2013).
- 41 In the journey of building such an interpretative interface, we moved in the context of the COESO project (see below) from video annotation to the interpretation of digital traces. The video-centric approach of MemoRekall is still possible, however we also offer a fundamental shift in perspective – that of a network of multimodal documents, where any artefact of any

kind may become the focus of one's project. The main research questions leading the development of this new version, named Arvest, are:

- How can software actively support and enrich a semantic line of interpretation combining qualitative and quantitative approaches around a collection of multimodal documents?
- How can software create research artefacts that participate in the researcher's analysis?
- How can software allow for the creation of "interpretative interfaces" that facilitate research and actively participate in analysis?
- How can we link multimodal documents and data extracted from them in the same environment in order to keep the context of the data and to go back and forth between close and distant readings?

42 The reduction of digital traces to data is not neutral: it is fragmentation, reduction, even decontextualization. In this context, it is essential, as Bruno Latour's text on the construction of scientific reference (*Latour 2006*) shows, to document all stages that lead from the work and its creative process to the trace, from the trace to the data, from the data to its visualization and interpretation; and to understand the implications of each stage in this chain in terms of data interpretation and trace reduction. In other words, to be able to go back to the source, to identify the various manipulations and instrumentations that led to a given visualization or interpretation. Each step leads to an interpretation, which must be documented as such.

#### 43 **Development theoretical framework**

44 In order to build a framework for development, we identified five main structuring challenges to overcome which we detail in the following lines.

#### 45 **Analyzing traces in a multimodal environment and moving away from a silo approach**

46 Taking into account the multimodal features of the documents we are confronted with as historians is considered an important turn within the general context of digital humanities (*Bateman 2008*), and more specifically in the context of computational humanities research (*Smits, Wevers 2023*). If "multimodal" seems to have replaced the term "multimedia" in recent times, it's not merely to update the vocabulary. The shift emphasizes the simultaneous connections and interactions between different media—such as text, images, sound, and video—rather than their simple juxtaposition. According to the Oxford English Dictionary, "multimedia" pertains to creative works that utilize or combine various forms of digital content, such as text, audio, video, and animation - the emphasis here is placed on the medium that supports the content. Conversely, "multimodal" refers to the incorporation or utilization of several different methods or systems to convey a message, integrating multiple semiotic modes of communication—such as text, images, motion, and audio—within a single artefact. Here we are dealing with a mode, a form, rather than a supporting medium. "Multimodal" underscores the inclusion of different modes of communication that cannot be superimposed with a media, such as movement or gesture in videos. As Bateman elucidates, "nowadays that text is just one strand in a complex presentational form that seamlessly incorporates visual aspects 'around', and sometimes even instead of, the text itself." This underscores that multimodal projects go beyond merely juxtaposing different media formats; they involve the orchestrated interaction of these modes to create a cohesive presentation. Bateman further emphasizes this by stating, "Combining these modes within a single artefact—in the case of print, by binding, stapling, or folding or, for online media, by

*'linking' with varieties of hyperlinks—brings our main object of study to life: the multimodal document."*

47 In the context of digital humanities, recognizing the multimodal nature of documents is crucial. Traditional siloed approaches often separate documents by file types—text, image, etc.—which limits a comprehensive analysis of the interconnectedness of different elements, especially in image and video analysis. The multimodal approach aims to integrate these different modes into a cohesive whole, highlighting how they work together to convey meaning. This broader perspective allows for a more comprehensive and dynamic understanding of how various semantic elements interact within a single artefact.

#### 48 **Manually and automatically annotating traces**

49 Whether through annotating video recordings or enriching textual data, digital technologies enable a dynamic interaction with semiotic content. Annotation, as defined by Manuel Zacklad, refers to "*any form of addition intended to enrich an inscription or recording in order to attract the receiver's attention to a passage or to complete the semiotic content by relating it to other pre-existing semiotic content or by an original contribution*" (Zacklad 2007). One of its main goals is to make the implicit explicit. Annotation is most of the time considered a private practice, for example in the first steps of a research process to mark important passages of a document, to highlight specific words or visual elements, or to track and grasp an emerging analysis. It can also be part of a research group practice, to collectively annotate a collection of documents according to specific needs and rules, as we see for example with crowdsourcing. But these annotations can also move to the public sphere: for example if we consider footnotes as an annotation practice, or if we decide to publish and edit private annotations. In computer science, openly publishing the annotations which led to the training of a machine learning algorithm is also a practice encouraged by open science.

50 Indeed, annotation practices have evolved significantly with the advent of digital technologies, expanding the possibilities of annotation beyond traditional textual forms to include for example graphic and audio-visual annotations. Digital annotation practices extend also to the concept of metadata, which play a crucial role in describing the content and characteristics of digital files. Metadata is a useful tool in many contexts, however it does present some limitations in regard to the challenge of adequately capturing the semantic richness contained within files. To address this, more advanced tools such as image recognition, topic modeling, and lexicometry are being explored in the digital humanities to extract more precise information from visual and textual data. Hence, to the private/public dichotomy of annotations, we also can add a manual/automatic one.

51 Building on the afore-discussed experience with MemoRekall, being able to move from private to public annotation, as well as integrating manual and automatic annotations is one of the challenges we wish to address with Arvest.

#### 52 **Navigating between close reading and distant reading**

53 Many research projects involve corpora with a small number of documents that need to be annotated manually. That's why we feel it's important to offer an interface that can be used to annotate small corpora, as well as larger corpora requiring automatic annotation or a combination of manual and automatic approaches.

54 The integration of annotation tools offers a more nuanced understanding of digital content, bridging the gap between close reading and distant reading. Close reading, the traditional method used by humanities researchers, involves meticulously deciphering the content of a document. In contrast, distant reading, theorized by Franco Moretti, involves analysing large literary corpora by observing relationships between texts rather than reading each one closely. Distant reading allows researchers to move "*from texts to models*", identifying patterns and relationships across a vast number of documents (Moretti 2005). Similarly, distant viewing, as defined by Taylor Arnold and Lauren Tilton (Arnold, Tilton 2019), extends this concept to corpora of still or moving images.

55 We wish to create a continuum that enhances our interpretative capabilities and supports the development of more sophisticated data visualization interfaces. This multi-scale approach allows researchers to toggle between detailed analysis and broader pattern recognition, ultimately fostering a deeper and more comprehensive understanding of digital content.

56 **Avoiding loss of context by linking data extraction and the original trace**

57 The issue of the link between data and its original context — its source — is today a major concern. To really grasp the scale of this problem and its cultural embeddedness, one may consider for example the marked absence of any cited source in the results of queries made with digital tools such as chatGPT. In history, the issue is crucial and at the very heart of the historian's practice: the practice of footnotes is there to link the stated fact and its source, preserved and accessible in a specific archive collection. "*Citing one's sources*" and "*going back to the source*" is the creed taught from generation to generation (Prost 2014).

58 In a digital context, this usage translates into the description of the methodology, the process that goes from sources to data and then to their visualization and interpretation. Such "*good practices*" are at the heart of FAIR recommendations and the reproducibility of research. Thus, today, in conferences and publications dedicated to digital humanities, many presentations and articles include a presentation of the workflow that led from sources to their interpretation (Carboni, Barras, Joyeux-Prunel 2023). The development of data papers follows the same direction. Yet, it is often difficult, for example at the time of data visualization, to go back to the source with a single click, as one might do with a footnote. It can be assumed that this problem will only grow with the availability of "data sets" that are more or less ready to use but cut off from their original sources. The problem is far from being ignored and there have been multiple initiatives by developers of document exploration interfaces and digital publications to reconcile the data and the source from which it is extracted. One notable example is the Impresso project, Media Monitoring of the Past, which allows moving from different visualizations of data related to subjects, people, or places mentioned in the texts and images of a multilingual newspaper corpus, to viewing and reading the source from which these data were extracted (Düring, Bunout, Guido 2024).

59 However, it must be acknowledged that a gap is widening at the moment of the transition from unstructured data (the sources) to structured data. This transition, the operation of extracting data from documents, can take place manually (by entering information into a spreadsheet) or computationally (for example, with machine learning or deep learning processes). By transforming the trace into data, the researcher performs a triple operation of simplification, reduction, and construction of the information available in the digital trace: simplification, because not all the complexity of a digital trace can be taken into account; reduction, because only a limited number of variables are considered; construction, because the preceding operations are deliberate actions carried out in function of specific research

questions and objectives. To be able to perceive the connections between traces through their fragments, it is necessary to produce an image (whether visual or auditory) that presents a synthetic vision of the data and their relationships. Taken separately, the data are devoid of meaning. They are mere noise, a confusing mess from which one must grasp "the overall sense of an obscured reality" (Ginzburg 1989). By accumulating them within an image, by recreating continuity between discontinuous elements, by replacing isolated elements within a whole, it is then possible to make them intelligible. If digital traces need to be instrumented to be read and analysed, the question remains how, and at what cost? Data consists of a minimal unit of information that is both simpler to preserve and to process in comparison with its unstructured trace. But this is done at the cost of cutting the traces into an infinity of fragments. The risk is the loss of the overall view of the trace and the context from which the data was extracted.

- 60 The hiatus between the trace and the data, the break in causality caused by capturing and coding, necessitates the creation of multimodal environments for research, particularly in history, which allow for reconnecting and restoring continuity between the data and the trace from which it was extracted, as easily and directly as using a footnote, or even more effectively. While footnotes often remain theoretical due to the logistical challenges that physical archive visits pose for the reader, our goal is to make this potential accessible by offering the possibility to directly consult the referenced sources. In this perspective, "*Data in Context*" could become the guiding principle of digital historiography, the digital equivalent of the traditional "*cite your sources*".
- 61 **Reintroducing narrative into visualizations in order to reconcile temporal and spatial representations.**
- 62 The integration of narrative into visualizations plays a pivotal role in reconciling temporal and spatial representations. Narrative, encompassing both time and space, transforms how we perceive and interact with data, allowing us to weave together disparate elements into a coherent whole. As Alberto Cairo (*Cairo 2016*) suggests, visualizations can be utilized at various stages of the research process: during exploration to derive new insights, in discovery and interpretation to uncover hidden patterns, and as an end product for disseminating research findings. This multipurpose use underscores the necessity of incorporating narrative to link complex documents, which exist not in isolation but as interconnected pieces of a larger collection.
- 63 The paradox of aggregation lies in its inherent fragmentation. While aggregation strives for completeness, it merely captures a fragment of the vast information available on the Web or within large databases. The real challenge lies within the need to increase content but to also simultaneously maintain a human scale, allowing the collection to be perceptible as a whole, in each of its elements, and in the connections between them. The interplay of the whole and the part, and the part and the whole, necessitates the reconstruction of a narrative that links these fragments.
- 64 In digital historiography, two types of narratives emerge: the narrative of the process steps and the narrative of the process results. The connotation step implied by the analysis of a data visualization serves as a narrative, reintroducing time into what primarily appears as static spatial representations. Contemporary interfaces rarely support dynamic representation, presenting us instead with computer-generated images that offer an overview of calculated results. Connotation becomes an unfolding operation, akin to carefully spreading out a crumpled piece of fabric. The narrative reintroduces hermeneutics and esthesis, countering the desire for a purely mathematical representation.

65 Narrative, therefore, is not merely a discourse of proof. Despite the allure of measurement in certain projects, calculation alone lacks objective criteria. Meaning emerges through successive interpretations and instrumentations, necessitating a narrative to explain the process and propose a reading of its results. The paradigm remains one of interpretation, not explanation. By reintroducing narrative into visualizations, we bridge the gap between temporal and spatial representations, fostering a deeper understanding and appreciation of the data we seek to analyse and interpret. This narrative-driven approach ensures that visualizations are not static endpoints but dynamic tools for continuous discovery and storytelling in the digital age.

## 66 Specs and prototype

67 We shall now give a brief overview of the specifications and functionalities of Arvest as a tool. As mentioned above, we made the choice to make Arvest fully compatible with IIIF. Even if this aspect will not necessarily be exposed to the average user, it is an important part of the tool and as such we shall now quickly go over some of the main advantages of IIIF and how they propulse Arvest as a platform. Technically, IIIF is comprised of two main parts: a set of open APIs that each have their own functionalities (an Image API for serving high resolution images; a Presentation API for serving media along with other elements like metadata and annotations; an Authorization Flow API for handling access control; a Content Search API for searching through textual annotations; and a Content State API for serving parts of a IIIF Manifest in a compact format); and a number of open source and proprietary IIIF-compatible viewers that can serve media. Let us briefly focus on the [Presentation API](#), which is currently in its third version. This part of the specification allows for the creation of a IIIF *Manifest*, which is a JSON format document which, using the W3C Web Annotations model, refers to a media resource object (or range of objects) as well as other elements such as metadata, annotations, required statements etc. The IIIF Manifest is a versatile tool that has several advantages:

- It is *lightweight*. As a textual JSON document, stocking and exchanging IIIF Manifests requires very little resources.
- It is a *referential* document. Each element in a Manifest can point to a URL which means that it is possible to remove the need for duplicate uploads of data and media. This is also useful in an academic context, where parts of the documents or the entire document can be precisely cited and perpetuated with DOIs.
- It is *interoperable*. All IIIF viewers must be able to interpret a IIIF Manifest, ensuring that content within the Manifest will always be accessible regardless of the technology used to open it.
- It is a *composite* document. With a IIIF Manifest, it is entirely possible to encapsulate within one lightweight, interoperable file all of the complementary elements like metadata and annotations we have discussed in this article.

68 To these advantages, we can also add the fact that hundreds of collections from various institutions around the world have already started making their content IIIF-compliant — meaning that users in a IIIF-driven environment already have access to millions of different digital objects from around the world. The majority of these objects are images (as the name suggests, IIIF was initially conceived to serve images), however the third version of the Presentation API now allows time-based media to be included. Many of the IIIF viewers however have not yet incorporated this functionality. This is where the work we are doing will also allow us to contribute to IIIF's community of open source developers — our time-

based media serving and annotation tools will be fully available on their own for the community to use as they wish independent of the Arvest environment.

69 As such, Arvest started life as a *fork* of the widely-used IIF viewer Mirador. We added video and audio playback functionalities, and also developed a plugin for image and audio-visual resource annotation. Building on the annotation functionalities of MemoRekall, the user can configure an annotation in the following way:

- Annotation *metadata*. An annotation can have any number of metadata attached to it (typical fields include *title*, *author*, *tags* etc.). By default, all metadata will be Dublin Core compliant, however we wish to enable any kind of ontological model. Every annotation can also target a specific region of the document in question, be it in the spatial or temporal dimensions. When linked to a region, a certain degree of interactivity with the document can be enabled, for example clicking on an annotation will jump playback of a time-based media to the beginning of the annotation's start time.
- Annotation *content*. As well as metadata, we can add content to an annotation. This can go from simple textual description augmented with html tags, to overlays that will be displayed on top of the related media. This content can be freely created and edited within Arvest.
- Linking to *external documents*. Following the inter-documentary approach to annotation, it is also possible to link to other documents. Like in MemoRekall, it is possible to link to static files or links on the web, but a powerful addition is the ability to link an annotation to another IIF Manifest. When this is the case, Arvest will allow the user to directly open the Manifest within the environment, and thus fluidly navigate a network of IIF Manifest. This functionality can also be used to make data-driven interfaces. For example, nodes on an image representing a network of documents can be annotated so that clicking on a node in the image will directly open the document in Arvest — thus allowing a rapid switch between distant and close reading perspectives. We demonstrate the power of this kind of interface in the third part of this article.

70 All of these Manifest editing functionalities will be wrapped within a multiuser environment, written in Django. This comes in answer to several aspects which we found could be significant hurdles to overcome for newcomers to IIF, especially individual users with low techno-fluency:

- *Connect and play*. Using IIF today can be complicated - often users will find themselves cloning git repos, running servers from the command line, editing JSON etc. Another alternative is to have an external provider come and set up a working environment for you or your institution - naturally this is not an option in many cases. Arvest will allow users to simply create a free account, connect to the web application and start using IIF without even having to know what it is.
- *Media storage*. Each user will have a certain amount of storage space where they will be able to stock media and Manifests. We will also provide an instance of PeerTube for the handling of large video files. Users will have full control over their content, meaning that they can expose their IIF-compliant content to the world, or keep everything private.
- *Collaborative work*. Users can share their projects between each other, and also share them as embeddable html elements. Users retain full access control of their projects, allowing them to be edited or not by other users.

71 As mentioned above, the user can consult a [first version](#) of the Arvest environment. It should be noted that, as of yet, no real work around UX design of the interface has occurred, and we mainly adhere to Mirador's native UX choices. Mirador also supports themes which can radically change the feel of the environment (accessed through the left-most toolbar, under *Workspace settings > Change theme*).

## 72 **ML-driven workflow integration**

73 The functionalities presented in the previous subsection allow for manual creation and modification of projects and annotations, principally allowing for work in a close-reading approach. We also wish, however, to integrate distant-reading and viewing techniques, allowing Arvest to be a complete environment to create content in these two complementary approaches and navigate between their output. Future development will seek to integrate curated ML algorithms directly within the tool, however in this first phase of development, we decided to propose an open API for the creation, consultation and modification of content within Arvest (be it media, projects or annotations) from afar. This allows us to integrate Arvest in complex ML-driven workflows without any extra overhead on the development end. As mentioned above, the multi-user environment is written in Django and designed to function with HTTP requests, allowing for the creation of a RESTful API. Furthermore, IIIF Manifests are interoperable JSON documents which can be created by any number of methods and be directly exploitable within Arvest.

74 With Arvest we are seeking to bridge the gap between non techno-fluent users and the kidneys of tools and environments discussed in the first part of this article. In the spirit of this project, we are currently assembling a [library of python notebooks](#) that allow for the creation of content that can be directly opened in Arvest using various ML algorithms and workflows. The notebooks are designed to be run with no installation needed directly online within [Google Colaboratory](#), and also with very little setup needed locally in [Anaconda](#). These notebooks are written with a pedagogical tone, and are designed to be novice-friendly, and robust enough to use in a workshop setting with the user's own content.

75 The notebooks are split into two different categories: first basic Arvest functionalities such as accessing and uploading media, reading and creating Manifests. We also offer utility scripts such as batch media uploads. The second category are common ML workflows that have been developed in the context of various case studies that are driving Arvest's development. We discuss one such project in detail in the following section, and other case studies will be the subject of detailed analysis for future publications. However we can already propose the following workflows:

- Automatic speech recognition using [vosk](#). Derived from a case study in theatrical studies, video or audio can be decomposed into words, and annotations are created for each word detected.
- Image embedding projection with the [Distant Viewing toolkit](#). Derived from a case study in history of art, a workflow similar to a tool like [PixPlot](#). The user can take a corpus of still images and use an image embedding neural network model and dimensionality reduction algorithms such as UMAP or T-SNE to project them into a space where images that are close to each other are similar. An interactive Manifest is created of the projection, where clicking on the image in the projection will open the original image in Arvest.
- Video decomposition using the Distant Viewing toolkit. Derived from a case study in history, a large corpus of archival video footage we decomposed into shots and then

projected into manifold spaces like described above in order to make navigation of the corpus easier.

- Actor-network analysis using networkx and link to other platforms. In the context of a case study around the archives of a music creation centre, we perform analysis on the contents of the archives that are stored on Nakala and Heurist. We can create network visualizations of the various records that can be navigated in Arvest, and when a node is linked to a media file, the file can be viewed directly.

76 One remarks that each of these case studies look at incorporating other digital humanities projects like vosk, the Distant Viewing toolkit, and platforms like Nakala and Heurist. Indeed, we intend that Arvest integrate itself into this digital humanities environment, and serve as a platform where various approaches and tools can work together seamlessly. This library of notebooks is also intended to give a map and a starting point for the various ML-driven workflows that are commonly used in digital humanities — it can not only be a resource to learn about feature extraction, dimensionality reduction, classification etc., but also a way of rapidly testing a workflow on a given corpus with immediate interactive output in a platform like Arvest, and a starting point in the code which can be modified.

## 77 **Using Arvest to interpret digital traces: deriving collaboration analytics from multimodal document networks**

78 Following our co-design approach, we developed the prototype of Arvest and detailed its specifications along with different case studies. Wishing to anchor development of the tool in real scholarly practice, we have investigated the affordances and limitations of this kind of environment when put to the service of a semantic line of questioning.

79 We present below a detailed case study centred around the Pilot 2 project Dancing Philosophy in the context of the EU Horizon 2020-funded COESO project (Collaborative Engagement on Societal Issues). This Pilot, led by dancer-choreographer Cosetta Graffione and philosopher Stefania Ferrando, looked to explore “*new ways of knowing through words and dance*” (Ferrando 2022). Ferrando and Graffione met on seven different occasions from October 2021 to June 2022 in France and Italy where they lead workshops with dance students, other artists, researchers and local citizens (Blin, Marranca 2022) (see also the blog documenting the project and its different steps: <https://coeso.hypotheses.org/498>. We collected the digital traces of this research creation, during its course and afterwards. The collection of traces consists of texts, photographs, videos, notebooks, and drawings (which have been digitized). How can we derive collaboration analytics from a collection of multimodal digital traces, and gain insight on the nature of the collaboration that occurred? How can automated analysis and the programmatic generation of data-driven navigation interfaces participate in a research project that looks to track an intangible, collaborative creative process? What can software offer us in terms of possibilities for interface and navigation of this information that will prolong our analysis and open new avenues for research?

80 Focusing on Pilot 2, we looked to discover how collaboration allowed participants to displace/develop/question each other’s respective fields and practices (in this case, contemporary dance and philosophy), and to understand the nature of the collaborative processes that occurred. In order to best approach this question, we based our research on recent work from Boullier and Pidoux (Boullier, Pidoux 2021) that was also carried out in the context of the COESO project. In their paper, Boullier and Pidoux offer a conceptual

framework for understanding the nature of a collaborative process. Methodologically, Boullier and Pidoux allow for the ongoing collection of data during a project. They intend that their methods and the artefacts they produce will participate in a feedback loop with a project's participants, allowing them to reflect on (and recalibrate) the nature of the collaborative work they are undertaking. Working with a small corpus of traces that documents Graffione and Ferrando's Pilot study, we wish to examine how Boullier and Pidoux's model can be applied once an on-going project has been completed.

## 81 Collaborative analytics

82 In their *collaborative analytics*, Boullier and Pidoux propose a number of cooperation features such as *skills, culture diversity, legal and ethical compliance* etc. Each of these features are given categorical values that impact the *cooperation typology* of a collaboration. The cooperation typology is visualized according to a compass model. A feature will be categorized into the *adaptive, plan oriented, institutional* or *revisable* poles of this compass. For each feature, different categories will correspond to different typological poles: for example, with *skills, experimental* corresponds to *adaptive, academic-expert* to *plan oriented* and *procedural* to *institutional*.

83 Boullier and Pidoux then construct a number of *indicators* for each of these cooperation features. Some features have one indicator attached to them (for example, *skills* has the indicator *skill type*), some several (for example *rhythm of citizen/research participation* has *frequency of conversation and contributions* and *frequency of meeting*). Each indicator has a number of different *data* that inform it: for example, the *skill type* indicator is derived from the data "*words to identify for every user and classify according to feature categories*". These data are further attached to *data sources* (for the previous example, the source is *message content*). Methods of analysis are proposed for each indicator, such as NLP, qualitative analysis, multidimensional analysis etc.

## 84 COESO Data

85 Our corpus is comprised of 34 documents, including videos, audios, photos, Laban notations, and written documents:

86 |#|Title|Type|Description|Date|Place| ---|---|---|---|---|--- 1|Desire|Video|Edited recording of final public performance after workshop.|2021-10-03|St. Exupéry Cultural Centre of Wissous (France). 2|Public workshop|Video|Edited recording of the final performance after workshop with members of the public.|2021-10-03|St. Exupéry Cultural Centre of Wissous (France). 3|Public performance|Video|Edited recording of final public performance after workshop.|2022-04-03|St. Exupéry Cultural Centre of Wissous (France). 4|Choreographic phrase recording|Video|Recording of CG working on a choreographic phrase.|2022-04-03|St. Exupéry Cultural Centre of Wissous (France). 5|Burning heart|Video|Edited recordings of work during the workshop.|2022-04-11 – 2022-04-12|Hauts-de-France Polytechnic University, Arenberg (France). 6|Wandering|Video|Edited recordings of work during the workshop.|2022-04-11 – 2022-04-12|Hauts-de-France Polytechnic University, Arenberg (France). 7|Improvised solos|Video|Recording of CG and SF working on solos during workshop.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 8|Teaser of Ferrando's solo|Video|Exraits of the *Impovised solos* video teasing SF's solo.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 9|Teaser of Graffione's solo|Video|Exraits of the *Impovised solos* video teasing CG's solo.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 10|Alma Ballet

Exercises|Video|Edited recording of different groups of students working on choreographies.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 11|Centre (Graffione’s description)|Audio|CG describes the movement around the word *Centre*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 12|Centre (Ferrando’s description)|Audio|SF describes the movement around the word *Centre*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 13|Organic (Graffione’s description)|Audio|CG describes the movement around the word *Organic*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 14|Organic (Ferrando’s description)|Audio|SF describes the movement around the word *Organic*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 15|Sober (Graffione’s description)|Audio|CG describes the movement around the word *Sober*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 16|Sober (Ferrando’s description)|Audio|SF describes the movement around the word *Sober*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 17|Truth (Graffione’s description)|Audio|CG describes the movement around the word *Truth*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 18|Truth (Ferrando’s description)|Audio|SF describes the movement around the word *Truth*.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 19|Alma Ballet workshop photos|Photos|Photos of the workshop that features CG, SF and the Alma Ballet students.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy) 20|Laban Kinetography of a choreographic phrase|Laban Kinetography (PDF)|Laban Kinetography of the phrase worked on in the *Choreographic phrase recording* video.|2022-04-03|St. Exupéry Cultural Centre of Wissous (France). 21|Laban Kinetography|Laban Kinetography (PDF)|Laban Kinetography of the movements around the words described in the various audio recordings.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy). 22|Alma Ballet Laban Kinetography|Laban Kinetography (PDF)|Laban Kinetography of the exercises performed by the Alma Ballet students.|2022-04-19 – 2022-04-23|Alma Danza Centre, Bologna (Italy). 23|Freedom and the practice of authenticity in Italian feminism.|Written document (PDF)|Text written by SF for a presentation at a conference.|2021-11-03|Univeritat de Barcelona, Barcelona (Spain). 24|“Starting with words” intentions|Written document (PDF)|Text stating the intentions for the workshop in Wissous.|2022-03-23 – 2022-04-02|St. Exupéry Cultural Centre of Wissous (France). 25|About silence|Written document (PDF)|Text written by SF about *Silence* (the word at the base of her solo).|Paris (France). 26|Graffione’s intentions|Written document (PDF)|Text written by CG about the solos and the words that triggered them.|2022-04-01 – 2022-05-01|Paris (France). 27|About the workshops|Written document (PDF)|A text recounting CG’s thoughts on the various workshops.|2022-05|Paris (France). 28|Graffione’s sketches|Written document (PDF)|CG work documentation: photo of pages from a book in Italian.|2021-09|St. Exupéry Cultural Centre of Wissous (France). 29|Founding texts|Written document (PDF)|Philosophical texts by Pina Bausch and Gualtieri.|2022-03-22 – 2022-04-03|St. Exupéry Cultural Centre of Wissous (France). 30|Graffione Intentions|Written document (PDF)|CG writes about intentions for the workshops.|2022-05-01|Paris (France). 31|Ferrando’s notes|Written document (PDF)|Photos of pages from SF’s notebooks. |St. Exupéry Cultural Centre of Wissous (France). 32|Graffione’s notes|Written document (PDF)|Photos of pages from CG’s notebooks.|2022-03-01|St. Exupéry Cultural Centre of Wissous (France). 33|Desire and relationships|Written document (PDF)|A text written by SF on desire and relationships. |EHESS-Paris and UPHF-Valenciennes (France), Savona (Italy). 34|La parola del corpo|Written document (PDF)|A text written by CG and SF about the project.|2021-11-01 – 2022-03-01|EHESS-Paris (France), Savona (Italy).

- 87 This corpus is a curated sub-selection of the entire COESO and Pilot 2 archives: although small, this collection of multimodal documents still offers many perspectives on some of the outcomes of the Pilot study and the creative process that occurred. Some characteristics of this collection are worth noting:

- They are *fragmentary* documents. Somewhat contrary to many of the kinds of documents described in Boullier and Pidoux's model, these documents emerge in a sporadic manner. We cannot hope to attain a complete overview of the project in an ongoing, continuous manner; only certain moments in time that find themselves entwined within these artefacts.
- They are *subject to mediation*. Contrary to the kind of raw data that Boullier and Pidoux manipulate, these documents have been edited, curated and selected by various actors.
- They are *aesthetic objects*. Many of these objects emerge from a creative process. Some document events, some are final reports or recordings of performances, some are between these poles. In any case, they differ from the communicative, day-to-day human interaction sources proposed by Boullier and Pidoux.

88 Despite these differences, is it still possible to attain collaborative metrics like those proposed by Boullier and Pidoux from this content? Can we measure collaboration from this data?

## 89 Cooperation features

90 Given the nature of collaborative work and the artefacts that emerge from it, we found it useful to conceive of a document as a node in a network. A first task was to examine the various indicators proposed by Boullier and Pidoux and see to what extent they can be adapted and measured when the data source is one of these nodes.

91 Here the reader may find a complete assessment of the cooperation features, categorical values, indicators and corresponding data sources proposed by Boullier and Pidoux and how we propose to adapt them to our corpora. The reader will notice an immediate tension between some of the metrics proposed and how these are to be categorized into typological categories.

92 Let us take an example with the feature *Dissemination type of results*, the indicator for which is *degree of field hybridation*. Boullier and Pidoux originally propose to derive this from *NLP analysis of reports, a list of actors that disseminate the product, where a product is disseminated, what is the type of product being disseminated and what the format of the product is*. These can be qualified as quantitative metrics, however there is no indication as to how they then inform the three categorical values of *open (adaptive)*, *divided according to fields (plan-oriented)* and *academic oriented (institutional)*. Indeed, it appears to be left up to the researcher to assess these fields and then categorize the results in a qualitative manner. In our assessment, we indicate how these categorizations could be considered for documents in the column *Node data + method of analysis*. Considering that even with quantitative gathering of certain metrics the final qualitative categorization will ultimately be left to the researcher, and given the small size of the corpus, the decision was made here to bypass the step of quantitative data collection and allow the researcher to make a categorization for each feature straight away.

93 The categorization was performed for each node in our document collection. Next, this data was taken and 'normalized' so that each document could be placed on Boullier and Pidoux's typological compass. In Figure 1 the reader can find an example of two different documents visualized in this manner (on the left, the video *Desire* and on the right *Ferrando's notes*). We also calculate the centroid of this resulting polygon so that it may be represented by one point on the compass. This allows for a composite visualization of the entire collection

which can be viewed in Figure 2. The reader can also consult the [full results](#) here for the full composite data file used to construct the final visualization.

96 The normalizations were obtained by taking the typology matrix for each feature and getting an average for each typology. Note that we do not include the features that were not informed for the document. The absence of certain fields from some documents can weigh the visualization somewhat – we have not found a solution for accounting for this in the visualization itself, however, in the [full results](#) the reader can also find the loss field for each document, which is a metric indicating the percentage of fields that were informed (a loss of 0.2 means that 80% of features were informed). This is a factor that is to be considered when interpreting the results.

## 97 **Collaboration analytics with Arvest**

98 Looking at these visualisations, different questions arise. For example, why are most of the documents oriented towards the *plan-oriented* typology? What are the three points in the *adaptive* category? What are the two outlying points? Before we begin to answer these questions, it is clear that a dynamic way of navigating this multidimensional data would be useful (the various dimensions are: feature value for each document; composite for each document; composite for the entire collection). This is where our tool Arvest can offer some solutions.

99 Levying Python scripting we can rapidly create interactive visualizations for the Arvest environment. In Figure 3, we see the data we have created loaded as a [manifest network](#) in [Arvest](#). On the left, in the composite visualization of the document collection, each point on the compass is linked to an annotation which allows the reader to open the corresponding document's Manifest in another pane in Arvest. On the right, one of these Manifests features the document in question, and also an annotation that paints the document's typological compass below it. This allows us to rapidly navigate this multidimensional data, pivoting between different views, and even simultaneously viewing the distant and close reading perspectives.

101 The collection tends towards the *plan-oriented* corner of the compass. This is defined by Boullier and Pidoux as consisting of collaborations of short-duration and of a relatively formalized nature. The pilot gave place to a series of short workshops which were led according to a formalized scientific method. This composite view and typological tendency confirm the suppositions that we could make about the collaborative work that occurred.

102 Taking a closer look, we have identified some notable points of interest which are labelled in Figure 4. We shall explore these documents using Arvest and offer some interpretation as to how they may have ended up in the places they occupy.

104 The documents in Grouping A are outliers compared to the rest of the collection of documents. They are in fact 8 documents which are all identical in typological nature, and therefore superimposed on the visualization. The 8 documents are the audio recordings made by Graffione and Ferrando that describe the movements depicted in the Laban notations around the four words *truth*, *organic*, *sober* and *center*. They all tend towards the *institutional* corner of the compass. When looking through the data, we see that this can be explained by the fact that the recordings were made specifically for a final report for the COESO Project (*Blin, Marranca 2022*). They are only accessible through a link that is made available through this scientific paper, and it can be argued that had they not have had to

write this report, these recordings (which ‘make up’ for an absence of video documentation) would not have been made. Furthermore, they each involve one actor, speaking in their own language.

- 105 The documents in Grouping B are comprised of videos documenting work that occurred during public workshops. They are the only documents in the collection that breach the line from *plan-oriented* towards the *adaptive* corner of the compass. The contextual network of the documents is growing through reaching out to the public. They do not stray too far from the pack – indeed, as open and adaptive as these workshops need to be, they are still run with a scientific and formalized methodology.
- 107 Grouping C describes two documents: photos of *Graffione* and *Ferrando’s notebooks*. Their positions on the compass throw light on an eventual limit to the methods we are exploring. When looking at the local typological compass (see Figure 5), we see that the documents in fact tend extremely towards the adaptive and institutional corners of the compass, hence the final position which is derived as the centroid of all informed fields. Note also that these documents have an extreme *loss* factor: 0.46, meaning that barely half of the fields could actually be informed. This pushes on the importance of having access to different perspectives when looking at this kind of multidimensional, data-driven interface as quantitative analysis can sometimes hide the realities of the dataset it is describing. We can perhaps explain this typology as being derived from a low amount of plan-oriented work (these are personal notebooks in which the actor writes in an ad hoc manner).
- 108 Grouping D is comprised of textual documents such as *Graffione’s intentions* and *Laban Notations*. These documents tend the most towards the *plan-oriented* corner of the compass. Indeed, they either derive from or state a somewhat rigid methodology that steers the document away from the adaptive side of the scale. The *Laban notations* notably also require an expert knowledge to be interpreted, and therefore tend the document away from *adaptability* and *openness*. These are also documents created by a sole author.
- 109 Finally, Grouping E is comprised of an ad hoc collection of documents and can be considered to represent the general tendency of the collection’s typology. We notably find textual documents that offer overviews of the Pilot’s workings and methods such as *Choreographic score*, *About the workshops*, *About silence*, *La parola del corpo* and *Desire and relationships* as well as the two video documents that are made as teasers for the project. Their central position can therefore be accounted for by their global perspective on the project.

## 110 **Conclusion**

- 111 To answer to the challenge of interpreting digital traces in a digital environment, combining qualitative and quantitative methods, traditional and innovative approaches, we propose Arvest, a solution dedicated first to historians but which can be of interest to other fields. This paper presents the history and the different development phases which structured the development, as well the main issues Arvest has faced and its specifications. The development, based on a co-design approach, is ongoing, along with case studies. We presented a case study based on the COESO project. We adopted Boullier and Pidoux’s model of collaborative analytics, and examined if it was possible to apply this model to a collection of multimodal documents. The possibilities offered by Arvest allowed us to navigate our multidimensional, data-driven interface and interrogate the documents in new ways. The multiple layers of analysis engendered an analytical process that encouraged new

ways of perceiving the corpus. Essentially, the visualizations were a means of rendering and engaging with the researchers' own subjective inflections and suppositions around a corpus in a structured manner.

- 112 A first version of Arvest is already available. It is part of the ongoing development of the ERC-funded project From Stage to Data, the Digital Turn of Contemporary Performing Arts Historiography (STAGE). Arvest will be used to model creative processes and particularly their collaborative dimension, taking into account not only the rehearsals but also all the data produced by the team members. The major point in terms of development going forward for Arvest is a need to dynamically create bespoke interfaces like the network and composite typological interfaces shown in the COESO case study. These interfaces are a powerful tool to access many different data-driven perspectives on the collection. As demonstrated with the typological compass, the nature of these interfaces will differ according to the type of analysis and research question being carried out. It will be of utmost importance to offer tools for the programmatic and manual creation of these interfaces within the tool going forward.

## 113 **Acknowledgments**

- 114 Funded by the European Union (ERC, STAGE, grant agreement no. 101097091). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.
- 115 With the support of Région Bretagne.
- 116 With the support of Ouest Valorisation, Société d'Accélération du Transfert de Technologies.
- 117

## **Bibliography**

Arnold, T. and L. Tilton. "Distant Viewing: Analyzing Large Visual Corpora". *Digital Scholarship in the Humanities*, vol. 34, no. Supplement\_1, pp. i3–i16, <https://doi.org/10.1093/llc/fqz013>.

Arnold, T. and L. Tilton. *Distant Viewing: Computational Exploration of Digital Images*. The MIT Press.

Arnold, T. and L. Tilton. *The Distant Viewing Toolkit*. 0.3.1, <https://distant-viewing.github.io/dvt/index.html>.

Aronson-Lehavi, S., et al.. "Comparative Video Annotation and Visual Literacy: Performance Analysis of Rina Yerushalmi's Theatre Language". *International Journal of Performance Arts and Digital Media*, vol. 17, no. 1, pp. 86–101, <https://doi.org/10.1080/14794713.2021.1893491>.

Bachimont, B., et al.. "Preserving Interactive Digital Music: A Report on the MUSTICA Research Initiative". *Proceedings Third International Conference on WEB Delivering of Music*, pp. 109–12, <https://doi.org/10.1109/WDM.2003.1233883>.

Bardiot, C. and M. Menassel. *Video Annotation for Circus Education. Memorekall and the COSMIC Project*. Rennes 2, <https://hal.science/hal-03828472>.

Bardiot, C., et al.. *Documenting Collaborations with Video Annotation: Memorekall Capsules (COESO D.2.4)*. <https://doi.org/10.5281/zenodo.6619635>.

Bardiot, C.. *Performing Arts and Digital Humanities: From Traces to Data*. ISTE / Wiley.

Bardiot, C.. "Theatre Analytics: Developing Software for Theatre Research". *Digital Humanities Quarterly*, vol. 14, no. 3,  
<http://www.digitalhumanities.org/dhq/vol/14/3/000476/000476.html>.

Bateman, J. A.. *Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents*. Palgrave Macmillan.

Blin, I.and D. Marranca. *Dancing Philosophy - Choreographic Score (COESO D.2.5)*.  
<https://doi.org/10.5281/zenodo.6788008>.

Boullier, D.and J. Pidoux. *Cooperation Quality Assessment: Development and Implementation of the Cooperation Analytics*. <https://doi.org/10.5281/zenodo.5599052>.

Cairo, A.. *The Truthful Art: Data, Charts, and Maps for Communication*. New Riders.

Carboni, N., et al.. "Pister Des Circulations Visuelles À L'échelle Mondiale". *Humanistica 2023*, Association francophone des humanités numériques, <https://hal.science/hal-04094170>.

Cross, N.. *Design Thinking: Understanding How Designers Think and Work*. Bloomsbury/Berg.

Düring, M., et al.. "Transparent Generosity. Introducing the Impresso Interface for the Exploration of Semantically Enriched Historical Newspapers". *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, vol. 57, no. 1, pp. 20–40,  
<https://doi.org/10.1080/01615440.2024.2344004>.

Drucker, J.. "Humanities Approaches to Graphical Display". *Digital Humanities Quarterly*, vol. 5, no. 1, <https://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html>.

Drucker, J.. "Performative Materiality and Theoretical Approaches to Interface". *Digital Humanities Quarterly*, vol. 7, no. 1,  
<https://digitalhumanities.org/dhq/vol/7/1/000143/000143.html>.

Ferrando, S.. *The Movement of an Embodied Thought: Pilot 2 Dancing Philosophy Report (COESO D.2.6)*. <https://doi.org/10.5281/zenodo.6788039>.

Flanders, J.and F. Jannidis, editors. *The Shape of Data in the Digital Humanities: Modeling Texts and Text-based Resources*. Edited by Julia Flanders and Fotis Jannidis, Routledge, Taylor & Francis Group.

Ginzburg, C.. *Clues, Myths, and the Historical Method*. Translated by Anne C. Tedeschi and John Tedeschi, John Hopkins University Press.

Graham, S., et al.. *Exploring Big Historical Data: The Historian's Macroscope*. Imperial College Press.

Groeger, L.and J. Schweitzer. "Transformational Leadership, Design Thinking and the Innovative Organization". *Papers from the 30th EGOS Colloquium 2014*, European Group for Organizational Studies EGOS, pp. 1–2.

Hou, Y., et al.. "Digitizing Intangible Cultural Heritage Embodied: State of the Art". *Journal on Computing and Cultural Heritage*, vol. 15, no. 3, pp. 1–0, <https://doi.org/10.1145/3494837>.

Krueger, R. A. and M. A. Casey. *Focus Groups: A Practical Guide for Applied Research*. 5th edition, SAGE.

Latour, B.. "Le "pédofil" De Boa Vista - Montage Photo-philosophique [1993]". *Petites Leçons De Sociologie Des Sciences*, La Découverte, pp. 171–224.

Manning, P.. *Big Data in History*. Palgrave Macmillan.

McCarty, W.. "Modeling the Actual, Simulating the Possible". *The Shape of Data in the Digital Humanities: Modeling Texts and Text-based Resources*, edited by Julia Flanders and Fotis Jannidis, Routledge, Taylor & Francis Group, pp. 264–84.

Milligan, I.. *The Transformation of Historical Research in the Digital Age*. Cambridge University Press.

Moretti, F.. *Graphs, Maps, Trees: Abstract Models for a Literary History*. Verso.

Moretti, F.. "operationalizing": Or, the Function of Measurement in Modern Literary Theory. Pamphlet 6. Literary Lab, Stanford, <https://litlab.stanford.edu/LiteraryLabPamphlet6.pdf>.

Noordegraaf, J., et al., editors. *Preserving and Exhibiting Media Art: Challenges and Perspectives*. Edited by Julia Noordegraaf et al., Amsterdam University Press, <https://www.oopen.org/record/530353>.

Oiva, M. and U. Pawlicka-Deger. "Lab and Slack. Situated Research Practices in Digital Humanities - Introduction to the DHQ Special Issue.". *Digital Humanities Quarterly*, vol. 14, no. 3, <http://www.digitalhumanities.org/dhq/vol/14/3/000485/000485.html>.

Prost, A.. *Douze leçons sur l'histoire*. Édition augmentée, Points.

Rosenzweig, R.. "Scarcity or Abundance? Preserving the Past in a Digital Era". *American Historical Review*, vol. 108, no. 3, pp. 735–62, <http://chnm.gmu.edu/resources/essays/scarcity.php>.

Rygiel, P.. *Historien À L'âge Numérique*. Presses de l'Enssib.

Smits, T. and M. Wevers. "A Multimodal Turn in Digital Humanities. Using Contrastive Machine Learning Models to Explore, Enrich, and Analyze Digital Visual Historical Collections". *Digital Scholarship in the Humanities*, vol. 38, pp. 1267–80, <https://doi.org/10.1093/llc/fqad008>.

Tremblay, P. A., et al.. *The Fluid Corpus Manipulation Toolbox (v. 1)*. Zenodo, <https://doi.org/10.5281/zenodo.6834643>.

Waterhouse, E., et al.. "'I Gave That Cue.' Integrating Dance Studies, Praxeology, and Computational Perspectives to Model Change in the Case Study of William Forsythe's *duo*". *International Journal of Performance Arts and Digital Media*, vol. 17, no. 1, pp. 160–81, <https://doi.org/10.1080/14794713.2021.1884803>.

Zacklad, M.. "Réseaux Et Communautés D'imaginaire Documédiatisées". *A Document (re)turn*, Peter Lang, pp. 279–97.