Immersive Datascapes (and other modes of knowing)
Project by Giulia Taurino

Media Studies Grant Report

Project presented

*Immersive Datascapes (and other modes of knowing)* is a research&creation project that explores the epistemology and design of online archives and databases deployed for new media research. It is meant to propose an interactive data visualization designed to study MIT Docubase, an ongoing curatorial initiative by MIT Open Documentary Lab for collecting data on “people, projects and technologies transforming the documentary in the digital age.” Its expected output is an immersive virtual experience that uses adaptive navigation to favor new modes of communicating with, accessing to, and learning about audiovisual archives.

Partners involved

*Immersive Datascapes (and other modes of knowing)* was developed by media researcher and artist Giulia Taurino, in collaboration with MIT Open Documentary Lab (Cambridge, MA - United States) and SAT - Society for Arts and Technology (Montreal, QC - Canada). On the one hand, as a research fellow at MIT Open Documentary Lab, I had access to MIT Docubase’s data, as well as to a series of academic resources. In particular, I received important and extensive feedback on the project development by William Uricchio, Founder and Principal Investigator of MIT ODL, Sarah Wolozin, Director of MIT ODL, and Katerina Cizek, Artistic Director, Co-Founder and Executive Producer of the Co-Creation Studio at MIT Open Documentary Lab. On the other hand, SAT offered the technical support (VR expertise, material, on-site and virtual space) for the creation of the immersive experience. I worked in close collaboration with Emmanuel Durand and Nicolas Bouillot from SAT métalab research cluster, who were fundamental in the creation of the project.

Timeline and Financial Support

The realization of this project has been possible thanks to the support of the The International Federation of Television Archives, which gave me the unique opportunity of being awarded the Media Studies Grant. The project was developed in two parts covered by the Media Studies Grant’s installments, respectively following
the work from June to August and from September to October 2021. The overall development length is 5 months.

- **Phase I - Data Discovery and Visualization.** Built upon previous work of data collection and study of the database (MIT Docubase), the first phase included data discovery and several iterations of visual models (spiral chart, ridgeline plot, network visualization, matrix/scatter plot visualization) by the grantee, Giulia Taurino.

- **Phase II - Immersive Prototype Development.** The second phase included a series of experiments in the development of an interactive prototype for an immersive on-site and on-line experience.

Overall, the grant helped me bring the project from the concept to the development phase and, even more, to a creative journey of exploration of new immersive methods for knowledge discovery that will be useful for the Media Studies community.

**Questions asked and case study**

The project started from a few questions:

- How can we document the very process of curating digital archives and databases?
- How can we frame their evolution, and define what type of knowledge they provide at different stages?
- What do they tell us? What’s missing?

Working on MIT Docubase as a case study, the project pays specific attention to new media archives and databases that include content created using audiovisual immersive media, interactive and internet-based technologies, and other forms of digitally mediated visual content. By doing so, it outlines the risks and challenges of born-digital records as to what concerns processes of formatting, archiving, accessing, preserving, and ultimately, curating new media repositories.

**Techniques and methodology**

Reasoning on questions related to what digital, multimedia archives curate, frame, define and miss, as well as on their archival uncertainties (cf. Thylstrup *et al.*, 2021), this project uses methodologies traditionally found in Digital Humanities, including creative coding (Python), data visualization techniques (Python, Javascript), and immersive technologies (Unreal Engine, Edition In Situ Software, Mozilla Hubs). In its conceptualization and theoretical framework, the project draws upon a variety of perspectives in critical making (Ratto 2011) and speculative design (Dunne and Raby 2013). In its core practice, it takes an experimental and exploratory co-creation approach to promote a participatory experience that engage communities of researchers, artists and media practitioners in a co-curation experiment.
Goals

The development goal set for this project is to create a full dome, interactive projection and visual guide for reading MIT Docubase and asking the larger public to actively re-invent its composition and engage with the database as a lively, dynamic, co-curated collection. This artistic, exploratory approach aims at introducing new points of view on ways of caring, curing, curating digital archives, in a deeper attempt to foster a process of unlearning and activate other modes of knowing. The immersive data visualization experience ultimately aspires to unveil invisible narratives contained in the database and to problematize their socio-cultural impact, while representing much more than a static repository: a documentary act itself.

Project activities

From the MIT Docubase dataset, I was able to extract information about over 300 documentaries released between 1989 and 2021, with metadata related to Title, Description, Country, Year, Topics, Technologies, and links to the audiovisual material available. What I was interested in showing in the visualization was the intersection between 4 dimensions that account for the techno-cultural transformations as linkages between country of production or origin of the content, year of release, technologies used and topics covered. Bearing this in mind, phase I (i.e. Data Discovery and Visualization) comprised several iterations and steps.

- **Phase I - Static Visualizations.** I initially explored static visual models such as the spiral chart, which highlighted the evolution of interactive documentary production over the years (as per MIT Docubase data), and the ridgeline plot, which focused on density of production per country (as per MIT Docubase data).

- **Phase I - Interactive Visualizations (Network).** For the first few iterations of an interactive visual model, I decided to explore the dataset through a simple network visualization in order to get a glimpse on the distribution of technologies across content and on their degrees of interconnectedness. What you see in fig. 1-5 in the Appendix is a series of screenshots from a network visualizations, explored in different color selections with white or black background and originally designed for a projection on the semi-spherical surface of SAT Satosphère space. As you notice in fig. 1 and 2, each node represents a technology used in the development of the documentaries found in the dataset. Some are less connected to other technologies, like Text Mining and some are more central and widely used technologies like photography or video (fig. 3). When we resize the nodes based on the degree of connectivity (fig. 4), we notice that Flash appears among the most common technologies, hinting at the fact that a consistent number of projects in the database were either removed or updated due to the termination of Flash-based
websites happened after 2020. This is an example of how this project does not focus on presenting a definitive visualization, but rather tries to document the traces of a dynamic state of knowledge in the curation of new media databases and archives.

- **Phase I - Interactive Visualizations (Matrix).** However insightful as well as adaptable to a spherical surface, the network visualization didn’t seem to be ideal for showing the complexity of each intersection between geographic location, year of release and technologies used. I wanted to understand and problematize the density of content as related to all three parameters, and possibly a fourth parameter of topics, and take advantage of the three-dimensionality of the space. Furthermore, a network visualization is in its most common definition and use a mapping of an existing network, with existing nodes, edges, and weights assigned to them. While it can account for centers and peripheries, it still doesn’t map the missing media content, objects, histories. I therefore started exploring the possibility of an interactive, 3d matrix visualization. More than a scheme of nodes and edges that gives a holistic overview of the database, matrices allow us to account for co-variations based on the link between variables, such as year and country, allowing many relationships to be explored in the scatter plot. They also appear to be more efficient in showing variations in the density of content as well as absences. Fig. 6 to 12 in the Appendix show the matrix visualization in the form of a 3d scatter plot across several iterations. In fig. 6, I present the first iteration of a black and white static version of the scatter plot. In fig. 7 I explored different distributions of the data points, always maintaining the relationship between year, country and technologies. I eventually opted for retaining the shape observed in fig. 8, which shows the year progression on the y axis on a range from 1989 to 2021, the production country on the x axis ordered from the most to the least recurrent geographic location, and the technology used on the z axis. In another set of iterations (fig. 9), I explored different options for color gradients in order to make visible the changes in topics per year, country, and technology. I eventually chose the colormap viridis (fig. 10). This colormap was designed to improve graph readability for users with common forms of color blindness and/or color vision deficiency. It is also perceptually-uniform, meaning that values close to each other have similar-appearing colors, a feature that persists even when converting to black-and-white. In this 3d model there is no fixed point of view, but a multiplicity of perspectives. Even when rotated upside down and looked at from a different angle as a sort of cascade rather than a peak (fig. 11, 12), this matrix vision gives justice to the missings in media production, to potential histories, and potential data points that could be or could have been.

The second phase in the research and realization of this project was the development of an interactive, immersive experience and adaptive navigation process for virtual environments. For this purpose, I developed a simplified interactive version of the
project ready to be tested for users navigation (fig. 13). More specifically, the second phase (Immersive Prototype Development) involved two steps as it follows.

- **Phase II - Immersive Prototype Development (Local Host).** This interactive prototype recreates the visualizations previously shown. In the 3d prototype we can efficiently scan the matrix, define pairings, identify strong or weak relationships, and locate missing data points. we can zoom in and take a closer look at the information related to each object (fig. 14) or zoom out and look at how a certain documentary positions itself in the grid and topic list (fig. 15). The hover animation also highlights the grid so to make the navigation smoother and facilitate the reading (fig. 16). Video 1 in the Appendix gives a glimpse on the level of interaction currently implemented. We hope to be able to add further interaction with the content itself, such as trailers when available or other visual material related to each documentary, so to encourage a close reading in addition to the distant reading.

- **Phase II - Immersive Prototype Development (Onsite + Online).** Together with Emmanuel Durand and Nicolas Bouillot of SAT, we are currently considering two main options for the development of the immersive environment. A first option would be to use SAT videomapping tool and software Edition in Situ to create a full-dome prototype scene in an immersive audio-visual installation at Satosphere. The second option is to reproduce a scene presenting the results of this project on SAT immersive platform based on Mozilla Hubs. In this case, the content would be accessible on the web, and could also be broadcasted in multiple immersive spaces.

**Results and Impact**

Media historians are required to think in terms of conjunctions between time and space, between cultures and technologies. This three dimensional matrix plotting responds to the need. Additionally, this visualization helps detecting the presence of outliers, or seemingly abnormal data. Instead of being discarded, they present an opportunity for further investigations into media content that remains separate from the mass. Rather than focusing only on mapping existing norms, recurrent features, interconnections and hubs in the aggregate, the matrix visual model restitutes meaning to the empty areas that might be the result of biases in the curation of the database or in the digital accessibility to certain technologies in a given country that generated a lack of production. In the evaluation of its results and impact, this visualization proved indeed to be a documentaristic effort that goes beyond the curatorial process of creating a dataset, beyond even the framing and defining design process required in data visualization, to ultimately assign a meaning to the missing points in our datasets and media cultures. In order to truly give sense to a data visualization that shows both an existing and a missing dataset, I wanted to physically bring the users into these empty zones between points by creating an immersive experience. The immersive environment that will be launched at SAT invites users to explore the data points
while also imagining data projections on a grid of potential values. Although the experience has not yet been opened to the public, we can gather some feedback from the initial prototyping. Following the presentation of the project covered by the Media Studies Grant at the FIAT/IFTA Conference (October 2021), the feedback received were positive across communities of media historians, archivists and curators. Graduate level students also showed their interest in the tool as both a way of learning and doing archival research. Among others, several questions revolved around the possibility to include in the visualization both a distant and a close reading. While it won’t be possible to include the complete media content and material in the visualization, I will consider the possibility to include a trailer connected to each data point so to stimulate a close analysis of the content, its aesthetic traits, tone and language. This feature will be easier to assimilate if we were to replicate this visualization based on an audiovisual archive of film, television, and/or radio material, which could be included in the hover animation.

Future Plans & Sustainability

*Immersive Datascapes (and other modes of knowing)* is very much an ongoing research project. In a way, it is intended to remain open, so to include new iterations and experiments focusing around the use of 3d and VR techniques for data visualization. As the project is still evolving, it will received additional funding for the immersive development thanks to a residency at Zù Atrium in Montreal. Here, I would like to outline the vision that will guide the project in its next phase of development. As I mentioned, the main scope that guides this project is to propose new modes of designing and exploring knowledge, while also promoting and facilitating immersive learning through the adaptive navigation of audiovisual archives and databases. After the soft-launch, I will work to turn the immersive experience into a participatory experiment of co-curation and co-creation. This include an artistic on-site performance where participants are ask to either propose existing content to add to the dataset or to imagine content that doesn’t exist yet for the missing data. The overall idea behind this Media Studies Grant project is to introduce new ways of doing media research based on digital collections. So I am open to explore options to replicate this immersive prototype based on archives and datasets other than MIT Docubase. Finally, this project is an invitation to redefine the historical value of archival audiovisual material in relation to both knowns and unknowns, presence and absence.
Appendix

Fig. 1. Network visualization of MIT Docubase on white background (series of screenshots of interactive zoom in, zoom out, resized nodes).

Fig. 2. Network visualization of MIT Docubase on white and black background (series of screenshots of transition from back to white).
Fig. 3. Network visualization of MIT Docubase on black background (series of screenshots of distant reading of the entire network, with interaction with the nodes).
Fig. 4. Network visualization of MIT Docubase on black background (screenshot of distant reading of the entire network, with resizing of the nodes and interaction).

Fig. 5. Network visualization of MIT Docubase on black background (screenshot of distant reading of the entire network).
Fig. 6. Matrix visualization of MIT Docubase on black background (series of screenshots in black and white).

Fig. 7. Matrix visualization of MIT Docubase on black background (series of screenshots showing different distributions and shapes).
Fig. 8. Matrix visualization of MIT Docubase on black background (series of screenshots showing shape retained).

Fig. 9. Matrix visualization of MIT Docubase on black background (series of screenshots showing different colormaps).
Fig. 10. Viridis colormap used for the matrix visualization of MIT Docubase.
Fig. 11. Matrix visualization of MIT Docubase on black background (screenshot showing axis labels).
Fig. 12. Matrix visualization of MIT Docubase on black background (screenshot showing axis labels in inverted order).
Fig. 13. Simplified interactive matrix visualization of MIT Docubase on black background (screenshot from the adaptive navigation).

Fig. 14. Simplified interactive matrix visualization of MIT Docubase on black background (screenshot from the adaptive navigation, zoom in and hover animation).
Fig. 15. Simplified interactive matrix visualization of MIT Docubase on black background (screenshot from the adaptive navigation, zoom out and hover animation).

Fig. 16. Simplified interactive matrix visualization of MIT Docubase on black background (screenshot from the adaptive navigation, zoom out and hover animation).
Video 1. Simplified interactive matrix visualization of MIT Docubase on black background (video recording from the adaptive navigation, zoom out and hover animation).