NAHR: A Visual Representation of Social Networks as Support for Art History Research

Houda Lamqaddam
KU leuven
Leuven, Belgium
houda.lamqaddam@kuleuven.be

Jan Aerts
KU Leuven
Leuven, Belgium
jan.aerts@kuleuven.be

Koenraad Brosens
KU Leuven
Leuven, Belgium
koen.brosens@kuleuven.be

Katrien Verbert
KU Leuven
Leuven, Belgium
katrien.verbert@kuleuven.be

ABSTRACT

In the field of art history, the analysis of community dynamics can give researchers precious insight on the subjects of their studies. Data on genealogy and community bonds can provide a rich understanding of the functioning of a community. Traditional family trees are not designed to support extra-familial links and often lack the time-bound aspect of these relationships, and timeline-style tools miss the mark on representing the network dimension of such structures. We introduce NAHR (Networks in Art History Research), a tool that visualizes small networks of families connected through marriage, god-parenthood and professional relationships, and that provides insight in the change of these dynamics over time.

Author Keywords

Visualization; digital humanities; art history; user-centered design; network visualization; genealogical data.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

For art historians, the analysis of personal and professional relationships can help understand the dynamics of a community. Researchers often have to reconstruct artists' trajectories through the data they have access to, describing their personal lives (parents' identities, spouses, children) as well as their artistic or professional collaborations (pupils, business partners, etc.). The access to a comprehensive representation of such data is therefore essential in validating researchers' hypotheses, answering existing research questions and triggering new ones. Because the described connections mostly revolve around family structures, we consider the literature surrounding genealogical data representation as a base for our work. We find that, although common, the traditional "family tree" visualization presents numerous shortcomings when attempting to represent complex familial stories.

Our contribution is a novel approach to the issue of representing changing complex community dynamics through time in the aim of supporting art history research. While this paper presents the designed and developed proof of concept, a live version is being made available for our collaborators. This will allow them to search the database for any stored person and visualize the networks of their family and relationship links. By designing the visualization to address the actual needs of art historians, we also demonstrate how visualizations need to be tailored for use in the digital humanities field.

RELATED WORK

The representation of genealogical data has been a topic of interest in many fields including history, sociology, and visualization. Several tools aim at tackling the difficulties of representing such complex, hierarchical and time-dependent data.

A few constraints have been described in the literature. The first is the difficulty of integrating temporal information in tree-like structures, and hierarchical genealogical information in timelines. Scalability is often cited [5] as another limit to visualizing family trees. The more community members a tool aims at representing, the higher the likelihood of clutter, and the more readability is impeded by edge-crossing. Moreover, family trees often do not depict full ancestry links, but focus on only one parent lineage to conserve the tree format. Traditional family tree representations also do not generally enable the portrayal of complex community and family links such as divorce, remarriages, and out-of-wedlock births.

Finally, the contemporaneity of persons and relationships is another critical element to understanding community relationships. The representation of this aspect is often overlooked.

NordiCHIÍ8, September 29-October 3, 2018, Oslo, Norway

DOI: https://doi.org/10.1145/3240167.3240199

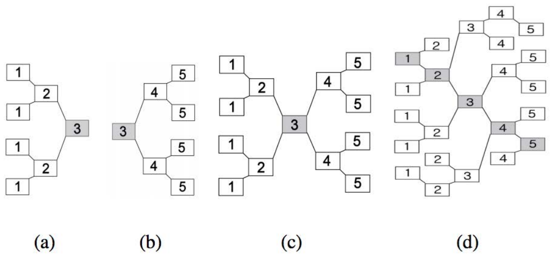

Figure 1. (a), (b), and (c) represent the ancestry, descendence and ego-centric genealogy representations. (d) was introduced by McGuffin as an alternative method to represent similar data.

This section defines two broad genres of visualizations as described in literature: some based on the family structure and relations between individuals, and others on the temporal aspect of the data. A third category that attempts to give equal importance to these two aspects is also described.

Family Tree based representations

One of the simplest representations of genealogical data is through trees and tree-like encodings (e.g. radial, treemap) [5]. These graphs represent the individuals as nodes and relationships as links, and are often based on the family structure layout [8] [5]. Three main use-cases - and therefore representations - have been described as scenarios for tree visualizations of genealogical data: ancestry, descendance, and ego-centric research [9] (see Figure 1). These representations present a clear view of family links and relationships, but they have important setbacks. First, they generally fail in including the temporal aspect of the data, that is critical in showing contemporaneity, which in turn is important in assisting the task of analyzing community dynamics. An additional constraint is the limited scalability of this type of graphs, as edge crossing grows rapidly when the number of individuals represented increases.

Timeline based representations

Other tools for genealogical data representations were designed to center the time aspect of the datasets rather than the family structures. One of the first examples of a family tree representation belongs in this group. The "Specimen of a Chart of Biography" [11] takes the shape of a timeline, where individual actors are represented by horizontal bars, starting in the year they were born and ending on their death year (see Figure 2). Since then, multiple tools with similar structures have been introduced, including in commercial software. Prominent examples are the Genelines interface from Progeny Genealogy [1] and the TimeLine view of the MacFamilyTree [13], which both base their representations heavily on time, and encode family relationships through annotations.

Combined family tree and timeline based representations

Some works have attempted to bridge the gap between the previous two categories. Rundall Munroe's "Movie narrative charts" [10] and Kim's TimeNets [6] are, for instance, based on timeline representations in which individuals' connections are shown through individual timeline movement. To tackle larger scale data, Reda et al's community representation tool also

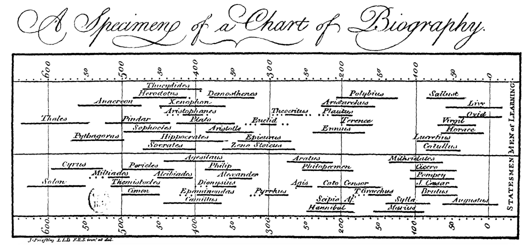

Figure 2. A specimen of Priestley's Chart of Biography which represented individual timelines over a six hundred year span.

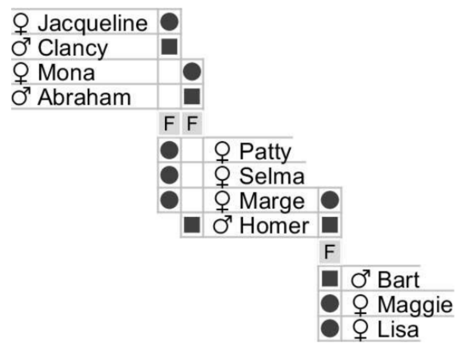

Figure 3. The Simpsons family structure represented with GeneaQuilts.

resembles a timeline where the nodes converge and diverge to form groups and community structures [12]. Zhao's EgoLines is another approach where a subway metaphor is used to represent ego-centric collaboration networks through time [14]. Specifically concerning genealogical trees, GeneaQuilts [2] presents a system for the exploration of large genealogy trees through a synthetic matrix representation (see Figure 3).

While these tools have addressed the issues of scalability and time-bound genealogical data representation, we have found that for art historians, the type of data and variety of tasks at hand require more adapted visualization tools. In this context, constraints such as sparse datasets, small sizes of populations of interest, and the strong dependency on the time dimension of each data point have created the need for a new approach to representing genealogical networks.

With NAHR, we contribute a new approach for the representation of complex genealogical and extra-familial relationships that accommodates different types of links, while centering their time-boundedness. This tool combines the familiarity of family tree visualizations to represent personal links with an incorporated way to explore the time dimension and its impact on the dynamics of a community.

DATA, USERS, AND TASKS

Methodology

The initial phase consisted of gathering user needs by conducting interviews with five expert users, including two art history professors and three PhD students. We also designed a survey that was distributed and filled by 11 students and professors of art history. These steps allowed us to derive user profiles, task descriptions and a better understanding of the data usage.

The following steps consisted of iterations between design of low-resolution paper prototypes and feedback from users. The last step was the delivery of an implemented proof of concept of the final prototype. The next sections describe the nature of the data represented, characteristics on the users, as well as a list of the extracted tasks.

Data

The data consists of genealogical trees with additional god-parenthood and professional relationships among persons. These extended *family* trees is what we refer to as communities. They are built around **persons** (also referred to as actors), that are each described with a name, dates of birth and death, gender and - in some cases - profession. Each actor is connected to other actors through 3 possible types of **links**.

- Marriage and descendance: each actor has a link to their parents, as well as potential spouse(s) and offspring.
- God-parenthood links: god-parenthood is an important marker of social links in the described social context. It allowed families to build connections to one another, or strengthen existing ones. Moreover, godparents can end up becoming masters for their godchildren or influencing their style, they are therefore critical to identify. The data reflects this, as each actor has a link to their godmother and godfather, if known.
- Professional links: few actors have known professions. Some of them, however, are known to have collaborated. Collaborations can be of different types, but our current dataset only contains joint ventures.

The data also includes time-bound aspects. Births and death dates, marriages, as well as beginnings and ends of professional collaborations are denoted as **events**.

The database that was assembled for this project consists of thousands of actors. For this project, however, we were given a small subset consisting of 32 actors spread over 3 generations. This dataset was selected by our users because it was the density and size of a typical community size that they would be interested in visualizing (see User Characterisation). The focus was therefore on building an interesting approach to representing a dataset of this nature.

User characterisation

Two broad categories of user backgrounds are targeted by this tool. The first and main one is the art history research community. This interface is made as a proof of concept to introduce an effective visualization to explore community dynamics. It is meant to present a usable tool allowing to explore the data and learn new aspects of the dynamics between actors, while at the same time acting as a sample project show-casing the added value of integrating visualization to the data analysis process in art history. The second category is a broader non-expert audience: the tool is also designed to attract funding for the project, and interest for the digital humanities field in general. In that aspect, the interface should be viewed as usable and *fun* by non-domain users.

A look at the literature on art history research, interviews and a survey ran on 11 art history researchers allowed us to make the following data usage-related observations:

- For most of the participants, the number of subjects studied at once was quite low. This was also addressed by users in interviews, as large dataset representations or full database visualizations were expressly discouraged. One user noted that "Many people from computer science make these tools for us where they want to show all the data that we have in one big graph, but that's simply not how we work". It seems indeed that researchers in the field of art history are more likely to focus on finding out more about a single artist, or a selected few, and a specific era. For 8 out of 11 persons surveyed, we found that a tool allowing to visualize groups of less than 50 persons at once was sufficient.
- The interest - and acceptance - for digital tools are still generally low. The literature confirmed what we found in our own research, which is that the interest in visualization tools seems to be a relatively new trend in the field of art history. In 2005, the Summit on Digital Tools at the University of Virginia found that only about six percent of humanist scholars "go beyond general purpose information technology and use digital resources and more complex digital tools in their scholarship." General purpose information technology here referring to tools for word processing, presentations, spreadsheets and conferencing [3]. Later, in 2012, Gibbs & Owens find that digital tools are still "a fringe element" in the humanities field. Although the interest seemed higher, a lack of usability in the existing tools was pointed out [4]. According to our survey and interviews, while most participants reported being already familiar with charting tools, the most cited one was the spreadsheet Excel. Its ease of use for chart creation was cited as a positive point, although it is clear it is not enough to produce adapted interactive visuals. This tells us that familiarity and acceptance might be critical aspects to cater for in any new tool we would like to introduce.
- According to our users, most of the current art history research workflow is carried offline. As an example to illustrate this, our main users work on physical archives from the national libraries of the cities of Brussels and Antwerp in Belgium. The process of retrieving data, therefore, involves tracking down records, transcribing information, and archiving said data in a local format for future referencing. For this reason, most participants expressed interest in incorporating a digital visualization tool to their data analysis workflow. While this seems to contradict the previous point at first sight, it appears to be a consensus that the field of digital humanities is "as buoyed by optimism as it is laden with skepticism" [7].

A tool that can bring added features to familiar representations of the data can therefore be of great value to the community. We believe that a tool that is designed in a user-centered approach and that centers the tasks performed by researchers has the potential to overcome the skepticism caused by a lack of usability and the use of unadapted generic tools.

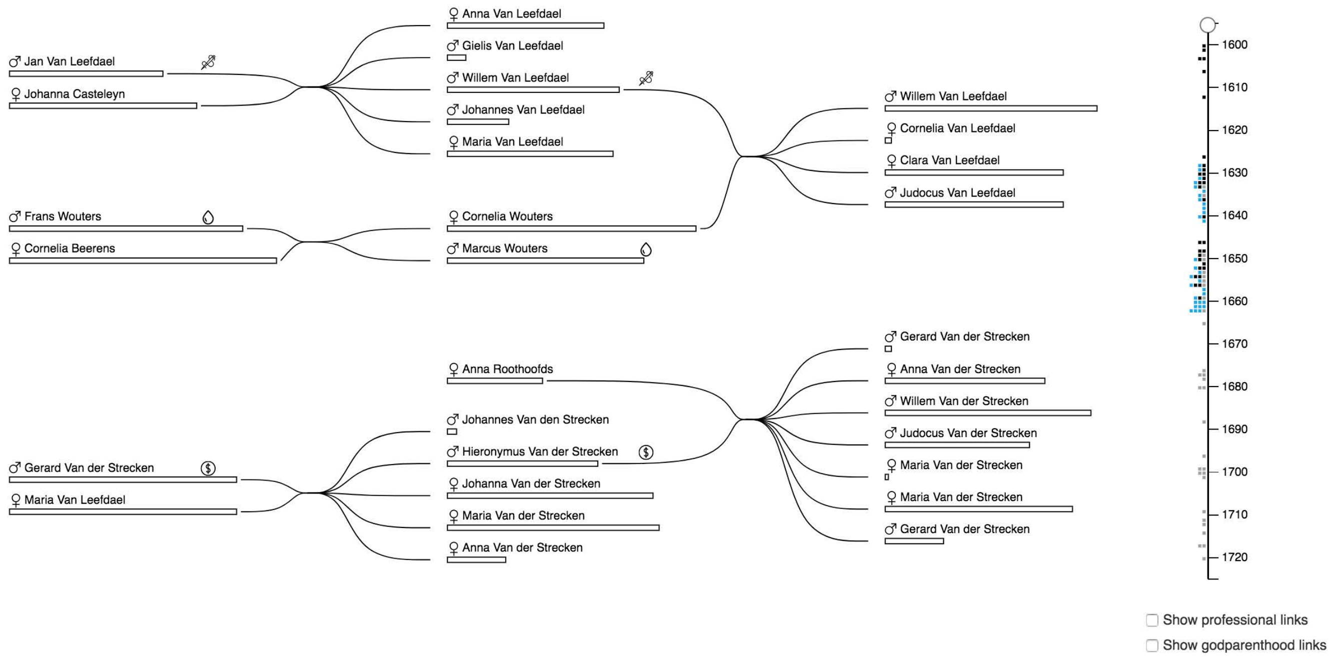

Figure 4. Main visualization before interaction.

Tasks

The conducted interviews and survey allowed us to extract three categories of tasks with which NAHR can assist. First, actor-related tasks are the ones centered around gaining knowledge on actors and relationships. Chronology-related tasks involve being able to navigate in the time dimension. Finally, in the research-focused tasks, we list the tasks that are related to the users' work-flow rather than the data. Overall, we established that users focused on analyzing specific subsets of persons and relationships, rather than the whole dataset. This allowed us to focus on detail views and interactions instead of designing larger overview visualizations.

Actor-related tasks

- (T1) *Viewing an actor in the context of their community relationships*: Users are interested in seeing one or more actors as they interact with their community in terms of marriage and descendance, as well as god-parenthood, and professional collaborations. The system should show all these relationships and make it easy for the user to distinguish between their nature, as well as choose which type of relationships to view or dismiss.

- (T2) *Understanding passed actors legacy*: Family links are critical community structures and therefore should be clearly exposed. The family bonds should be visible even after the members' deaths as their place in the community remains existing and their impact may stay valid.

- (T3) *Estimating the relationship density of an actor*: An actor's influence on the community may be estimated from the number of links they have to other members. As this changes through time, a user can have an idea of the change in dynamics around a user by looking at how much the density of their ties varies. The system should allow a user to view that information over time, as well as in a specific point in time.

Chronology-related tasks

- (T4) *Freezing the story in time:* Users are sometimes interested in looking at the ties and dynamics between members of a community frozen in a specific year. The system should allow them to explore the ties in a chosen point in time.

- (T5) *Focusing on the time periods with most activity:* Some periods in time witness more events than others. The system should assist the users in knowing the years with most changes in dynamics.

Research-focused tasks

- (T6) *Having access to references of evidence of presented data*: Users who want to pursue their research on a particular actor, or investigate the validity of the data displayed in the visualization, should be able to access a reference to the origin of this information. The system should present users with evidence of the data and direct them to its initial sources.

- (T7) *Saving and sharing an image of a community in a point in time*: Users work in collaboration to analyze their datasets. They want to be able to share a specific configuration of community links at a point in time to colleagues in order to share an insight or discovery. The system should allow them to save "images" of the current configuration they are looking at to save for future retrieval or to forward to one another.

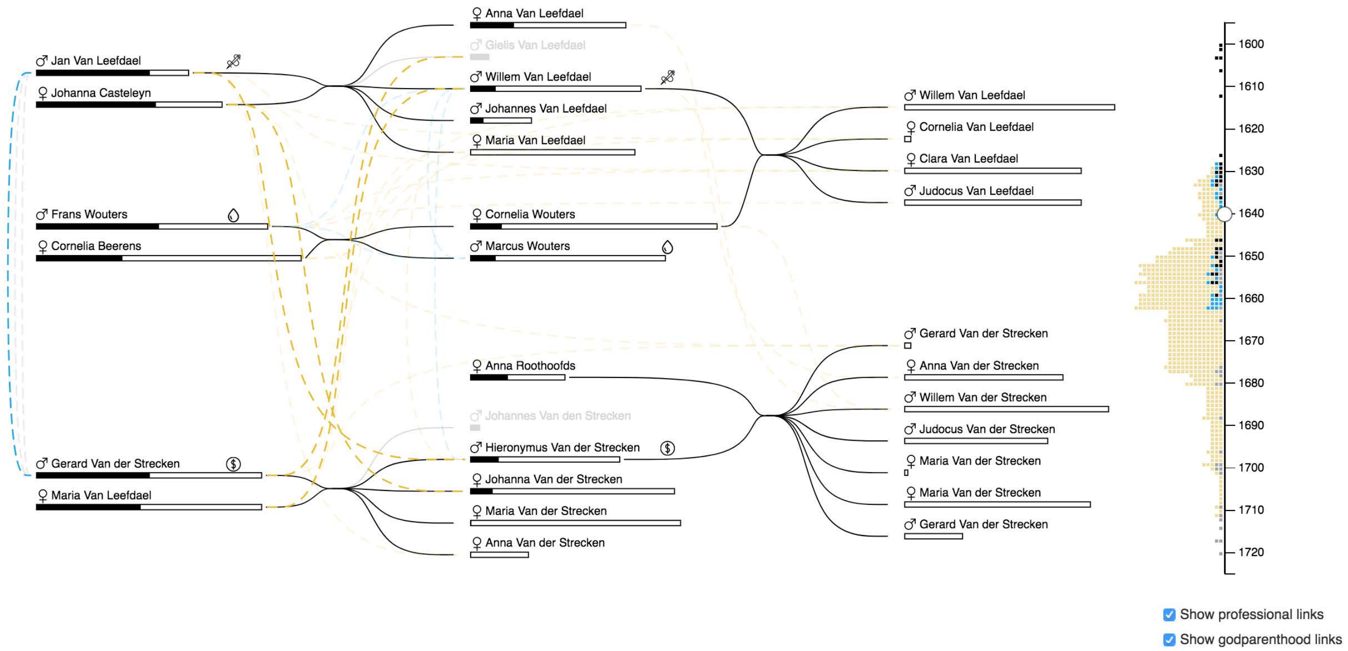

Figure 5. Main visualization after year selection and links enabling.

DESIGN

The main interface consists of a graph visualization describing a family structure where the nodes are actors in the dataset, and links are ancestry relationships. Figure 4 shows the visualization before any interaction, when only the community structure is drawn. Figure 5 shows the tool after a desired year was selected, and the links were enabled, showing only what the community links were like in said year.

In this part, we will define how families are represented, how the time aspect is sewn into the tree view, and what is encoded in the nodes and links of the graph. Each task being supported is referred to at the description of the feature supporting it with the notation (T*i*), *i* being the task number defined in Tasks.

The final subsection contains a reading of the graph in Figures 4 and 5 to illustrate how different features work together to tell the story of the dataset.

Structure

Graph Structure

The genealogical aspect of the data being central to our users' research, we chose to address T1 by representing the data in a common family tree mapping to leverage the familiarity this representation has. This decision was validated after an initial family tree representation of the dataset was found to be easily readable, albeit not sufficient to capture all information included in the data.

Parents of the first generation available in the dataset are drawn one below the other in separate regions denoting separate families. Each generation is aligned on the *y* dimension, and siblings are vertically ordered in birth date chronological order.

The decision to use a generational view of the family links instead of a representation aligned on an absolute view of time was a preference indicated by users, as generations are important tools to analyze the data according to our interviews.

Finally, the family tree was drawn horizontally instead of vertically to allow for a readable node width across generations, we found that this did not impede readability.

Time

We chose to use the *x* and *y* dimensions for the network aspect of the data only in order to ease graph comprehension. Therefore, an extra dimension was needed to include the time aspect. For that purpose, a vertical slider was added, allowing the user the following two tasks:

- Selecting a current year: dragging the handle to a specific year updates the nodes and links to represent the current state of persons and relationships in that year (T4).
- Providing an overview of event frequency through the years: The year selection slider also acts as a bar chart where each square represents an event happening in the corresponding year. The squares are color-coded in the same way as the links (Figure 8) to ensure differentiability between different types of events, namely births and deaths, professional cooperations and god-parenthood links (T5).

Nodes

A person is represented by an annotated bar - or life bar - containing the information pertaining to them. Each bar is a rectangle of fixed height, where the width is proportional to the lifespan of the person. For each selected year in time, an actor's life bar will proportionally fill to indicate their current age. It will be grayed out after death to reduce its saliency (see

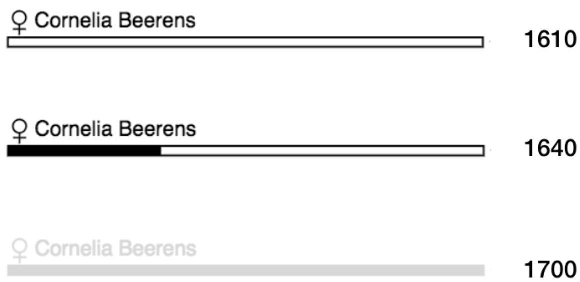

Figure 6. Node representation before an actor's birth, during their life and after their death.

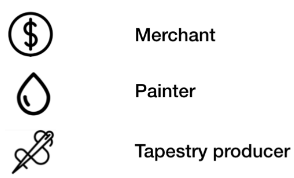

Figure 7. Signification of profession-related icons used in the visualization.[2]

Figure 6). This allows the family to retain all its members, alive, future, or deceased, while giving saliency to the current family structure only (T2).

The life bar is annotated with the name of the actor, a symbol representing their gender, and an icon representing their profession, if known. Figure 7 contains a legend of the used symbols.

More information about a person is also visible upon hovering (see Figure 9 (a)). Once the reference documents used to build the database are fully digitized, this tool-tip will also contain links to view to said documents (T6).

Links

As described in the Data section, three types of relationships are defined in our dataset. Different link encodings represent these relationships as can be seen in Figure 8.

The family link is drawn in a typical "family tree" manner. God-parenthood and professional relationships can be toggled on or off the screen to allow for more readability, and to support different tasks. These relationships are represented with colored dashed lines to differentiate their appearance from the family ones. They are drawn using thick lines to balance the lighter color and maintain visibility.

We retain all links visible after actors pass away and collaborations end to support T2 and T3. The saturation encoding is therefore used to express the validity of a link at selected year.

As with the nodes, hovering will display a tool-tip containing additional information regarding the relationship (see Figure 9 (b)).

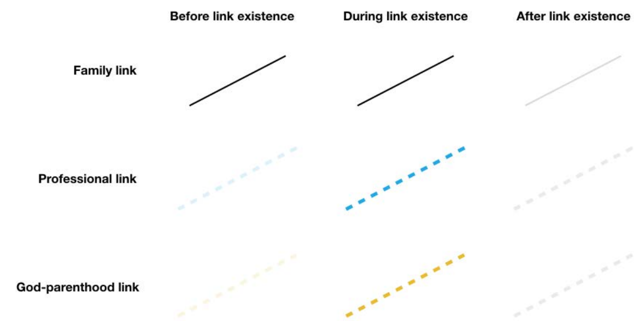

Figure 8. Representation of link states before, during and after existence, according to the type of relationship.

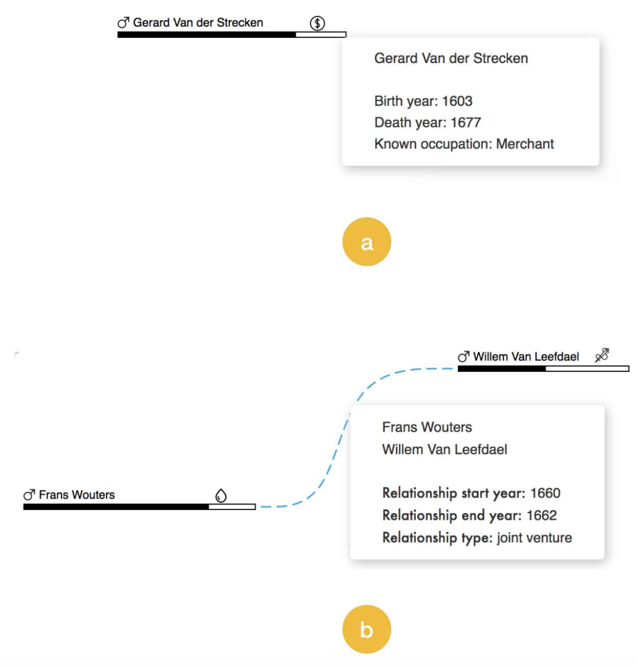

Figure 9. Hovering on a node (a) or a link (b) will display a tooltip containing additional information about the element.

Reading the graph: A walk-through

After having described the features of the tool, we suggest a walk-through of the graph in Figures 4 and 5 to present a synthetic illustration of how the elements function together in this use case.

Users are first presented with the view in Figure 4. It is a sideways family tree representing 3 families across three generations. Each of the three couples in the leftmost column are parents of a different family whose offspring are linked to them through curved black lines. A third generation is also shown and linked to its parents the same way. On the right, we can see a vertical slider where the handle is in initial position before any event is recorded. The values along its length are the years for which we have data on the persons displayed. We

[2]The icons used to represent professions are made available by their authors on flaticon.com. The merchant and tapestry producer icons were designed by Freepik, the painter icon was designed by Good Ware

can see that we have historical records on the studied persons between the years of 1600 and 1720. The year slider is also annotated with a vertical modified bar chart representing the number of events recorded for each year. We can see for instance that the first birth (in black) was recorded in 1600, and the last death (in gray) in 1720.

If they would like to see the ongoing community dynamics at a specific point in time, the user only needs to slide the handle to the chosen year, this will update the filling of the bars representing each person in the graph to reflect their current age. Figure 5 shows the shape of this community in the year 1640. We can see that the first generation is around the middle of their lives and have already had their first children. Two of which (Gielis Van Leefdael and Johannes Van der Strecken) have been grayed out, meaning they have already passed away after a short time alive. The rest of the persons have not been born yet at this time. Hovering on any of these persons will indicate more specific information about birth year, death year, and profession.

If we toggle the professional and god-parenthood links, we can also see the extra-familial dynamics happening at the same time. We can first see that there were several joint ventures happening between Jan Van Leefdael and Maria Van Leefdael's husband: Gerard Van der Strecken, one of which was taking place during 1640. We can also have insight as to the two families turning to one another for god-parenthood roles for their children. In the background, we can see future links that will gain in visibility when the equivalent time point is selected.

RECEPTION

The tool was first introduced to our five expert users for feedback, where it was used for a couple of weeks before comments were heard. It was later presented in art history symposia [3] and workshops to audiences of art historians specifically, as well as broader publics of humanities researchers, where it was received as an innovative and easy-to-use solution. More feedback received about specifics of the tool is described below.

- *Structure representation:* Most of the users were presented the tool after having first seen the same data in a traditional family tree representation, then using a force directed layout. They reported the genealogy structure in our tool was easy to understand, as it uses the codes of a traditional family tree. The users appreciated the ability to scroll through time and view concurrent events and links through interaction with the tool, as well as the overview provided by the event frequency view embedded in the slider.

- *Nodes and links representation:*

  The representation of links was well received. Family structure and professional links gave a clear idea of the community connections. God-parenthood links were thought to still have the potential to be confusing as they are more prevalent in the dataset, therefore are more likely to result in clutter.

  The display of the dynamics around three generations in one frame was seen as a positive aspect, although the need was raised to allow following these family trees further in the past or the future.

[3] Venues the tool was presented in include the Baroque Tapestry and the Rome of the Barberini Symposium (USA, 2017), the Annual Meeting of the Renaissance Society of America Symposium (USA, 2018), and the Equip & Engage: Research and Dissemination Infrastructures for the Humanities Symposium (Belgium, 2018).

CONCLUSION AND FUTURE WORKS

In this project, we developed a visualization tool aimed at representing small networks of extended families and communities for art history researchers. We found that researchers in art history needed small scaled, highly adapted systems rather than large scale data exhaustive visualizations. In the process of developing the NAHR interface, we found that classic family tree representations were in themselves not sufficient to support the users' tasks, but that a redesigned interactive family-tree based visualization that added time as a third dimension was successful in supporting user tasks and earning initial acceptance.

This tool is to be integrated into an online database containing all the actors found in our users archives, where users will be able to search for specific persons and visualize their family trees. Next steps will therefore be to develop ego-centered views built around each actor to allow users to explore their networks. We are hoping this will enable our users to have a seamless experience navigating back and forth between written descriptions of actors and the visual representation of the same data on NAHR. We also want to investigate collaboration aspects and implement features that can address T7.

Because our users are only interested in studying a limited number of actors at once, the scalability factor was not addressed in the design we presented. However, were that use-case to evolve, visual and interaction scalability will have to be addressed.

Future work also includes formal user testing to compare NAHR with other existing methods and improve its design.

ACKNOWLEDGEMENTS

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

REFERENCES

1. 2018. Genelines Sample Charts. (2018). `http://progenygenealogy.com/`

2. A. Bezerianos, P. Dragicevic, J. D. Fekete, J. Bae, and B. Watson. 2010. GeneaQuilts: A System for Exploring Large Genealogies. *IEEE Transactions on Visualization and Computer Graphics* 16, 6 (Nov 2010), 1073–1081. DOI:`http://dx.doi.org/10.1109/TVCG.2010.159`

3. Bernie Frischer, J Unsworth, A Dwyer, A Jones, L Lancaster, G Rockwell, R Rosenzweig, and S Wells. 2006. Summit on digital tools for the humanities: Report on summit accomplishments. *Charlottesville, VA* 23 (2006).

4. Fred Gibbs and Trevor Owens. 2012. Building Better Digital Humanities Tools: Toward broader audiences and user-centered designs. 6 (01 2012).
5. Kerstin Keller, Prahalika Reddy, and Shimul Sachdeva. 2010. Family Tree Visualization. *Vis.Berkeley.Edu* (2010). https://pdfs.semanticscholar.org/bbe7/05676937445f2c434e8295cefbf4758ad8c2.pdfhttp://vis.berkeley.edu/courses/cs294-10-sp10/wiki/images/f/f2/Family_Tree_Visualization_-_Final_Paper.pdf
6. Nam Kim, Stuart Card, and Jeffrey Heer. 2010. Tracing genealogical data with TimeNets. (01 2010).
7. Julie T Klein. 2015. *Interdisciplining digital humanities: Boundary work in an emerging field*. University of Michigan Press.
8. Tiancheng Li, Elton S SIQUEIRA, Patrick Kabongo, and Li Weigang. 2016. On Chinese and Western Family Trees: Mechanism and Performance. 5 (01 2016), 11.
9. Michael J. McGuffin and Ravin Balakrishnan. 2005. Interactive visualization of genealogical graphs. *Proceedings - IEEE Symposium on Information Visualization, INFO VIS* (2005), 17–24. DOI: http://dx.doi.org/10.1109/INFVIS.2005.1532124
10. "Munroe". 2009. "Movie Narrative Charts". XKCD. (2009). https://xkcd.com/657/
11. Priestley. 1765. A Chart of Biography. Warrington Press. (1765).
12. Khairi Reda, Chayant Tantipathananandh, Andrew Johnson, Jason Leigh, and Tanya Berger-Wolf. 2011. Visualizing the Evolution of Community Structures in Dynamic Social Networks. 30, 3 (2011). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.462.8150&rep=rep1&type=pdf
13. "Synium Software". 2018. "MacFamily Tree". (2018). http://www.syniumsoftware.com/macfamilytree
14. Jian Zhao, Michael Glueck, Fanny Chevalier, Yanhong Wu, Azam Khan, J Zhao, M Glueck, F Chevalier, Y Wu, and A Kahn. 2008. Egocentric Analysis of Dynamic Networks with EgoLines. (2008). https://hal.inria.fr/hal-01325788/document