



# HOMELESS PAINTINGS OF THE ITALIAN RENAISSANCE

CS171 FINAL PROJECT  
PROCESS BOOK

Ben Zauzmer  
Linda Song  
Jason Mahr

# TABLE OF CONTENTS

- Initial Project Proposal
- Overview and Motivation: Provide an overview of the project goals and the motivation for it. Consider that this will be read by people who did not see your project proposal.
- Related Work: Anything that inspired you, such as a paper, a web site, visualizations we discussed in class, etc.
- Questions: What questions are you trying to answer? How did these questions evolve over the course of the project? What new questions did you consider in the course of your analysis?
- Data: Source, scraping method, cleanup, etc.
- Exploratory Data Analysis: What visualizations did you use to initially look at your data? What insights did you gain? How did these insights inform your design?
- Design Evolution: What are the different visualizations you considered? Justify the design decisions you made using the perceptual and design principles you learned in the course.
- Implementation: Describe the intent and functionality of the interactive visualizations you implemented. Provide clear and well-referenced images showing the key design and interaction elements.
- Evaluation: What did you learn about the data by using your visualizations? How did you answer your questions? How well does your visualization work, and how could you further improve it?

# INITIAL PROJECT PROPOSAL: OVERVIEW & MOTIVATION

Ben is taking a class called “Homeless Paintings of the Italian Renaissance.” The goal of the class is to explore a dataset of photos of Italian Renaissance paintings that have been lost to history. The original paintings may have been lost, stolen, destroyed in World War II, mislabeled, or hidden in a private collection. Ideally, we will also find a few of the paintings along the way, though the main goal of this project will be to do cool data visualization projects.

Linda is majoring in CS and minoring in History of Art and Architecture, so she is looking forward to combining her twin interests with this unique and previously unexplored data set. Jason is majoring in CS and is looking to explore how data can be visualized in a meaningful and aesthetically pleasing manner through this project.

# INITIAL PROJECT PROPOSAL: PROJECT OBJECTIVES

The main goals will be to let users learn more about the distribution of the paintings in the set across many different axes: time, painting location, last known location, artist, last known date, etc. The benefits include a general knowledge of art historical collections, and perhaps insights into finding some of the lost paintings of this collection.

# INITIAL PROJECT PROPOSAL: DATA

The data is provided by the Homeless Paintings course, originally encoded by Bernard Berenson's archive in Italy (in a villa owned by Harvard). The data is split into 11,233 separate json files – one for each painting – which unfortunately are not exactly uniform so some cleaning will be required. The files are online but password protected, but we can get a link with permissions for our TF if necessary.

# INITIAL PROJECT PROPOSAL: DATA PROCESSING

Mostly, the json files are in two different formats, so we will run a loop through the files to determine the format and then run the appropriate functions to extract data as necessary. We will take dates, locations, artists, and anything else that might be interesting to explore. We expect to use Python to scrape the files (since they are only online), and then write the information into a CSV.

# INITIAL PROJECT PROPOSAL: VISUALIZATION

For the dates that paintings were created and then lost, it would be great to create bar charts with brushing. For the locations that paintings were created in and then lost in, we are considering making maps of the world with the paintings plotted on top. For both of these, we would like to have interactive filtering so the user doesn't have to use the full dataset for each.

# INITIAL PROJECT PROPOSAL: MUST-HAVE & OPTIONAL FEATURES

## Must-have features:

- We need to have visualizations that cover every aspect of the metadata: date, artist, location, etc. Bar charts, line graphs, maps, and 2D plots are all viable options.

## Optional features:

- There is also potential to do analyses based on the paintings themselves, using aspects such as RGB, saturation, hue, brightness, etc.

# INITIAL PROJECT PROPOSAL: SCHEDULE

Date	Name	Assignment
3/23	Ben	Scrape and clean data
3/30	Ben, Linda, Jason	Each create one visualization
4/6	Ben, Linda, Jason	Each create a second visualization
4/10	Ben, Linda, Jason	Edit and improve each other's work

# OVERVIEW AND MOTIVATION

- As stated in our initial project proposal, the main motivation behind creating a visualization of this data comes from the massive, unorganized, dataset found and under utilized in Ben's art history class of the Italian Renaissance. Given the large dataset and the lack of organization and efficiency in handling the data, we sought to restructure the data and visualize it in a way that would make the data more accessible and comprehensible to our audience.

# RELATED WORK

- Homework 2 and Homework 4 were influential for our design as baseline templates to start visualizing the large dataset at hand.
- In addition, much inspiration was taken from mbostock's D3 visualization gallery, in particular the Bubble Chart (<http://bl.ocks.org/mbostock/4063269>) and Multitouch Drag (<http://bl.ocks.org/mbostock/9631744>), as well as nsonnad's Object Constancy with Multiple Sets of Time-Series (<http://bl.ocks.org/nsonnad/4175202>).

# QUESTIONS

- The main question that we had with our data set and from our feedback is the question of who would want to access this data and what story it would tell. Through this feedback, we added a historical background element to the visualization, allowing the data to be held up against historical facts and events that occurred in the time of the Italian Renaissance in order to convey a greater understanding of the context of when and where the paintings were created.

# DATA

- From Ben's Homeless Paintings in the Italian Renaissance class, we have a json data set of paintings that contains the date, artist, geographic location, associated topics (i.e. elements in the painting, such as Jesus, necklace, John the Baptist, etc.), hue/saturation of the photo taken of the painting (which are all in black and white), and notes scribbled on the back of the photo. This was in individual json files per painting, so we used python to scrape the data into one set to work with in the visualization.
- In addition, important events within the history of the Italian Renaissance were scraped off of Wikipedia using python to be able to directly compare the dates of the paintings to the events that could have inspired or affected artists to create art.

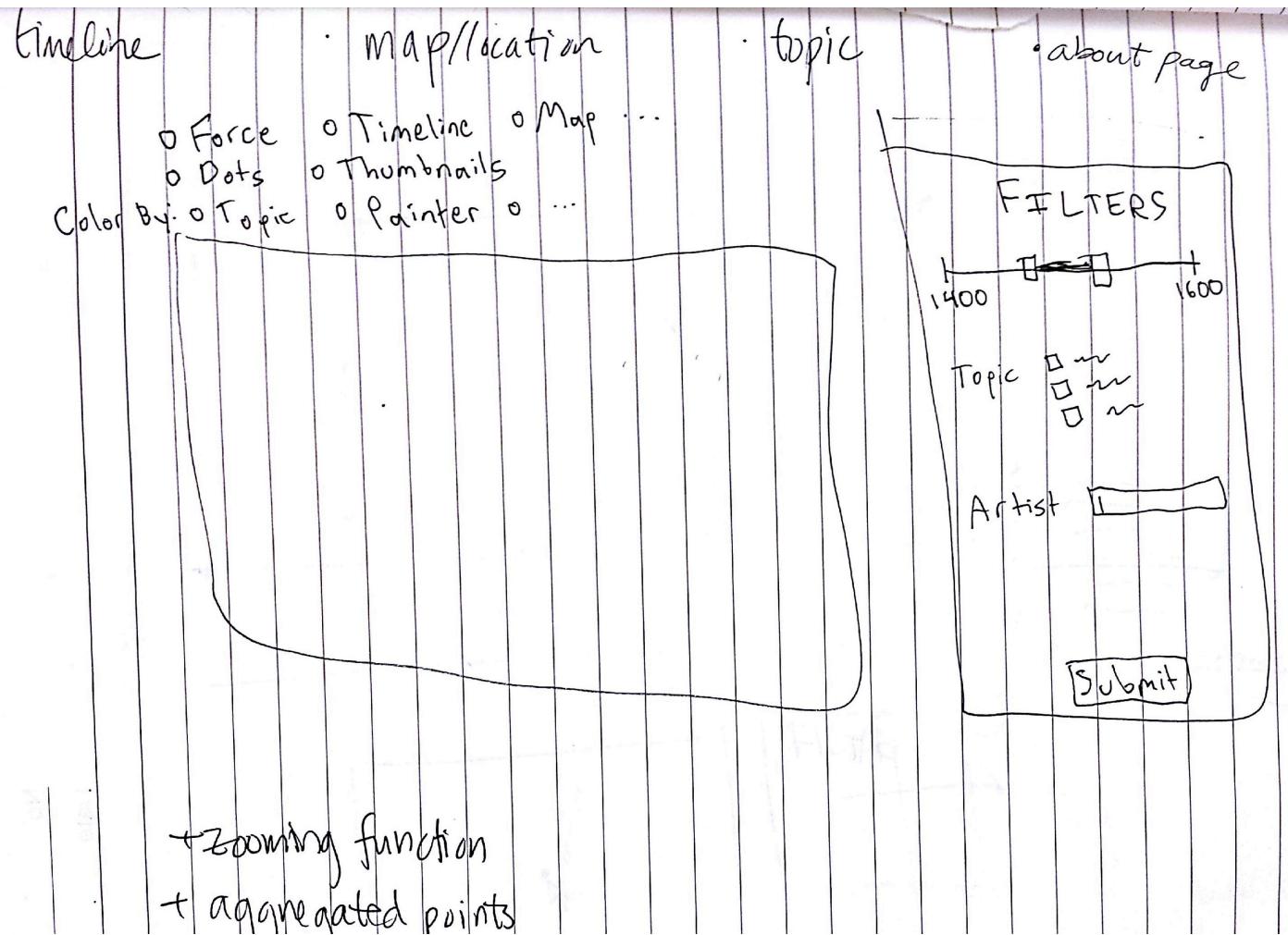
# EXPLORATORY DATA ANALYSIS

- Because we don't have much of a storyline for the visualization, much of the feedback that we got from Design Studio 4 was to create associations and correlations that could have arisen from the paintings lost/found to make the data more interesting for the average user. Consequently, there are thoughts of possibly comparing the overall history of the Italian Renaissance and its historical main points against the dates of the paintings and when they were lost to see if we can find some correlation that would explain the lost records of the homeless paintings.

# DESIGN EVOLUTION

- The following two images show the initial design sketches that would visualize the paintings in an informative way:

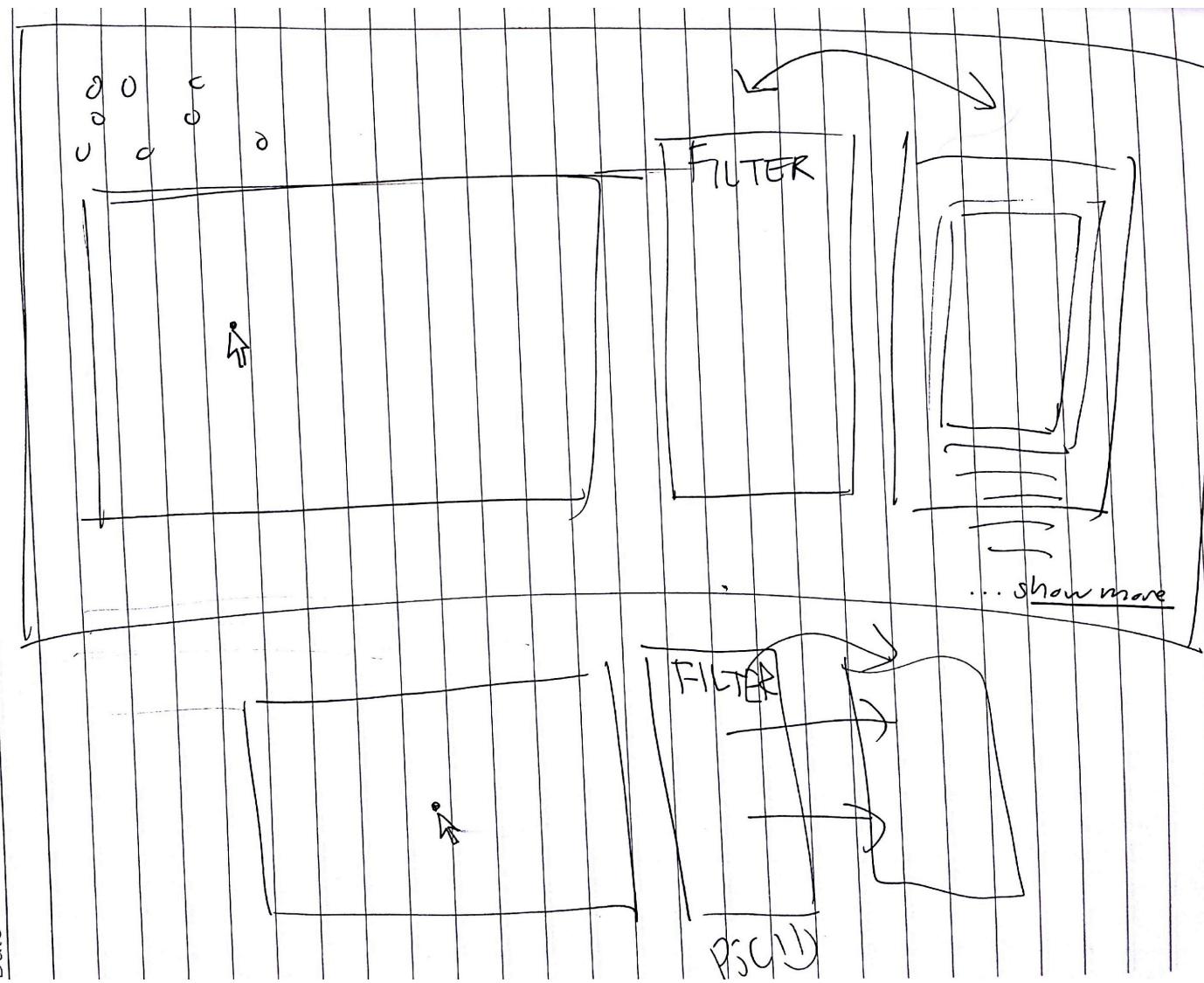
## 1<sup>ST</sup> IDEA



Initially, we based our visualization off of the second problem set where we use radio buttons to switch back and forth between different visualizations, ranging from the world map to timeline. The overlapping nodes on the map would aggregate into larger nodes to prevent cluttered usage. There would also be a filtering function where the date of the paintings shown can be adjusted along with the topic and the artist.

## 2<sup>ND</sup> IDEA

In addition to the first idea, we also thought of adding a second column next to the filtering column that would show the actual painting chosen and relevant details underneath the image. The arrows indicate the columns perhaps transitioning or changing location, but we aren't exactly certain what would be the best way to arrange the filter/detail/map yet.

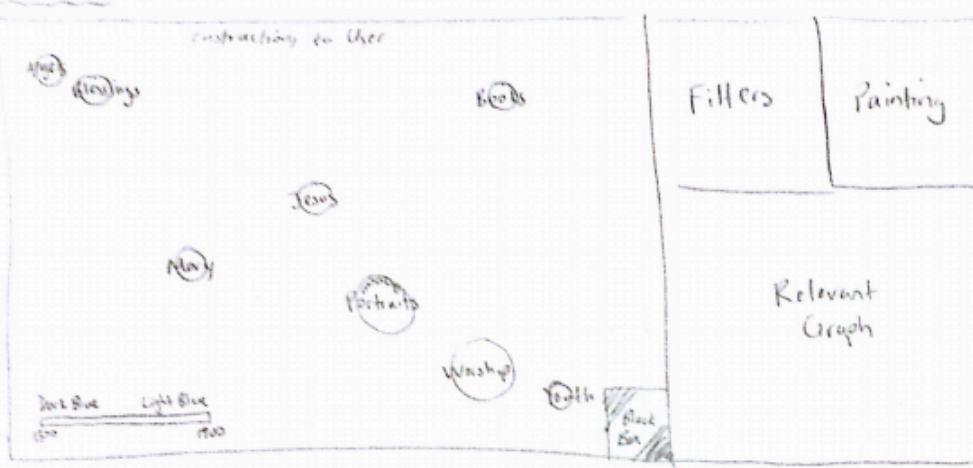


# WORD CLOUD

The ideal implementation of the word cloud would show the radius of the topic circle to correspond with the number of paintings with that topic. In addition, the circles would have a color based on the average year that the topic appears. The circles would be movable for easy comparison and will hopefully be able to be sorted by average year or frequency.

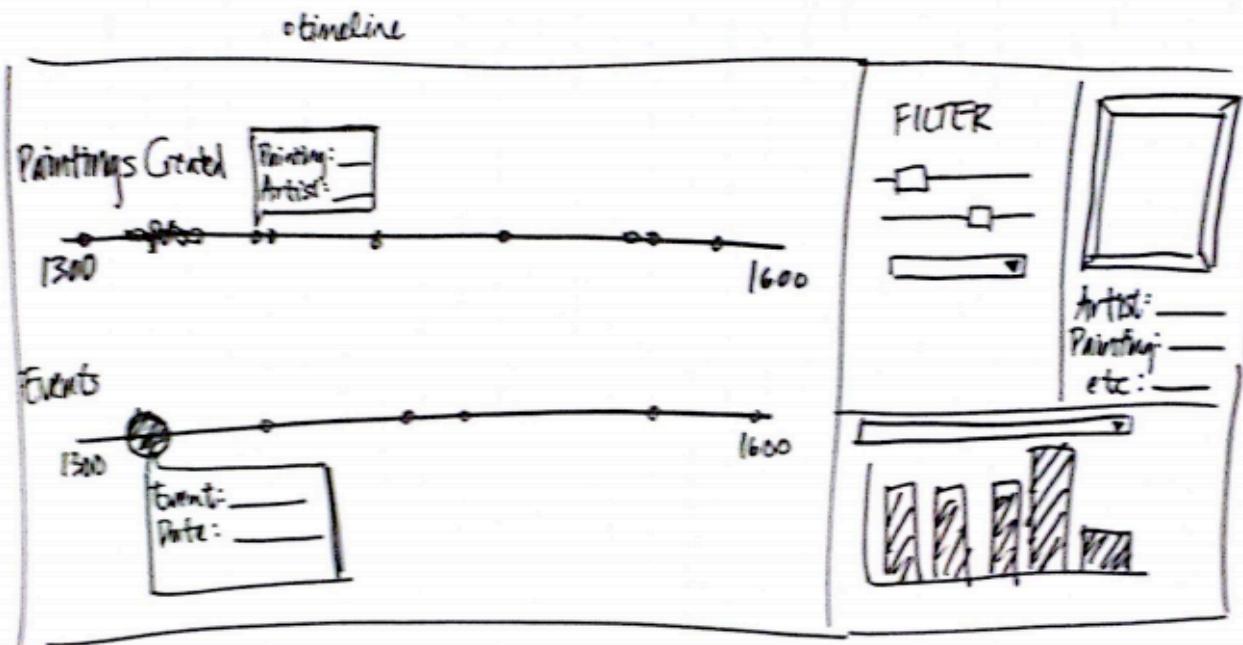
Dragging multiple circles to a box in the corner would spawn a circle with the topics combined.

## Word Cloud



- Size based on number of paintings with the topic (larger = more paintings)
- Color based on average year that topic appears (some dark color to some light color)
- Circles will be movable for easy comparison and hopefully will be able to be sorted by avg. year or frequency.
- Small white nodes in topic circles can be clicked and will act like normal nodes
- Dragging multiple circles to the box will spawn a circle at some position with size and color dictated by above rules

## TIMELINE



For the timeline component of our visualization, the idea is to have parallel timelines that compare the paintings created versus a timeline of the major events in the Italian Renaissance. The nodes would still be responsive to the detail portion of the visualization next to the filter, and the nodes on the timelines would display appropriate details about the painting or the event when hovered over.

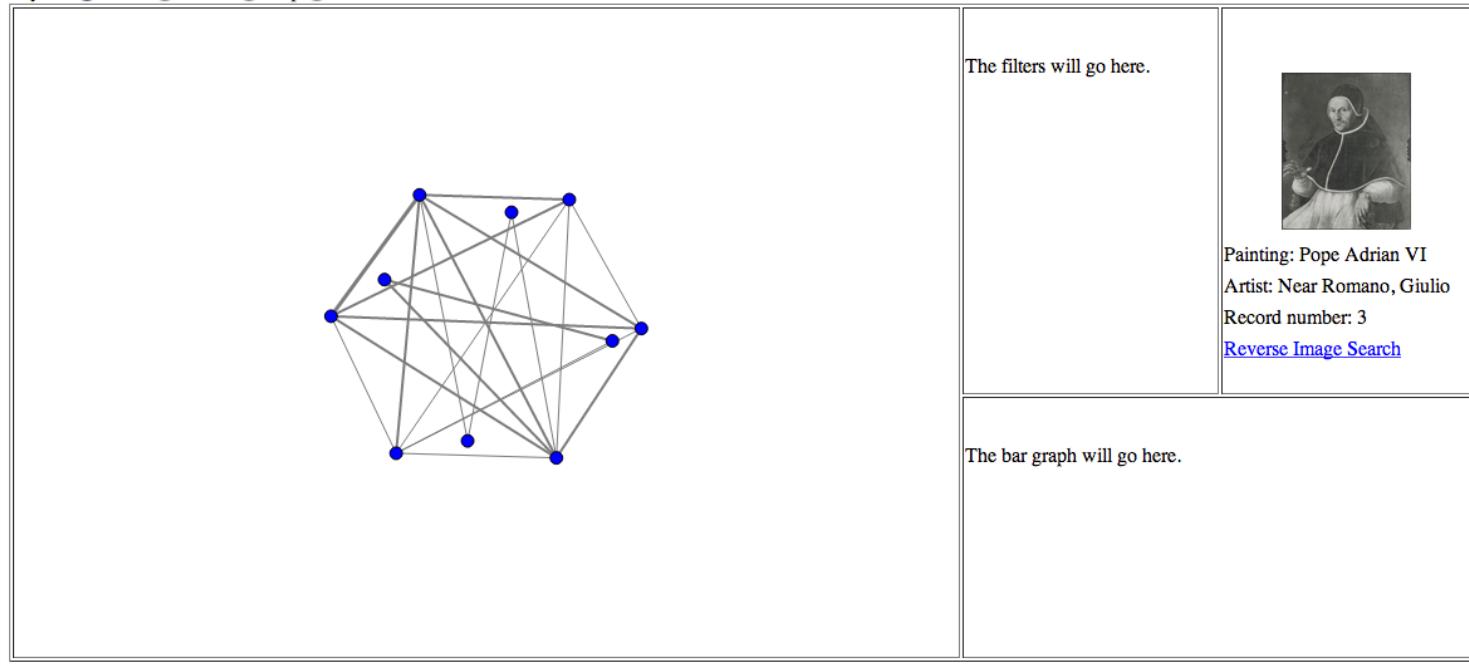
# IMPLEMENTATION

- To implement our visualizations, we utilized much of the second problem set with the graph/node and the radio buttons and the fourth problem set with the geomap of the US. Since ~11,000 points is a lot and can significantly delay the visualization, conditions were made through different filters such as the date of the painting, the artist, or simply the total number of paintings shown, to ensure that not all of the data points are shown at once. In addition, colors would also be added to depict the artist at hand or the time period chosen to allow for easy comprehension and compartmentalization of the data for the user.

# 1ST SCREENSHOT

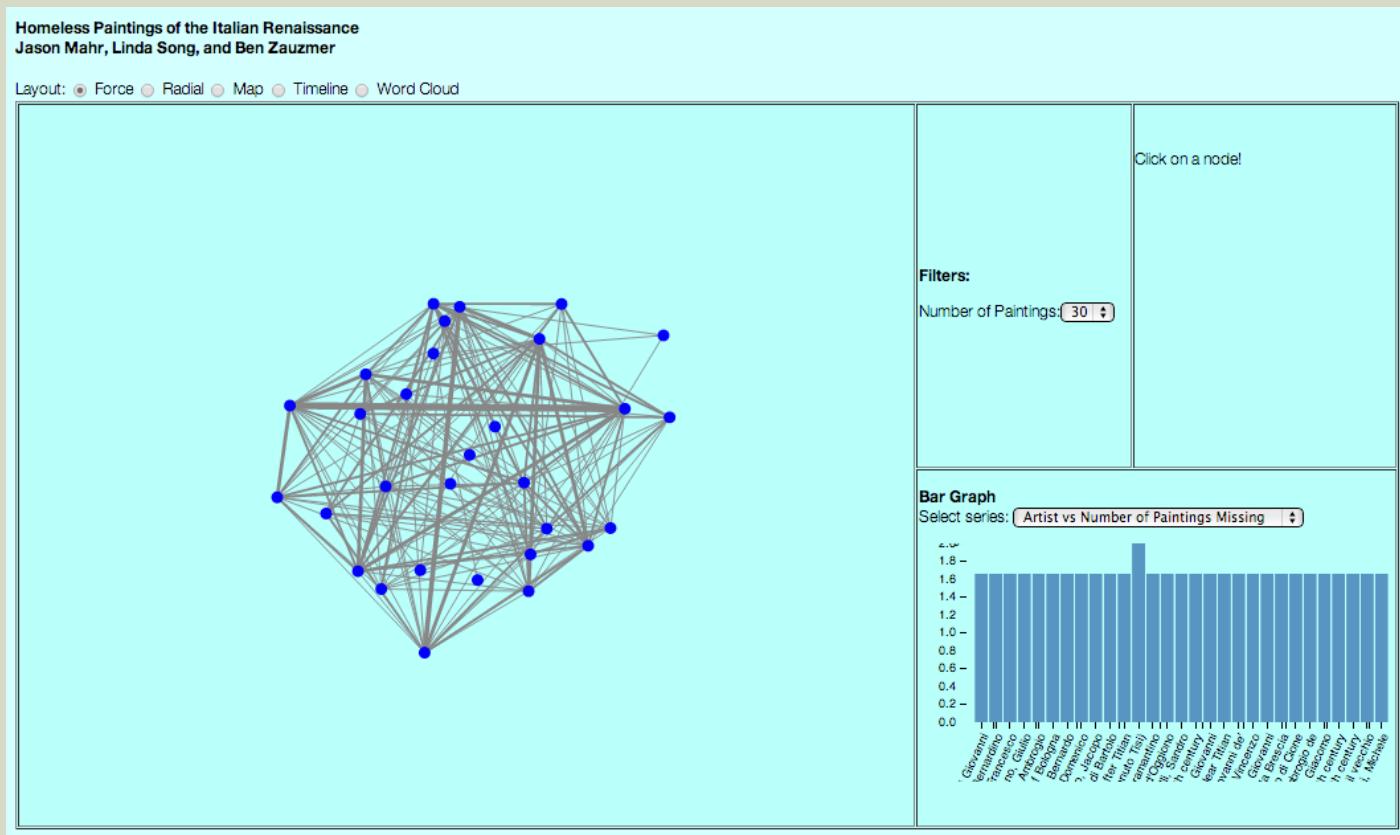
Homeless Paintings of the Italian Renaissance  
Jason Mahr, Linda Song, and Ben Zauzmer

Layout:  Force  Radial  Map  Timeline



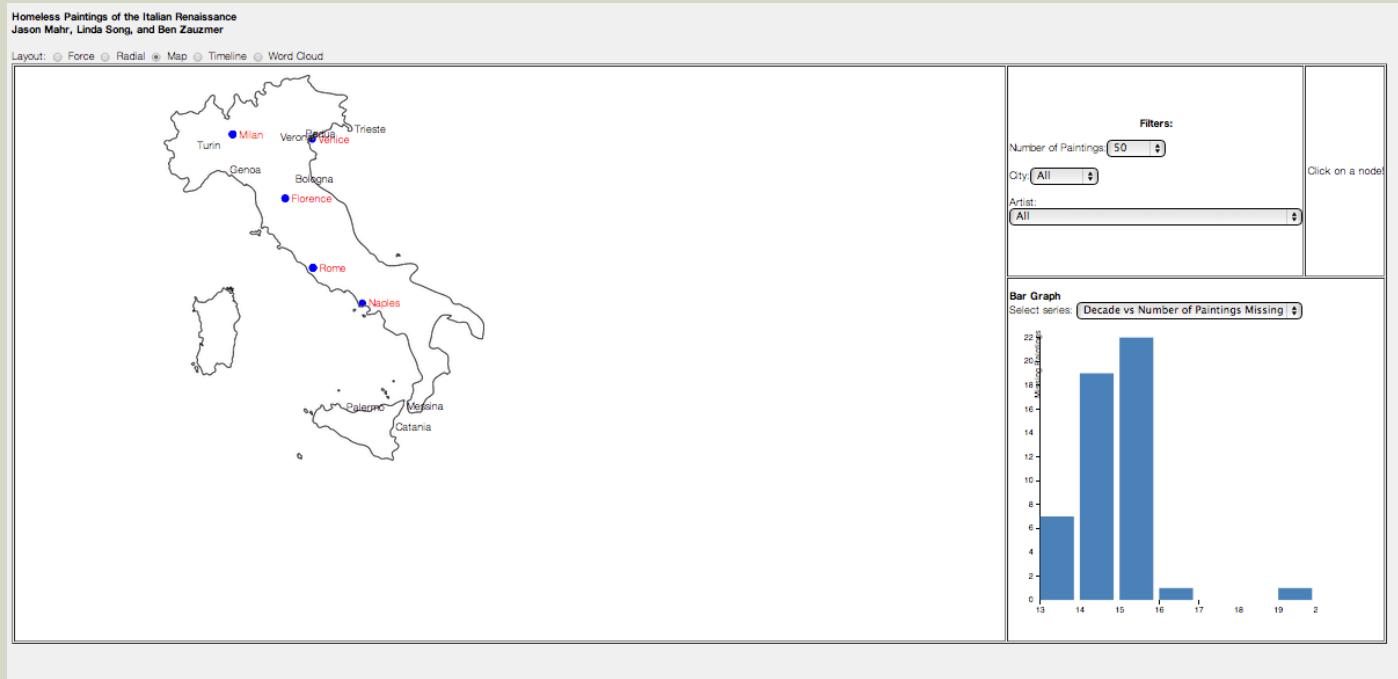
The first screenshot shows the initial structure and layout of our visualization. We took a sample amount of data (sample size of 10) to be able to work with the data without overloading the browser. The basic layout was inspired by the 2<sup>nd</sup> problem set of visualizing github data into a graph with nodes and links, adapting the force and radial functionalities of the nodes as a starting point. A Google reverse image search was also implemented for the user to be able to search for the image itself online with ease on a new tab.

# 2<sup>ND</sup> SCREENSHOT



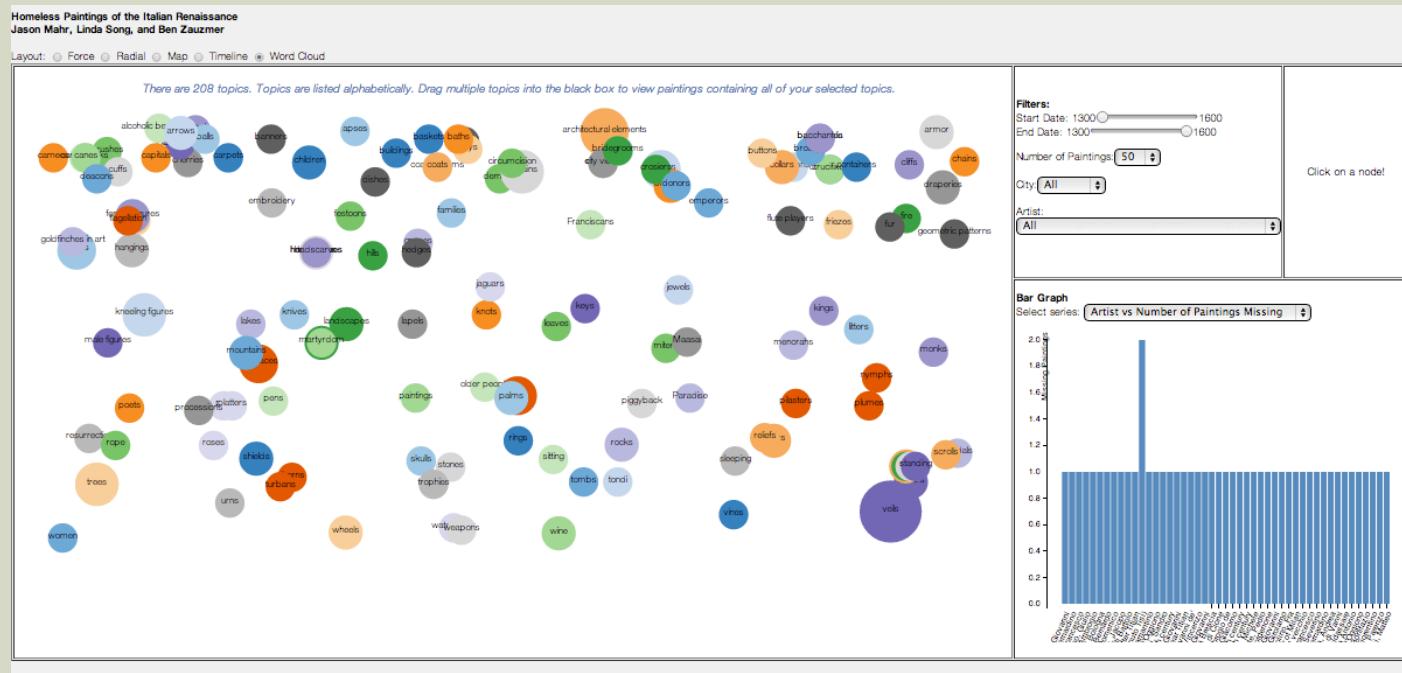
The second screenshot of our visualization illustrates the bar graph in the bottom right hand corner beginning to take shape as it depicts the total number of paintings lost in the data per artist. The filters are also beginning to be set up to give control to the user the amount of data to be shown at once.

# 3RD SCREENSHOT



This screenshot shows the map of Italy and its corresponding paintings in the dataset as well as a different bar graph that compares the number of paintings missing to the century the painting was created. The filters are also advancing in complexity as the cities where the painting was last seen can be filtered out as well as the artists who are in our dataset. This location also caused the map to be relatively useless to be zoomed in as the locations are not GPS accurate but are rather the names of cities in Italy instead.

# 4<sup>TH</sup> SCREENSHOT



As designed in the previous slide, this screenshot is the word cloud made of the topics that describe the painting at hand by key words that depict elements or events within the painting or such as Jesus, infants, or worship. The colors in this screenshot aren't indicative of anything in particular, but the plan is to make the colors more relevant by having the color depict the average year that the paintings with that topic have and using a gradient to show the relevance to that topic to the current year.

# EVALUATION

- The visualization of our data allows us to visually compare the number of paintings missing and gain a better conceptual understanding of how the missing paintings relate to each other and places them into a greater context. Through the map, the user is able to see where paintings from which time period are generally missing from, making connections to how the geographic location of the painting relates to the amount of missing paintings. In addition, the parallel timeline puts the missing paintings in context with events that were surrounding the paintings at their completion. Our visualization allows for this cross comparison and overall a more complete conceptual understanding of how time, the artist, topic, and location impacted the paintings' fate as the user interacts with the filters that dynamically change all components of the visualization. However, the visualization is far from perfect; in the next steps of improvement, a more definite interaction between the differing visualizations and how the visualization handles the data as the filtering is less focused would be adjusted so that the visualization can handle all ~11,000 paintings at once without overloading the browser.

# CONCLUSION

- In conclusion, this visualization allows for users interested in the Italian Renaissance and its missing paintings to place the missing paintings into a greater context that assists their understanding of where and potentially why the paintings are missing. This could further help the current efforts to find the homeless paintings of the Italian Renaissance and the historians' efforts to locate and restore the works of art to be studied and appreciated by artists and art appreciators alike.