CESTA is a pioneering research hub where we apply digital tools and methods to create new knowledge in interdisciplinary humanistic inquiry.
Summer is a unique season at CESTA. Our space in Wallenberg Hall is enlivened by a diverse cohort of undergraduate researchers, who spend the bulk of this “break” learning, experimenting, documenting, and ultimately, growing in community. At the outset, our interns discuss what they hope to get out of this experience and how to best synthesize their efforts and their learning. In lieu of research posters, students created condensed research stories for their projects, producing a narrative and visual lookbook of the CESTA summer program.

This *CESTA Summer Research Anthology* is an artifact of this summer’s cohort while gesturing to our community’s past and future. Multi-year projects are nearing conclusion and others are just beginning. On behalf of Brian, Erik, and the entire CESTA community—enjoy the stories in this collection and find more about these and other projects online at [cesta.stanford.edu](http://cesta.stanford.edu).
<table>
<thead>
<tr>
<th>Project Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>by Cathy Yang and Mark York</td>
<td></td>
</tr>
<tr>
<td>GRAND TOUR PROJECT</td>
<td>4</td>
</tr>
<tr>
<td>Exploring Early Modern Travelers through Digital Databases</td>
<td></td>
</tr>
<tr>
<td>by Ryan Tan, Mark York, and Brian Grenadier with Giovanna Ceserani</td>
<td></td>
</tr>
<tr>
<td>EARLY MODERN MOBILITY</td>
<td>10</td>
</tr>
<tr>
<td>Roads and Routes: Transregional Mobility in the Early Modern Era</td>
<td></td>
</tr>
<tr>
<td>by Evan Kim and Elliot Miller with Rachel Midura, Leo Barleta, Katie McDonough, and Paula Findlen</td>
<td></td>
</tr>
<tr>
<td>SPATIAL HISTORY PROJECT: URBAN DISCONTENTS</td>
<td>20</td>
</tr>
<tr>
<td>Precarious Footholds: Housing Vulnerability and Urban Change in Global Perspective</td>
<td></td>
</tr>
<tr>
<td>by Hadassah Betapudi, Janine Fleming, Michelle Julia Ng, Sierra O Ke Akua Burgon, and Aryan Singh with Zephyr Frank, Erik Steiner, and Leo Barleta</td>
<td></td>
</tr>
<tr>
<td>SPATIAL HISTORY PROJECT: LANDTALK</td>
<td></td>
</tr>
<tr>
<td>LandTalk: Human Experiences of Landscape Change</td>
<td>20</td>
</tr>
<tr>
<td>by Erick Enriquez with Erik Steiner, Deborah Gordon, and Cody Leff</td>
<td></td>
</tr>
<tr>
<td>SPATIAL HISTORY PROJECT: IMAGINED SAN FRANCISCO</td>
<td>24</td>
</tr>
<tr>
<td>Planned and Unplanned: Mapping Urban Power in San Francisco</td>
<td></td>
</tr>
<tr>
<td>by Krain Chen with Ocean Howell, Erik Steiner, Brian Kersey, and Cody Leff</td>
<td></td>
</tr>
<tr>
<td>EARLY CHRISTIANS UNDER ISLAMIC RULE</td>
<td>30</td>
</tr>
<tr>
<td>Thomas of Marga’s Book of Governors</td>
<td></td>
</tr>
<tr>
<td>by Palmer Manes and Zuyi Zhao with Michael Penn</td>
<td></td>
</tr>
<tr>
<td>LITERARY LAB: THE HUGO AND NEBULA AWARDS</td>
<td>34</td>
</tr>
<tr>
<td>Understanding Book Awards through Big Data</td>
<td></td>
</tr>
<tr>
<td>by Shana Hadi with Laura McGrath, J. D. Porter, and Mark Algee-Hewitt</td>
<td></td>
</tr>
<tr>
<td>LITERARY LAB: TECHNE PROJECT</td>
<td>38</td>
</tr>
<tr>
<td>Found in Translation: Building an “Atlas” of World Literature</td>
<td></td>
</tr>
<tr>
<td>by Eunji Lee with Laura McGrath, J. D. Porter, and Mark Algee-Hewitt</td>
<td></td>
</tr>
<tr>
<td>JOSQUIN RESEARCH PROJECT</td>
<td>40</td>
</tr>
<tr>
<td>Understanding Renaissance Music Through Density and Dissonance</td>
<td></td>
</tr>
<tr>
<td>by Young Fenimore Lee with Jesse Rodin and Craig Sapp</td>
<td></td>
</tr>
<tr>
<td>GLOBAL MEDIEVAL SOURCEBOOK</td>
<td>44</td>
</tr>
<tr>
<td>The Global Medieval Sourcebook: An Online Portal to the Middle Ages</td>
<td></td>
</tr>
<tr>
<td>by Nina Du and Tina Zhang with Kathryn Starkey and Mae Lyons-Penner</td>
<td></td>
</tr>
<tr>
<td>TEXT TECHNOLOGIES: SOPES</td>
<td>48</td>
</tr>
<tr>
<td>Ordinary People, Extraordinary Stories</td>
<td></td>
</tr>
<tr>
<td>by Karunya Bhramasandra with Elaine Treharne</td>
<td></td>
</tr>
<tr>
<td>TEXT TECHNOLOGIES: MAPPING MANUSCRIPTS PROJECT</td>
<td>52</td>
</tr>
<tr>
<td>Digitally Tracing the Distribution of Medieval Manuscripts</td>
<td></td>
</tr>
<tr>
<td>by Timothy Karoff with Mateusz Fafinski and Elaine Treharne</td>
<td></td>
</tr>
<tr>
<td>Updates from other ongoing CESTA projects</td>
<td>56</td>
</tr>
</tbody>
</table>
Introduction

by Cathy Yang and Mark York

Undergraduate Research Interns are the heart of CESTA’s community, working closely with faculty and staff, and contributing directly to innovative research at the intersection of technology and the humanities. This summer, we’ve had a wonderful cohort of students from majors across campus including Religious Studies, Symbolic Systems, Civil Engineering, English, History, and Computer Science. As student researchers, our cohort has worked on a variety of projects in collaboration with faculty leads from across and beyond Stanford University. This anthology showcases the contributions of students working on faculty-led research projects that combine technology and the humanities to foster scholarly discussions across different disciplines.

With mentorship from our project leads, we used digital tools and methods to investigate cultural records, objects, and historical phenomenon through space and time. Over ten weeks, we explored big ideas by bringing diverse disciplines and methods together from the humanities and sciences. From mapping massive urban areas and translating ancient poetry, to visualizing the migration of people and ideas, we worked with each other to cultivate invaluable academic and professional skills. We have developed skills in critical thinking, creative problem solving, hypothesizing, collaboration, and ethical decision making.

This has been an eventful summer of research and learning here at CESTA. We participated in a wide variety of workshops on tools and techniques at the forefront of the digital humanities. We learned how to use spreadsheets to organize our metadata, to engage in social network analysis, to craft a professional website and portfolio, to create powerful maps and visualizations through GIS and Tableau, and were introduced to coding in Python and R. We also explored historical documents firsthand at Stanford Library’s
Special Collections and toured the Cantor Arts Center. These insights have enriched our intern experience, and trained us as digital humanities researchers.

We would like to thank all the CESTA staff—Amanda Wilson Bergado, Brian Kersey, GP LeBourdais, and Erik Steiner—for all their help and support in facilitating student research and making this anthology possible. We would also like to thank the faculty members and graduate students for their scholarly work and collaboration, as well as Leo Barleta, Scott Bailey, Quinn Dombrowski, Claudia Engel, Stace Maples, David Medeiros, Rachel Midura, and Vincent Nicandro for their continued support of humanities students and scholarship. We are proud of what we, as a community, have done over the summer, and we are excited to share it with you here. To learn more about student interns and their projects as well as publications, events, and academic programs at CESTA, explore our website at cesta.stanford.edu.

With gratitude,
Your Communications Team
*Cathy & Mark*

---

**Student Profiles**

**Cathy Yang**
Cathy Yang is a senior majoring in Art Practice and minoring in East Asian Studies. Cathy is fascinated by the ability of CESTA to bridge technology and the humanities. Outside of CESTA, she enjoys sketching, reading, and drinking tea.

**Mark York**
Mark York is a Communication major and Creative Writing minor now in his junior year. Outside the lab, he is a lover of animation, books, dungeons and dragons... basically anything nerdy. He has a pocket-watch collection, and dabbled in making a YouTube channel once or twice. He hopes to publish a young adult fantasy book series.
The Grand Tour Project is dedicated to applying computational approaches to better understand eighteenth-century travels to Italy. These journeys involved tens of thousands of travelers, seeking to experience Italy’s ancient past, classical ruins, and art heritage, as well as contemporary music, culture, and political systems. Collectively known as the Grand Tour of Italy, the importance of these journeys, undertaken just at the dawn of modernity, is hard to overestimate, as they contributed to a massive reimagining of politics and the arts, of the market for culture, and of practices of professionalism. The social and political elites of the time gained ‘citizens of the world’ status by their Italian travels, while the major scholars and artists of the age found in their own tours invaluable material and inspiration.
Since 2008 the Grand Tour Project has worked to create and leverage digital tools, analysis, and visualizations to bring us closer to the diverse travelers, elites and otherwise, who collectively constituted the world of travels to Italy in the eighteenth century. We have been working with the more than five thousand entries in A Dictionary of British and Irish Travelers to Italy 1701-1800, compiled from the Brinsley Ford Archive by John Ingamells (Yale University Press 1997), digitally transforming it to create a dynamic searchable database of more than six thousands digital entries along with digital visualizations of these travelers’ journeys and lives.

The project is now preparing for publication, which will include the interactive database, complete documentation and history of the database creation, essays by leading scholars of eighteenth-century history and culture exemplifying research completed using the database, and pedagogical material to bring the Grand Tour in the classroom in the form of a digital humanities course. This summer, while working towards publication, some of the most pressing questions for the project have been how to make its material most accessible to readers, from advanced researchers to students, and what is at stake in producing a digital work that originates from a printed book.

Dictionary of British and Irish Travelers to Italy 1701-1800
Katherine Read


Katherine Read was described late in life by Fanny Burney as melancholy, shrewd and clever, ‘most exceedingly ugly’ and neglectful in her dress. In 1752, aged twenty-seven and having already studied under Quentin de la Tour in Paris, she went alone to Italy to complete her training. On 16 June 1753 she wrote to her eldest brother from Rome: ‘I have had no money but from you since I came abroad… I am obliged to board, otherwise I could live at a third of the expense; this you may believe is no small vexation!’ In the same letter she said she was studying under L.-G. Blanchet, and copying portraits by Dolci and Van Dyck: ‘I apply so constantly and take every decent method of improvement that I think it must be impossible I can miss. I am but in a manner beginning to be known here’. She had met Cardinal Albani and was copying four heads by Rosalba Carriera from his collection. In her next letter of 8 June 1753 she said she had painted two Italian Princesses and was soon to paint Prince Caesarini’s brother, but payment, she complained, was made only with trinkets, ‘for you know that Italians despise people so much that they are obliged to do anything for money’. Lord Charlemont, however, would pay her for the profile she was to paint.
What is gained but also what might be lost when moving from traditional paper-bound to digital printing? A major feature we worked on this summer is a set of digital visualizations that address this question, aiming to give the readers of our project a fuller sense of the database. What inspired us here is thinking through the difference between the printed text—the 1997 Dictionary—and the database. When you hold the Dictionary in your hands, you have an immediate perception of its weight which conveys a sense of its content, and as you scroll through its pages, you also get a sense through your
long experience as a reader of printed texts, of how many, how long, and how varied its entries are. While the digital database offers immediate access to more precise information—you find out instantly for example the exact number of the entries which is enumerated as 6,005—it is harder to appreciate, in list form, the degree of variation between different entries, even just at the level of word count. This summer we worked on the “Dot Chart of Travelers” (see Figure 3 for its sketch), transforming the word count of the entries into an interactive visualization that lets you appreciate the full scope of the

FIGURE 3
Dot Chart of Travelers
6,005 traveler entries (represented as dots) can be filtered and colored by categories to get a better sense of the shape of the Dictionary data.
Grand Tour Project

The Grand Tour Project is dedicated to applying computational approaches to better understand eighteenth-century travels to Italy. These journeys involved tens of thousands of travelers, seeking to experience Italy’s ancient past, classical ruins, and art heritage, as well as contemporary music, culture, and political systems. Collectively known as the Grand Tour of Italy, the importance of these journeys, undertaken just at the dawn of modernity, is hard to overestimate, as they contributed to a massive reimagining of politics and the arts, of the market for culture, and of practices of professionalism. The social and political elites of the time gained ‘citizens of the world’ status by their Italian travels, while the major scholars and artists of the age found in their own tours invaluable material and inspiration.

Since 2008 the Grand Tour Project has worked to create and leverage digital tools, analysis, and visualizations to bring us closer to the diverse travelers, elites and otherwise, who collectively constituted the world of travels to Italy in the eighteenth century. We have been working with the more than five thousand entries in A Dictionary of British and Irish Travelers to Italy 1701-1800, compiled from the Brinsley Ford Archive by John Ingamells (Yale University Press 1997), digitally transforming it to create a dynamic searchable database of more than six thousand digital entries along with digital visualizations of these travelers’ journeys and lives.

Figure 2 [below] shows, as an example, the digital entry pertaining to the painter of the picture in Figure 1. The project is now preparing for publication, which will include the interactive database, complete documentation and history of the database creation, essays by leading scholars of eighteenth-century history and culture exemplifying research completed using the database, and pedagogical material to bring the Grand Tour in the classroom in the form of a digital humanities course. This summer, while working towards publication, some of the most pressing questions for the project have been how to make its material most accessible to readers, from advanced researchers to students, and what is at stake in producing a digital work that originates from a printed book.

What is gained but also what might be lost when moving from traditional paper-bound to digital printing? A major feature we worked on this summer is a set of digital visualizations that address this question, aiming to give the readers of our project a fuller sense of the database. What inspired us here is thinking through the difference between the printed text—the 1997 Dictionary—and the database. When you hold the Dictionary in your hands, you have an immediate perception of its weight which conveys a sense of its content, and as you scroll through its pages, you also get a sense through your database—all travelers’ entries are represented with a dot—and its variety in a number of significant categories, such as gender (by which you can color and group the dots), word count (by which you can size), and more. This chart allows you at a single glance to appreciate the full extent of the database but also to manipulate this representation interactively to get a better sense of the shape of the data in terms of important variables.

These new visualizations have been a major product of this summer’s work, alongside ongoing refinement and checking of the integrity of the data for its successful publication, and editing and enhancement of the pedagogical material, which was produced in the context of the digital humanities course taught last Winter about the Grand Tour jointly taught by Giovanna Ceserani and Rachel Midura. We are excited to see our contributions included in the project’s forthcoming publication which will offer it for use well beyond Stanford.

Acknowledgments

We would like to express our appreciation to our project lead, Giovanna Ceserani, who provided feedback, engagement, and support throughout the summer. Rachel Midura provided invaluable guidance in preparing the pedagogical material for publication. In addition, we would also like to thank CESTA staff Amanda Wilson Bergado, Brian Kersey, and Erik Steiner for their support and direction.
Italian travels, while the major scholars and artists of the age elites of the time gained 'citizens of the world' status by their reimagining of politics and the arts, of the market for culture, hard to overestimate, as they contributed to a massive of these journeys, undertaken just at the dawn of modernity, is Collectively known as the Grand Tour of Italy, the importance ancient past, classical ruins, and art heritage, as well tens of thousands of travelers, seeking to experience Italy's eighteenth-century travels to Italy. These journeys involved Grand Tour Project he is dedicated to applying producing a digital work that originates from a printed book. readers, from advanced researchers to students, and what is at stake in questions for the project have been how to make its material most accessible to summer, while working towards publication, some of the most pressing Grand Tour in the classroom in the form of a digital humanities course. This research completed using the database, and pedagogical material to bring the by leading scholars of eighteenth-century history and culture exemplifying database, complete documentation and history of the database creation, essays The project is now preparing for publication, which will include the interactive digitally transforming it to create a dynamic searchable database of more than six compiled entries in eighteenth century. We have been working with the more than five thousand and otherwise, who collectively constituted the world of travels to Italy in the tools, analysis, and visualizations to bring us closer to the diverse travelers, elites Since 2008 the Grand Tour Project has worked to create and leverage digital thousands digital entries along with digital visualizations of these travelers' journeys and lives. Figure 2 [below] shows, as an example, the digital entry pertaining to the painter of the picture in Figure 1. content, and as you scroll through its pages, you also get a sense through your have an immediate perception of its weight which conveys a sense of its Dictionary—and the database. When you hold the Dictionary in your hands, you a set of digital visualizations that address this question, aiming to give the paper-bound to digital printing? A major feature we worked on this summer is What is gained but also what might be lost when moving from traditional long experience as a reader of printed texts, of how many, how long, and how varied its entries are. While the digital database offers immediate access to level of word count. This summer we worked on the "Dot Chart of Travelers" number of the entries which is enumerated as 6,005—it is harder to appreciate, more precise information—you find out instantly for example the exact database—all travelers' entries are represented with a dot—and its variety in a shape of the data in terms of important variables. Also to manipulate this representation interactively to get a better sense of the allows you at a single glance to appreciate the full extent of the database but group the dots), word count (by which you can size), and more. This chart also to number of significant categories, such as gender (by which you can color and database—all travelers' entries are represented with a dot—and its variety in a shape of the data in terms of important variables. Also to manipulate this representation interactively to get a better sense of the allows you at a single glance to appreciate the full extent of the database but group the dots), word count (by which you can size), and more. This chart also to number of significant categories, such as gender (by which you can color and... Mark York is a Communication major and Creative Writing minor now in his junior year. Outside the lab, he is a lover of animation, books, dungeons and dragons... basically anything nerdy. He has a pocket-watch collection, and dabbled in making a YouTube channel once or twice. He hopes to publish a young adult fantasy book series.

Student Profiles

Ryan Tan
Ryan Tan is a junior with an interest in computer science, statistics, comparative literature, philosophy, and music. He loves to discuss music, is interested in sentiment analysis, and eventually wants to direct a play on campus.

Brian Grenadier
Brian Grenadier is a sophomore (Class of 2022) majoring in History. He is especially interested in applying modern computational techniques (textual and statistical analysis) to exciting historical questions. He loves playing golf and tennis and is absolutely delighted to be part of the Stanford community.

Mark York
Mark York is a Communication major and Creative Writing minor now in his junior year. Outside the lab, he is a lover of animation, books, dungeons and dragons... basically anything nerdy. He has a pocket-watch collection, and dabbled in making a YouTube channel once or twice. He hopes to publish a young adult fantasy book series.
Roads have always been an important component of the existence and maintenance of communities, economies, and governments: without roads, information cannot be communicated, and resources cannot be transmitted. What do the methods and patterns of transportation reveal about the early modern period (1500-1800)? This is the main research question that *Early Modern Mobility*, a transregional study of mobility, wishes to address. Studying the history of mobility as the patterns through which individuals travel and the methods by which they navigate reveals much about society at the time. The evolution of systems of mobility therefore
Postal Routes
An example postal route from Augsburg (Augusta) to Cologne (Colonia), with intermediate destinations and a count of postal waystations denoted. Source: Ottavio Codogno, *Nuovo itinerario delle poste per tutto il mondo* (Milan: Girolamo Bordoni, 1608).

reflects changes of political and economic possibilities of contemporary societies. Additionally, knowledge of the routes and roads by which people travel is crucial in developing a conception of the space in which individuals lived. In the early modern period, such routes were charted in sources like itineraries, whose gradual international development at the time reflected the increasing conception of a connected, globalized Europe.

By using technologies such as GIS and temporal networking, we were able to create lucid visualizations of mobility patterns that we would not be able to otherwise conceptualize. Putting these early modern mobilities into scope requires the collection of novel comprehensive datasets, which in turn creates sub-projects based on the nature of the data. One of these, *The Early Modern Digital Itineraries Project*, focuses on investigating travel routes published in itineraries, cataloging their locations, and determining how publications of such routes varied between authors over time, thus shaping temporal networks. Another sub-project, *The Breton Corvée Project*, aims to digitize a geographical dictionary of Brittany.
Itinerary Network
A network visualization of connections between cities documented in itinerary routes published between 1630 and 1650. Different colors indicate different modern nationalities.
loads have always been an important component of the existence and maintenance of communities, economies, and governments: without roads, information cannot be communicated, and resources cannot be transmitted. What do the methods and patterns of transportation reveal about the early modern period (1500-1800)? This is the main research question that Early Modern Mobility, a transregional study of mobility, wishes to address. Studying the history of mobility as the patterns through which individuals travel and the methods by which they navigate reveals much about society at the time. The evolution of systems of mobility therefore reflects changes of political and economic possibilities of contemporary societies. Additionally, knowledge of the routes and roads by which people travel is crucial in developing a conception of the space in which individuals lived. In the early modern period, such routes were charted in sources like itineraries, whose gradual international development at the time reflected the increasing conception of a connected, globalized Europe. By using technologies such as GIS and temporal networking, we were able to create lucid visualizations of mobility patterns that we would not be able to otherwise conceptualize. Putting these early modern mobilities into scope requires the collection of novel comprehensive datasets, which in turn creates sub-projects based on the nature of the data. One of these, The Early Modern Digital Itineraries Project, focuses on investigating travel routes published in itineraries, cataloging their locations, and determining how publications of such routes varied between authors over time, thus shaping temporal networks. Another sub-project, The Breton Corvée Project, aims to digitize a geographical dictionary of Brittany from an eighteenth-century road construction boom fueled by a corvée labor requirement. This allows us to compare historic and modern locations from administrative documents, and provides insight into the decision-making of parishes, interactions between communities, and technical and economic discussions of the time. A third sub-project, Empire in the Backlands, records the geographic mobility of approximately 30,000 individuals from colonial Brazil, as well as a disjunct dataset of Power of Attorneys entries. To come to an understanding of mobility, it is necessary to make use of digital sources.

We’d like to thank Rachel Midura, Leo Barleta, and Katherine McDonough for their guidance and assistance throughout the process. We’d also like to thank Brian Kersey and Amanda Wilson Bergado for their support.

---

**Student Profiles**

**Evan Kim**

Evan Kim is a junior majoring in Symbolic Systems. He is interested in both technology and the humanities, having studied computer science, philosophy, and literature at Stanford. He also has an interest in music, as he plays both the cello and piano. Additionally, he participates in Stanford’s chamber music program as a cellist. For fun, Evan likes to read novels and watch movies and documentaries.

**Elliot Miller**

Elliot Miller is a senior majoring in Mathematical and Computational Science with a plan to coterm in Computer Science. His favorite part of his internship was getting to know the community and sharing in the excitement of their work. He’s a member of the Wind Ensemble, the Stanford Band, Jazz Orchestra, and a moderator of Stanford Humans vs. Zombies.
Precarious Footholds: Housing Vulnerability and Urban Change in Global Perspective

by Hadassah Betapudi, Janine Fleming, Michelle Julia Ng, Sierra O Ke Akua Burgon, and Aryan Singh
Project Team: Zephyr Frank, Erik Steiner, and Leo Barleta
Urban development produces uneven patterns of housing. These patterns evolve out of complex factors that include urban planning institutions, infrastructure, environmental conditions, and socioeconomic dynamics. Patterns in housing affordability, in living conditions, and in the social distribution of vulnerability of tenancy reveal socio-economic boundaries between zones of displacement in contemporary cities. Although certain underlying factors such as zoning and public infrastructure can be assumed to be present and operating in every city, the interaction of these forces and their relative importance across global regions remains largely unstudied in a comparative and historical perspective.

This study aims to contribute to understanding the factors influencing urban housing vulnerability and displacement through comparative analysis grounded in geospatial data and new techniques of machine learning and computer vision. Over the summer, our team focused on leveraging visual data about cities available in the form of Google Street View (GSV) images, asking the way they can contribute to the comparative analysis of housing vulnerability and displacement across diverse cities around the globe. Furthermore, we were also interested in examining how new technologies such as computer vision can help in this task.

Sources and Data
We first aimed to reveal patterns and relationships among variables about housing vulnerability and urban life in what we define as “average” or “unexceptional” cities, cities with populations between 500,000 and 3,000,000, which were not capitals, and lacked exceptional political status, demographic composition, wealth, or other attributes. Combining satellite imagery and
OpenStreetMap data, we randomly selected GSV images of ~10,000 points in residential streets within each metropolitan area. The images offer a visual summary of the studied cities from the viewpoint of the street and, through different methodologies, our team identified how they related to urban vulnerability. Using Google’s Computer Vision software, we analyzed generated “tags” and “threshold” values of the images. A “tag” generated for an image indicates that a given object tagged is present in that image, while the threshold value indicates the software’s confidence in the tag. It is important to note that all tagging (both object recognition and image classification) was carried out based on Google Computer Vision and AutoML platforms.

**Methods**

Our team developed three different approaches to use machine learning to examine this massive visual archive of cities. First, we analyzed the collection of images using computer vision models that are able to recognize features in the landscape and generate a series of tags describing objects, context, and other aspects of individual images. In our second approach, we focused on a statistical analysis of the tag data generated. This helped determine which tags tended to cluster. In contrast to the other two approaches that used existing computer vision tools, we also trained our own computer vision models to detect vulnerability. After visual analysis of hundreds of these images, we created a working
hypothesis that these images comprised three different characteristics associated with urban vulnerability and are worth further exploration. Our first observation in the images was that they represent areas with different residential densities. Second, we investigated whether the landscape was stable or under a process of transformation (construction, visible signs of vacancy, decay, etc.) and thus more likely to prompt dislocation. Lastly, we tagged for the perception of capital investment in the image, which included physical infrastructure, maintenance, and other signs of wealth/poverty.

**Preliminary Results**

Where our tagged images with “vulnerable” characteristics overlap with known areas of distress/displacement, we could have confirmation that the GSV tags correspond to a meaningful concept of vulnerability and
displacement. For instance, if the Philadelphia tags cluster in the same areas where we have data for vacant buildings—that correlation could reveal vulnerability. A second kind of confirmation of vulnerability would be one where the government has designated an area as distressed, in conflict, etc. We have seen some examples of this in Curitiba. In expanded research, we can test to see whether the GSV tags for vulnerability correspond to these areas.

**Next Steps**

The enormous collection of images available through Google Street View offers an entry point to visualize issues associated with urban vulnerability and displacement. These images could be analyzed in several different ways, but digital methods such as those used in this project allow for comprehensive examination of large datasets. The team expects that, with the refinement of the computer vision models and its theoretical underpinnings, we could plot in maps some of the visual signs of that we consider associated to vulnerability.

---

**Student Profiles**

**Hadassah Betapudi**

Hadassah Betapudi is a sophomore who learned about CESTA through the Humanities Research Intensive class, taught by Dr. Elaine Treherne and Dr. Caroline Winterer. She is pre-med, majoring in Political Science. She is interested in pioneering data-driven humanities to study poverty and inequality, medicine, and health policy.

**Janine Fleming**

Janine Fleming is a sophomore who is planning to major in Science, Technology and Society. She is working on the *Urban Discontents* project, studying and comparing vulnerability in different cities around the world. She is interested in human-computer interaction, as well as how technology can be used for social good. In her free time, Janine loves painting, learning how to DJ, playing volleyball, and visiting the beautiful Pacific coastline.
Student Profiles

Michelle Julia Ng

Michelle Julia Ng is a junior majoring in Computer Science + History at Stanford University. This summer, she is working on the Urban Discontents project at CESTA, where she uses data and mapping technologies to explore patterns of vulnerability in urban areas. Michelle is interested in using technology for social good, and she enjoys being able to exercise creativity in her research.

Aryan Singh

Aryan Singh is a sophomore with an interest in urban planning, infrastructure and public policy. He is working on performing statistical analyses of data for the Urban Discontents team this summer and hopes to explore the relationship between public transport and housing vulnerability in the future. Outside of CESTA, he is a member of Stanford’s Social Entrepreneurship Club and enjoys playing soccer for the men’s club team.

Sierra O Ke Akua Burgon

Sierra Burgon is a sophomore (Class of 2022) majoring in Human Biology with a concentration in Health and Health Policy. She is working on the Urban Discontents project, which seeks to produce insight on patterns of factors contributing to housing vulnerability across a sample of cities around the world.
LandTalk: Human Experiences of Landscape Change

by Erick Enriquez

Project Team: Erik Steiner, Deborah Gordon, and Cody Leff

LandTalk.org is a crowd-sourced collection of reports of changing landscapes from across the globe. Over the summer, I sought to create a platform where people share their observations about familiar but changing spaces by sharing videos, photographs, and perspectives of how local environments have evolved. Throughout this summer, our team’s primary goals were to develop and improve the project website that enables these people to share their stories, to develop a curriculum that facilitates student-driven collection of these conversations, and to explore methods of analyzing the data. We also investigated methods to extract, shape, and visualize the data so we could extrapolate meaningful patterns and conclusions.

I focused on the collection, design, implementation, and analysis of data to improve the website. I spent the first two weeks setting up my development environment with Docker, Node.js, Sublime, and WordPress before diving into code...
Mapping LandTalk Conversations

LandTalk has collected over 350 interviews from around the world with people who have lived in an area for more than 20 years. Observers discuss the changes they have witnessed over time in familiar, everyday landscapes: cities, parks, backyards, towns, trails, and beaches.

on GitHub and learning HTML, CSS, JavaScript, and PHP. By week three, I reformatted the online content and enabled a full-screen mode for the site’s maps of changing landscapes. While redesigning our web pages, I learned to use Sketch to create site mock-ups. I structured workflow with Erik Steiner, my project lead, that involved submitting design ideas to him, receiving feedback, and adjusting accordingly. Through this collaboration, we ensured the consistency and quality of site content by manually reviewing place names, coordinates, photographs, and transcriptions. Toward the end of the summer, I worked less on GitHub issues and instead focused on creating and cataloguing reproducible methods for data analysis. This process, known as Topic Modeling, will help us discover patterns and insights as our corpus continues to grow. Initial results are shown on the next page.

Rather than drawing from pre-existing documents, manuscripts, maps, and materials, our website’s primary function is to collect and display the data we hope to use for future analysis by a broad audience. On the LandTalk site, we provide a curriculum for various age groups and objectives that teachers across the world can implement in their classrooms. This curriculum guides students through
changing landscapes from across the globe. Over the summer, I sought to create a platform where people share their observations about familiar but changing spaces by sharing videos, photographs, and perspectives of WordPress before diving into code.

Docker, Node.JS, Sublime, and development environment with the first two weeks setting up my to improve the website. I spent the conversations, and to explore student-driven collection of these curriculum that facilitates share their stories, to develop a website that enables these people to develop and improve the project.

Our team’s primary goals were to evolve. Throughout this summer, how local environments have changed. While redesigning our patterns and conclusions. This process, known as reproducible methods for data collection of reports of GitHub issues and instead focused on GitHub and learning HTML, CSS, JavaScript, and PHP. By week

<table>
<thead>
<tr>
<th>“Looking back”</th>
<th>“Good Old Days”</th>
<th>“Play”</th>
<th>“Neighborhood”</th>
<th>“Yard”</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>kind</td>
<td>neighborhood</td>
<td>area</td>
<td>trees</td>
</tr>
<tr>
<td>moved</td>
<td>time</td>
<td>kids</td>
<td>people</td>
<td>yard</td>
</tr>
<tr>
<td>spend</td>
<td>people</td>
<td>play</td>
<td>neighborhood</td>
<td>house</td>
</tr>
<tr>
<td>family</td>
<td>big</td>
<td>children</td>
<td>development</td>
<td>plants</td>
</tr>
<tr>
<td>spent</td>
<td>area</td>
<td>family</td>
<td>local</td>
<td>front</td>
</tr>
<tr>
<td>spends</td>
<td>pretty</td>
<td>neighbors</td>
<td>live</td>
<td>tree</td>
</tr>
<tr>
<td>activities</td>
<td>moved</td>
<td>house</td>
<td>place</td>
<td>grass</td>
</tr>
<tr>
<td>kids</td>
<td>ago</td>
<td>playing</td>
<td>community</td>
<td>flowers</td>
</tr>
<tr>
<td>enjoys</td>
<td>good</td>
<td>basketball</td>
<td>lived</td>
<td>plant</td>
</tr>
<tr>
<td>misses</td>
<td>guess</td>
<td>homes</td>
<td>stores</td>
<td>garden</td>
</tr>
<tr>
<td>likes</td>
<td>place</td>
<td>houses</td>
<td>city</td>
<td>planted</td>
</tr>
<tr>
<td>ago</td>
<td>bit</td>
<td>street</td>
<td>built</td>
<td>lawn</td>
</tr>
<tr>
<td>quiet</td>
<td>remember</td>
<td>games</td>
<td>park</td>
<td>backyard</td>
</tr>
<tr>
<td>spending</td>
<td>part</td>
<td>baseball</td>
<td>restaurants</td>
<td>large</td>
</tr>
<tr>
<td>playing</td>
<td>lived</td>
<td>football</td>
<td>walk</td>
<td>planting</td>
</tr>
<tr>
<td>living</td>
<td>stuff</td>
<td>older</td>
<td>houses</td>
<td>animals</td>
</tr>
<tr>
<td>county</td>
<td>day</td>
<td>built</td>
<td>streets</td>
<td>bushes</td>
</tr>
<tr>
<td>miss</td>
<td>houses</td>
<td>space</td>
<td>nature</td>
<td>oak</td>
</tr>
<tr>
<td>neighbors</td>
<td>great</td>
<td>bikes</td>
<td>living</td>
<td>side</td>
</tr>
</tbody>
</table>

“City”

town
city
town
activities
areas
parks
population
small
buildings
downtown
terms
roads
life
large
community
smaller
describes
growth
past
infrastructure
public

<table>
<thead>
<tr>
<th>“Nature”</th>
<th>“Development”</th>
<th>“Growing up”</th>
<th>“Weather”</th>
</tr>
</thead>
<tbody>
<tr>
<td>land</td>
<td>restaurants</td>
<td>people</td>
<td>bloomington</td>
</tr>
<tr>
<td>wildlife</td>
<td>shopping</td>
<td>time</td>
<td>trees</td>
</tr>
<tr>
<td>area</td>
<td>mall</td>
<td>school</td>
<td>snow</td>
</tr>
<tr>
<td>birds</td>
<td>stores</td>
<td>growing</td>
<td>weather</td>
</tr>
<tr>
<td>trails</td>
<td>north</td>
<td>place</td>
<td>living</td>
</tr>
<tr>
<td>lots</td>
<td>store</td>
<td>family</td>
<td>growing</td>
</tr>
<tr>
<td>property</td>
<td>grocery</td>
<td>grew</td>
<td>animals</td>
</tr>
<tr>
<td>trees</td>
<td>shops</td>
<td>friends</td>
<td>plants</td>
</tr>
<tr>
<td>animals</td>
<td>road</td>
<td>kids</td>
<td>winter</td>
</tr>
<tr>
<td>time</td>
<td>area</td>
<td>home</td>
<td>plants</td>
</tr>
<tr>
<td>deer</td>
<td>built</td>
<td>work</td>
<td>miss</td>
</tr>
<tr>
<td>noticed</td>
<td>food</td>
<td>visit</td>
<td>deer</td>
</tr>
<tr>
<td>home</td>
<td>movie</td>
<td>today</td>
<td>remember</td>
</tr>
<tr>
<td>developed</td>
<td>centers</td>
<td>good</td>
<td>india</td>
</tr>
<tr>
<td>development</td>
<td>hills</td>
<td>high</td>
<td>winters</td>
</tr>
<tr>
<td>enjoy</td>
<td>mom</td>
<td>life</td>
<td>cold</td>
</tr>
<tr>
<td>homes</td>
<td>traffic</td>
<td>walk</td>
<td>ago</td>
</tr>
<tr>
<td>natural</td>
<td>commercial</td>
<td>city</td>
<td>summer</td>
</tr>
<tr>
<td>surrounding</td>
<td>large</td>
<td>beautiful</td>
<td>storms</td>
</tr>
<tr>
<td>bike</td>
<td></td>
<td>live</td>
<td>gardening</td>
</tr>
</tbody>
</table>
interviewing friends and family in their neighborhood so that these students can engage in conversation about changing landscapes, record them, and upload them to the site for others to investigate.

**Future Directions**
The *LandTalk* project aims to continue to expand its user base through new curricular materials and outreach to teachers at all levels. In the next phase, the project will develop a new section of the site that highlights exceptional stories, discusses teacher experiences, provides additional reference material, and reports on the initial analysis of the corpus.

Learn more at *landtalk.org*.

---

**Acknowledgements**
Throughout this summer, CESTA gave me the opportunity to learn new skills and hone old ones while also providing me a space full of curious, inquisitive, and brilliant minds. The work I have done this summer both on the project and on my own has given me confidence to contribute to meaningful projects. This would not have been possible without the help of Cody, whose patience and guidance helped me to get past hurdles in two hours that may have stalled me for two summers, as well as Amanda who has given us access to the wonderful tours, workshops, and snacks that got me through the summer. Erik brought me onto the *LandTalk* project with him and allowed me to focus on the aspects of the project I was most interested in contributing to and learning from.
Planned and Unplanned: Mapping Urban Power in San Francisco

by Krain Chen

Project Team: Ocean Howell, Erik Steiner, Cody Leff, and Brian Kersey

This project traces the history of urban planning in San Francisco, placing special emphasis on unrealized schemes. Rather than using visual material simply to illustrate outcomes, *Imagined San Francisco* uses historical plans, maps, competing design proposals, architectural renderings, and photographs to show what might have been. By enabling users to layer a series of urban plans, the project presents the city not only as a sequence of material changes, but also as a contingent process and a battleground for political power.

In researching, we ask: How can we make these geospatial and historical analyses of San Francisco more accessible to the public?

Significant Proposals

The Burnham Plan (1905) [above] might be the single most important plan in the project. Daniel Burnham is well known as the central planner of the 1893 World's Columbian Exposition in Chicago (also known as the White City). This fair created a miniature city, in today's Jackson Park, that mimicked Paris as it had been remade by Napoleon III under the direction of Baron Haussmann. So impressed were San Francisco officials, that they invited Burnham to make a plan for their own city. The plan presented a re-imagined San Francisco: San Francisco as Paris, San Francisco as a neoclassical imperial metropole. Though no aspect of the plan was actually realized, it influenced everything...
that came after it, serving to condition the range of possibilities that people could imagine. Many ideas that were first proposed in Burnham’s Plan emerged again and again in subsequent planning efforts.

The plans for the Golden Gateway (1960) [above] are also very important. This is an iconic part of the downtown skyline, serving to define the image of the city. The Golden Gateway was also one of the earliest urban renewal plans in San Francisco and so was an important arena in which the relationships between citizens and governmental agencies were renegotiated. Using federal funding, the municipal government here established the capacity to remake large areas of the city, whole cloth, in ways that would have been unimaginable only ten years earlier.

Golden Gateway
Redevelopment Project
This plan by Eichler Homes was one of nine proposals to redevelop and improve quality of life in the area. Source: Eichler Homes, Inc., 1960
An Interactive Site
To resolve our central question of accessibility, this summer we developed a re-design of the *Imagined San Francisco* website. Our challenge was to design for a range of audiences, including researchers, students, and the general public.

We aimed to provide full access and easy navigation to the scholarly narratives and digital map database. The geo-referenced database includes a basemap of the 1906 earthquake and fire burnt area, the Homeowners’ Loan Corporation’s original 1937 redlining map, and public transit plans from several decades.

The new design is more interactive, with dynamically rendered maps focusing the user’s attention. It also allows for both complete user control, with the Proposal Maps List, and more guided experiences, through the various narratives written by Professor Ocean Howell. We hope this website will be used by anyone who has an interest in the history of San Francisco, city planning, politics, and neighborhood activism. We also hope that the georeferenced maps we are making accessible will be used by a broad, cross-disciplinary academic community.

Each map has its own opacity slider, allowing users to interact with the urban plans to their own purpose. These are layered on top of different basemaps, such as modern San Francisco [see left] or a neighborhood boundaries map.

In some areas, the plans are particularly dense, due to the variety of proposed design schemes, coming from a variety of competing political interests. This highlights the ways in which specific geographic areas held outsized importance in determining the future of the entire city.
Re-designed Proposal List
This new design [below] places emphasis on the maps themselves, with the short descriptions and narratives as complementary information that is directly linked to the plans and visuals. There are plans ranging from the mid-19th to the late-20th century. In the previous version of the publication, this diverse range of maps went uncategorized. The urban plans are now grouped by era, demarcating the years when a specific set of planning sensibilities and priorities prevailed. Examples include “Imagining Utopia: An era when San Franciscans dreamed of a perfect city of the future: 1905-1933,” and “The Era of Radical Technocracy: Highways and Urban Renewal schemes sweep away portions of the old city, delivering San Francisco to an uncertain future: 1946-1963.” With such categorization, we make the relationship of each plan to the overall history of San Francisco more distinct. The user is also invited to explore the politics behind each era, some of rapid change and others of minimal development. Rather than simply listing the plans alphabetically, this reorganization better integrates the intellectual goals of the project into the design of the site.
**Address Search Bar**

In particular, we created features, such as an address search bar, for researchers interested in community history or for users simply curious about the site where their home now sits. We also built in functionality to allow users to automatically explore all plans which affected an address.

---

**Narratives**

The narratives provide a more guided experience through the digital publication, providing context and analysis of the various proposals. As the reader follows along with the text, the view updates to show the relevant plans, supplementing the narrative and often allowing the user to draw further connections. We plan to eventually make the scholarly narratives a collaborative effort, where other researchers of San Francisco’s planning history or related fields may submit essays. We hope these narratives will be read by anyone who has an interest in San Francisco history, city planning history, the power of neighborhood activism, and any related topics.

*Imagined San Francisco* hopes to go live in Spring 2020.
Acknowledgments
Many thanks to project lead Ocean Howell for his guidance on the project and support for my ideas, and for providing the Burnham and Golden Gateway plan descriptions above; to Spatial History Project Co-Director Erik Steiner for his encouragement and for providing the site design mock-ups; to Cody Leff for his mentorship and for sharing his programming expertise; to Amanda Wilson Bergado and Brian Kersey at the Center for Spatial and Textual Analysis for their invaluable support this summer; to the past project RIs for their work and notes; to my CESTA RI cohort for making my time on this project so enjoyable; to Kim Durante at Stanford Libraries and Stace Maples and David Medeiros at the Geospatial Center for their help; and to VPUE for funding my work.

Student Profile

Krain Chen
Krain is a sophomore potentially majoring in Computer Science with a minor in Architectural Design. Her two main roles on the Imagined San Francisco Project are to develop the website as part of a team and to georeference undeveloped maps of the city. Last summer, she worked as a background extra in movies and TV shows such as Legends of Tomorrow, so working at CESTA this summer is a new experience.

San Francisco and Stanford University both occupy unceded Ohlone land.
Thomas of Marga’s *Book of Governors*: Insights into Early Christians under Islamic Rule

by Palmer Manes and Zuyi Zhao

Project Lead: Michael Penn

In the *Early Christians under Islamic Rule* project, we created a dynamic and big-picture understanding of the *Book of Governors* by using digital humanities tools, while developing a model to apply these methods to future research. Written by Thomas of Marga, the *Book of Governors* is a ninth-century Christian text that spans over four hundred years and documents the lives of monks who lived in the Monastery of Beth Abhe, and the history of the Church of the East (an Eastern Christian denomination, located primarily in present day Iraq). Through applying social network analysis, we identified the text’s assumptions and attitudes toward different people, investigated distinct characters and their relationships with one another, applied geospatial analysis, and created a geographic database. We focused on creating five distinct visualizations: a heat map, district map, timeline visualization, monastery social network visualization, and a standard distance circles map to accomplish these goals.
The Early Christians under Islamic Rule project, we created a dynamic and big-picture understanding of the Book of Governors by using digital humanities tools, while developing a model to apply these methods to future research. Written by Thomas of Marga, the Book of Governors is a ninth-century Christian text that spans over four hundred years and documents the lives of monks who lived in the Monastery of Beth Abhe, and the history of the Church of the East (an Eastern Christian denomination, located primarily in present day Iraq).

Through applying social network analysis, we identified the text’s assumptions and attitudes toward different people, investigated distinct characters and their relationships with one another, applied geospatial analysis, and created a geographic database. We focused on creating five distinct visualizations: a heat map, district map, timeline visualization, monastery social network visualization, and a standard distance circles map to accomplish these goals.

Next Steps
Foremost, the relationship spreadsheets requires further fine-tuning in terms of node attributes and edge categories, in order to cultivate a database that can answer complex questions about the Book of Governors. Visualizations of the social networks of the text’s marginal groups, such as women and non-Christians, can provide further insight into the world of Thomas of Marga. Visualization of specific important journeys made in the text could be useful, especially for the purposes of comparison.

Acknowledgments
We would like to acknowledge and thank Amanda and Brian, and CESTA, for allowing us the space and resources needed to complete this project. David Medeiros and Stace Maples at the Stanford Geospatial Center gave invaluable advice for the geospatial component of the project, especially to work with QGIS and ArcGIS. Rachel Midura and Quinn Dombrowski provided helpful guidance for the use of the social network visualization program Gephi. We would also like to thank Dr. Michael Penn, our project lead, for allowing us to work on such an interesting project and for being such an invaluable resource for us.

Methods
To compile social networks and geospatial databases, we mined the text for names of people, locations, and instances of interactions between characters and places. We then logged our findings in spreadsheets. The social network database comprised two main spreadsheets: a master list of all characters (nodes) and another containing interactions and relationships (edges) in the Book of Governors. From these lists, we noted node attributes (e.g., gender) and edge categories (e.g., familial relationships) allowing for more dynamic forms of social network analysis. We used Gephi to calculate statistics and produce visualizations of the interactions and relationships, using concepts like degree and betweenness centrality to identify hubs and clusters of interaction and to pinpoint the most important nodes.

For the geospatial component, we relied primarily on ArcGIS, QGIS, and Tableau. We used QGIS to create both heat maps and district maps by both occurrence (a unique singular mention of a place) and frequency (the number of times a location is mentioned over the course of the text). Drawing on secondary sources, we then drew district maps, which represent important ecclesiastical boundaries to visualize locations for which we knew a region but not exact coordinates. We used Tableau to create a timeline visualization and ArcGIS to create a standard distance circle map.

Findings
We discovered that Thomas includes a surprising number of names as his written sources, providing material for a potential network of literature. We also found that very few women are present in the Book of Governors, comprising approximately five percent of all four hundred seventy nodes. Figures like God, Satan, and Nestorian Patriarchs played varyingly influential roles, depending on what statistics we used. Geographic results remain very preliminary, but we found intense clustering in the Adiabene/Mosul region, particularly between the Upper and Lower Zab Rivers. The text almost never mentions the Roman Empire or the Greco-Roman world, besides passing references to Athens and Turkey. Egypt, the birthplace of Christian asceticism, only appears a few times. Nestorian activity further east, in China and India, is described vaguely, mainly referred to as two amorphous ecclesiastical constructs.
Next Steps
Foremost, the relationship spreadsheets requires further fine-tuning in terms of node attributes and edge categories, in order to cultivate a database that can answer complex questions about the Book of Governors. Visualizations of the social networks of the text’s marginal groups, such as women and non-Christians, can provide further insight into the world of Thomas of Marga. Visualization of specific important journeys made in the text could be useful, especially for the purposes of comparison.

Acknowledgments
We would like to acknowledge and thank Amanda and Brian, and CESTA, for allowing us the space and resources needed to complete this project. David Medeiros and Stace Maples at the Stanford Geospatial Center gave invaluable advice for the geospatial component of the project, especially to work with QGIS and ArcGIS. Rachel Midura and Quinn Dombrowski provided helpful guidance for the use of the social network visualization program Gephi. We would also like to thank Dr. Michael Penn, our project lead, for allowing us to work on such an interesting project and for being such an invaluable resource for us.

Student Profiles

Palmer Manes
Palmer Manes is a junior (Class of 2021) majoring in Religious Studies. He is working on the geospatial component Book of Governors project, which involves assembling a dataset of every geographic location in the text and finding ways to visualize this dataset and the geographic interrelations within it.

Zuyi Zhao
Zuyi Zhao is a junior (Class of 2021) majoring Philosophy & Religious Studies. With a focus on the interactions between the various characters of the Book of Governors, Zuyi is responsible for the analysis and visualization of the social networks within the text, exploring the ways in which different and similar communities and identities intersect.
Understanding Book Awards through Big Data

by Shana Hadi
Project Leads: Laura McGrath, J. D. Porter, and Mark Algee-Hewitt

When awards have varying selection criteria and voter composition—often distinguished on the basis of what’s “popular” and “prestigious” in a given year—what do these awards really measure? What do these awards say about their stories? In speculative fiction, the two most well-known awards, the Hugo and Nebula, offer perspective to this research question. Presented within a yearly convention, the Hugo and Nebula Awards unite the larger gathering of fans and authors through a collectively chosen “new canon.” Beyond honoring the stories with the greatest print and literary reach, they reflect the tastes of their voters through canonizing the “best of the year.”
Any reader willing to purchase a Convention membership can vote in the Hugos, and I wondered if they reward popular works that consciously adopt speculative fiction tropes, while the Nebulas, which require voters to be published authors, would have a literary bent and reward works with more experimental qualities. Would there be an increasing diversity of topics and characters in the more recent selections, especially as more female authors join the lists? And are these themes consistent or volatile?

I collected 194 Hugo- and Nebula-nominated short stories spanning 20 years from 1999-2018. This corpus has over 925,000 English-language words from 123 unique authors, with 99 Hugo stories and 126 Nebula stories. Out of these, 30 stories have been cross-nominated for the Hugo and Nebula, with three stories having won both. I found that the Nebula stories had more occurrences of “she,” “shes” (she’s), “her,” “herself,” “woman,” and “women,” with a high degree of certainty. The word “shed” in this case would most likely refer to “she’d” (with the apostrophe removed), but the remaining words suggest that the Nebula short stories are more likely to feature female main characters, especially with the collective results of these words and their high observed/expected scores.
Female Pronoun Usage in Hugo and Nebula 1999-2018 Nominated Works

<table>
<thead>
<tr>
<th>Token</th>
<th>Corpus</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>her</td>
<td>Hugo</td>
<td>5,536</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>6,867</td>
</tr>
<tr>
<td></td>
<td>Nebula</td>
<td>8,518</td>
</tr>
<tr>
<td>she</td>
<td>Hugo</td>
<td>5,566</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>6,612</td>
</tr>
<tr>
<td></td>
<td>Nebula</td>
<td>7,855</td>
</tr>
<tr>
<td>woman</td>
<td>Hugo</td>
<td>301</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>360</td>
</tr>
<tr>
<td></td>
<td>Nebula</td>
<td>431</td>
</tr>
<tr>
<td>shed (<em>she’d</em>)</td>
<td>Hugo</td>
<td>263</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>305</td>
</tr>
<tr>
<td></td>
<td>Nebula</td>
<td>353</td>
</tr>
<tr>
<td>herself</td>
<td>Hugo</td>
<td>237</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>Nebula</td>
<td>379</td>
</tr>
<tr>
<td>shes (<em>she’s</em>)</td>
<td>Hugo</td>
<td>208</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>274</td>
</tr>
<tr>
<td></td>
<td>Nebula</td>
<td>361</td>
</tr>
<tr>
<td>women</td>
<td>Hugo</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Expected</td>
<td>113</td>
</tr>
<tr>
<td></td>
<td>Nebula</td>
<td>143</td>
</tr>
</tbody>
</table>

This figure presents the ratio of female pronouns in the short stories nominated for the Hugo (orange) to the Nebula (green), found through the technique of most distinctive words. The expected values are denoted in green.

Why the Nebulas over the Hugos? Perhaps it’s the composition of the fanbase. It’s possible that members of Hugo conventions may skew towards men raised on science fiction. Most pre-2010 Hugo-nominees come from three magazines established in the mid-twentieth century: Asimov’s Science Fiction, Fantasy and Science Fiction, and Analog Science Fiction, which have varying generations of readers. Any fan willing to purchase a membership can vote, and the fanbase has historically leaned towards white men interested in science.

In contrast, the composition of active members of the SFWA may be more well-balanced, with author-voters who will pay greater attention to stories that experiment with the craft of writing. At the very least, the recent Nebula-nominees have a significant proportion of female authors, and five female authors have won the Nebula in a row from 2013-2018, and 15 out of 20 Nebula-winners were written by women. The observed distinctions between the Hugo and Nebula Awards for Best Short Story may reflect different compositions of their audiences and tastes.
Next steps
Next steps could explore how these stories—the speculative fiction “canon”—would compare to those canonized from the realist tradition, such as nominated works for the O. Henry Prize or Pen/Hemingway Award. Which work is the “best” continues to be up for question, especially as tastes and themes change to reflect and challenge society. We may not know what makes the canon until fifty years later, but until then, at least we’ll have lots of excellent reading.

Acknowledgements
Many thank yous to Literary Lab’s Laura McGrath and J. D. Porter for their generous mentorship, guidance, and fascinating research stories. I also want to thank Green Library curator Dr. Rebecca Wingfield for helping me find several scarce short stories, Mark Algee-Hewitt and Nichole Nomura for their thought-provoking “Novel World-Building: Science Fiction” presentation, and CESTA and Literary Lab for the opportunity to work on exciting projects throughout the summer.

Student Profile
Shana Hadi
Shana is a junior majoring in English and Computer Science and contributed to various Literary Lab projects this summer. She especially enjoyed studying concepts of popularity and prestige, as well as the effects of translation on short stories. When she is not reading books, she helps manage Arts&Life at the Stanford Daily, writes creative fiction, wanders cities, and drinks a staggering amount of tea.

This article is an excerpt from a longer post that will appear on the LitLab Techne Blog at litlab.stanford.edu/techne/.
The challenge any translator faces is telling a story in a language that is not a product of the same culture, history, and circumstances that the original text came from. This results in a “fingerprint” or “signature” of the original (“source”) language that the translator often unconsciously carries over to the destination language (to clarify, this is distinct from mistranslation, which is simply an error in understanding the meaning or context of the text in its source language). Over summer, the Literary Lab team focused primarily on the syntactic branch of this question: are there subtle yet significant differences in the way words and parts of speech are used in translated short stories, and do those differences affect the reader’s experience of the text?

Fingerprints aren’t necessarily simply syntactical, though. A country’s culture and customs determine what and how its people need to communicate to each other; in turn, those needs shape their language. What signatures exist in non-Western languages? Whose speakers have lived a history and culture more removed from America and Britain? And where do those traces stem from? I chose to work with Korean-to-English texts this summer.

In Korea, showing respect to elders and those in positions of authority is such a priority that it is ingrained in the very way its language is structured. Do these features manifest in English translations of Korean works, and do the English...
translators accurately convey the complex and unique relationships between characters of different ages and ranks in Korea?

**Findings**
Recently, the influence of translations to and from European languages has resulted in the creation of an equivalent for “she”: 그녀 / 그녀들 (geunyeo / geunyeodeul). Because both 그 and 그녀, as third-person pronouns, are not commonly used, and 그녀 has only been around relatively for a relatively short time, I expected the prevalence of gendered pronouns to be higher in original English language short stories than in translated ones. Interestingly, “she,” “her,” and “he” were all tagged as distinctive words; “she” and “her” appear more in English, but “he” appears more in Korean. The effect of using titles over names is also reflected in the part-of-speech tags. Singular proper nouns (which would include names) appear 1.24 times more in English short stories than in Korean translated ones (p-value of 1.233e-64).

**Next Steps**
The ultimate end goal for this project could be an “atlas” of world literature: a study of how signals from translation between languages affects foreign readers’ understanding of that language.

---

**Acknowledgements**
I’d like to give a huge thank you to the Literary Lab and Dr. Laura McGrath and Dr. J.D. Porter for their mentorship and guidance throughout the summer. I would also like to thank Quinn Dombrowski and everyone else working on the translation project for their advice and input and CESTA for supporting my research.

This article is an excerpt from a longer post that will appear on the LitLab Techne Blog at litlab.stanford.edu/techne/.

---

**Student Profile**

Eunji Lee

Eunji Lee is a sophomore that is planning to major in Computer Science. She is interning for the Literary Lab and is contributing to various projects. She is currently working on the translation project, aiming to compare the syntax and themes of English short stories to those that have been translated from other languages. She has also helped identify literary agents and the authors they represent and has summarized various definitions of popularity for the Lit Lab’s other projects.
Acknowledgments

Thanks to Professors Jesse Rodin and Craig Sapp for providing me with the great opportunity to work with them on this amazing project; to Brian Kersey and Amanda Wilson Bergado for being so supportive of the CESTA internship cohort; and to Timothy Karoff, Palmer Manes, and Michelle Julia Ng for helping develop Jellybones right at CESTA.

Learn more at josquin.stanford.edu

The Josquin Research Project (JRP) is an open-access website that hosts a growing collection of complete scores from Renaissance vocal music. As a repository for these scores, the project will make the music fully searchable through melody and rhythm. The JRP also provides analytical tools to research individual works, the style of composers, and musical lingua francas. Our goal is to bring “big data” into conversation with traditional analytical methods.

Understanding Renaissance Music Through Density and Dissonance

by Young Fenimore Lee

Project Leads: Jesse Rodin and Craig Sapp

The Josquin Research Project (JRP) is an open-access website that hosts a growing collection of complete scores from Renaissance vocal music. As a repository for these scores, the project will make the music fully searchable through melody and rhythm. The JRP also provides analytical tools to research individual works, the style of composers, and musical lingua francas. Our goal is to bring “big data” into conversation with traditional analytical methods.
This summer, I worked with research leads Professor Jesse Rodin and Professor Craig Sapp to integrate current analytic tools with the website interface. Specifically, I developed the note tree tool, ribbon tool, and began development of a melodic search tool.

Note Tree
The note tree tool enables researchers to visualize note sequences from a large corpus of scores. Then, they can narrow criteria based on composer and genre to create a tree displaying the relative frequencies of each note, following the root note. If the tree depth is greater than one, the tree will iterate recursively, treating each child of the root note as another root note.

Researchers can also use this tool to create visual analyses of the relative frequencies of certain note sequences. For example, among all of the works by Ockeghem on the site, from the root note C, the ascending note sequence C-D-E is overwhelmingly the most common. A short glance at any note tree with a depth of at least two can identify note sequences of study interest within a certain composer’s body of work and statistical differences between musical genres within a composer’s corpus.
**Acknowledgments**

Thanks to Professors Jesse Rodin and Craig Sapp for providing me with the great opportunity to work with them on this amazing project; to Brian Kersey and Amanda Wilson Bergado for being so supportive of the CESTA internship cohort; and to Timothy Karoff, Palmer Manes, and Michelle Julia Ng for helping develop Jellybones right at CESTA.

Learn more at josquin.stanford.edu

---

**Ribbon Tool**

The ribbon tool shows the respective ranges of pitch and attack density performed by voices in a work through a visualization of the melodic contours and attack densities of specific musical scores.

The ribbons can be displayed separately or overlapping. Overlapping ribbons are useful in visualizations that demonstrate how voices begin to merge ranges during specific sections of a piece.
Next Steps
We are currently developing a melodic search tool for general note intervals. We are also expanding the ribbon tool to display more musical parameters simultaneously: rhythmic activity, melodic ranges, and musical "arrivals" (similar to commas and periods in prose). Because arrivals are to some extent subjective, we are developing an interface through which users can enter and display their own analyses in real time.

Acknowledgments
Thanks to Professors Jesse Rodin and Craig Sapp for providing me with the great opportunity to work with them on this amazing project; to Brian Kersey and Amanda Wilson Bergado for being so supportive of the CESTA internship cohort; and to Timothy Karoff, Palmer Manes, and Michelle Julia Ng for helping develop Jellybones right at CESTA.

Learn more at josquin.stanford.edu

Student Profile
Young Fenimore Lee
Young is a junior and prospective English major. He is the son of two Korean musicians and was born and raised in the Chicago metropolitan area. He identifies as queer and loves indie rock, indie folk, emo, post-hardcore, and math rock. His main project is his music journalism publication Jellybones.

Learn more at cesta.stanford.edu
Global Medieval Sourcebook
A Digital Repository of Medieval Texts

An Online Portal to the Middle Ages
by Nina Du and Tina Zhang
Project Leads: Kathryn Starkey and Mae Lyons-Penner

While working on the Global Medieval Sourcebook project, we challenged two major barriers for the use of medieval documents by professional scholars, aspiring scholars, and interested non-scholars: language and access. This summer, our work has mostly consisted of three parts: transcription, translation, and web design for the forthcoming collections of medieval Chinese and Latin texts.

The first barrier to studying medieval sources is that they are only accessible in their original...
languages. Another barrier is that the manuscripts are handwritten in many different scripts with large numbers of abbreviations. Even with skills in these languages, many scholars will find them illegible. And while more and more medieval sources are being made available digitally, most online repositories lack contextual and introductory information, and are intended only for specialists. Our additions of contextual information will make these sources more available to scholars, students, teachers, and the interested public.

Our goal of the *Global Medieval Sourcebook* is to create a free, open access, and open source teaching and research tool for the simultaneous viewing of medieval texts in their original language, in English translations, and as digitized manuscripts.

One of the greatest difficulties we encountered in transcribing these Latin manuscripts was the diversity of different abbreviations and scripts. We consulted Medieval Studies librarians and our project leads to decode the abbreviations. With their help, we were able to present a complete transcription. Our most important contribution has been compiling a collection of song lyrics (*ci*) from the Song Dynasty (960-1279) in medieval China. We transcribed and edited the Chinese versions of these lyrics for digital publication and for them to be as literal and fluent in English as possible.
possible. One of the biggest challenges in translating these lyrics was the differences in modes of expression between English and Chinese. If we had only translated these lyrics literally, many of them would not make complete sense because of the extensive employment of metaphors in the Chinese language and certain cultural differences. In order to counter this problem, we attached extensive notes to each lyric to explain what the versions convey. In addition, we wrote an introduction to the lyrics to facilitate audience understanding of the cultural context and historical background.

In order to make the website more approachable, we are currently working to create maps with routes marking the *ci* poets’ travel as complements to the general introductions to specific authors. This will allow users of the website to more easily visualize the geographical distribution of different works. One of the most important tasks ahead is to continue the publishing process for the *ci* collection. Once we have finalized the texts, the next step is to encode them in XML-TEI. After we have successfully encoded everything, we will publish these lyrics in bundles to the website.
Acknowledgements
We would like to express our greatest appreciation to our project leads, Mae Lyons-Penner and Professor Kathryn Starkey, who have helped us edit our work and gave directions on what we should focus on when working through the large corpus of Song Dynasty *ci* lyrics. In addition, we would also like to thank Dr. Benjamin Albritton, Rare Book Curator for Special Collections with the Stanford Libraries, for his generous help on explaining Latin abbreviations. Last but not least, we would like to express great appreciation to Professor Ronald Egan for his useful suggestions regarding the *ci* collection.

The *Global Medieval Sourcebook* project can be found at sourcebook.stanford.edu.
Stanford Ordinary People Extraordinary Stories (SOPES for short) is a project concerned with preserving ordinary lives of the past, the stories of “normal people” that would otherwise be forgotten. The project is interested in the mundane and the prosaic rather than the glamorous. It emphasizes the importance of safeguarding personal archives like collections of letters, postcards, diaries, and other ephemera. Without individual archives and the desire to retain our own intimate histories, mainstream archives risk becoming lackluster and monolithic. The project therefore seeks to diversify what we define as an archive and validate individual attempts to write history.
SOPES is inspired by unknown people of the past. In our research, we ask, "Who knew whom, and how?" The immediate goal of the project is to create a database of digitized personal ephemera with all the relevant metadata like the sender’s and receiver’s name and address as well as the details on the postmark. After creating databases on for each group, the project investigates the real-life connections between each individual by generating family trees and social networks.

Social Networks within the Frahm Family.

Project lead, Professor Elaine Treharne, collected ephemera to create a SOPES archive. I took the lead in transcribing and digitizing the materials and sorting them by family and organization. Among my samples, I categorized two chief groups: the ephemera of the Frahms family and a collection of autograph books. Category 1 was broken up into two parts, one for the Frahms’ letters and one for Jean Woods’ diary. Category 2 was broken up into three parts, one for Jessie Thorson’s autograph book, one for Martha’s, and one for Miss F. D.

Jean Frahm’s Diary. Source: Treharne Collection
Narrative of the Frahms Family

The heads of the family were Laura and Henry Frahm, and they had 3 daughters and 5 sons: John, Lena, Arthur, Ray, Ida, Florence, William, and Ted, in that order. The children were born in Idaho; the evidence suggests that Florence, Ida, and some of their siblings attended college in Gooding during the time these letters were written. Their parents, Laura and Henry, lived in Salt Lake City at the time.

The letters from the 1930s suggest that Laura headed a failing livestock business (alluding to the Depression during that time) and that Henry’s health was consistently poor. The cycle of letters from the 1920s reveals that Ted either left or wholly did not attend college to help his mother take care of his father. He was not altogether happy about this arrangement; he would write his sisters frequently (often promising to write them daily) and would ask them to tell his close friend, Bert J. Worden, to write him as well.

A tiresome inside joke among the family is the amount of divorces they appear to go through—Ted seems to be the newest member of this club. Jean’s diary is captivating—her progression through her early years of college and through her increasingly serious relationship with Lee Woods is fascinating. On January 1, 1951, Jean proclaims, “This is Lee’s and my year. We are going to make it as perfect as we can—resolve ‘not to quarrel any more.’” This resolution evidently does not hold for very long; as I progressed through the diary, almost every entry mentions that Lee was often angry with her for no reason. The tone of the entries as the year goes on becomes ominously forlorn.

In January, Jean seems thrilled about the holidays, her family, and her boyfriend. She seems depressed about going back to university—not an uncommon sentiment. However, school soon becomes the least of her worries as she slips into a crisis about her relationship with Lee. Evidently, her marriage was decided before the start of the year, and there was no getting out of it. Even on January 12, just nine days after the last time she declared her undying love for Lee, she was “befuddled” and torn about her decision to marry him. A few of her friends attempted to convince her out of it, simply entrapping her further in her pit of confusion. She does, of course, go through with the wedding on June 24, 1951.

The latter half of her diary is sad. Jean slowly begins adjusting to her life as a housewife, stuck cooking and cleaning for a man who does not often satisfy her, emotionally and physically. All along, her mind revolts against her situation, but her thoughts, having no power, mutate into a mild but desperate melancholia. Many entries terminate with Jean mentioning how low she felt during the day, mostly because of Lee’s irrational anger. She seems mired in the mold of the Beauvoirian housewife.
Young’s. I indexed every entry and letter, then I catalogued them by date, location, writer, and transcription. Afterwards, I made a social networks for each group in the form of detailed databases like family trees.

From our research into their correspondences, I crafted a narrative of the Frahms family [facing page]. I studied a diverse range of primary sources from 19th and 20th century America including a set of 30 letters from the 1920s, a diary from 1951, and three autograph books from the 1880s, 1890s, and 1900s.

Next Steps
Our next steps for this project is to create relational databases. Later, we will place these sources in conversation with one another through a website displaying all our databases.

Acknowledgements
I would like to thank Professor Elaine Treharne for allowing me to aid in her in her journey of cataloguing her materials. I would also like to thank the staff and team at CESTA for constantly offering help, expertise, and community. 

Student Profile
Karunya Bhramasandra is a sophomore planning to major in English and minor in Global Studies with a specialization in South Asian Studies. Her avid interest in personal writing fuels her passion to transcribe and form connections between the letters and their writers for her research project, SOPES. In her time off, she goes on drives, starts reading books but fails to finish them, and Facetimes friends from her hometown.

Karunya Bhramasandra

Student Profile
Karunya Bhramasandra is a sophomore planning to major in English and minor in Global Studies with a specialization in South Asian Studies. Her avid interest in personal writing fuels her passion to transcribe and form connections between the letters and their writers for her research project, SOPES. In her time off, she goes on drives, starts reading books but fails to finish them, and Facetimes friends from her hometown.

Karunya Bhramasandra

Student Profile
Karunya Bhramasandra is a sophomore planning to major in English and minor in Global Studies with a specialization in South Asian Studies. Her avid interest in personal writing fuels her passion to transcribe and form connections between the letters and their writers for her research project, SOPES. In her time off, she goes on drives, starts reading books but fails to finish them, and Facetimes friends from her hometown.
Our goal in the Mapping Manuscripts project is to map the production of medieval manuscripts through digital methods. Traditional mapping techniques for manuscript production are problematic because scholars know little about specific dates and locations of production. While scholars can approximate the date and production location of a manuscript through close examination, their conclusions can be as broad as “Germany,” or “9th or 10th century.” While pinpoints on a map can’t accurately represent this kind of fuzzy metadata, it can be useful when applying digital mapping methods. Mapping Manuscripts experiments with different methods of mapping.
imprecise manuscript metadata for creation, for research and teaching tools, and for more precisely identifying locations of copying hubs for medieval manuscripts. The inspiration for this project came in part from the work of the cool people who copied manuscripts a thousand years ago; we want to understand them and their work in our research, and to trace their legacies by investigating how their handwritten manuscripts spread across Europe.

My project lead Mateusz Fafinski (Freie Universität Berlin) and I broke the project down into two main parts: gathering metadata and computationally mapping that metadata. I worked on the first part, collecting manuscript metadata. Since we are only need information about the manuscripts to map them, I didn’t need to look at the manuscripts themselves. Instead, I worked with catalogues of medieval manuscripts that contained manuscript metadata. I mainly used metadata from Bernhard Bischoff’s catalogues of ninth-century
Our goal in the Mapping Manuscripts project is to map the production of medieval manuscripts through digital methods. Traditional mapping techniques for manuscript production are problematic because scholars know little about specific dates and locations of production. While scholars can approximate the date and production location of a manuscript through close examination, their conclusions can be as broad as “Germany,” or “9th or 10th century.” While pinpoints on a map can’t accurately represent this kind of fuzzy metadata, it can be useful when applying digital mapping methods.

Mapping Manuscripts experiments with different methods of mapping imprecise manuscript metadata for creation, for research and teaching tools, and for more precisely identifying locations of copying hubs for medieval manuscripts. The inspiration for this project came in part from the work of the cool people who copied manuscripts a thousand years ago; we want to understand them and their work in our research, and to trace their legacies by investigating how their handwritten manuscripts spread across Europe.

My project lead Mateusz Fafinski (Freie Universität Berlin) and I broke the project down into two main parts: gathering metadata and computationally mapping that metadata. I worked on the first part, collecting manuscript metadata. Since we only need information about the manuscripts to map them, I didn’t need to look at the manuscripts themselves. Instead, I worked with catalogues of medieval manuscripts that contained manuscript metadata. I mainly used metadata from Bernhard Bischoff’s catalogues of ninth-century medieval manuscripts, which contained information on the current location, production location, date, and signature of thousands of ninth-century manuscripts. From these catalogues, I digitized the metadata on every manuscript they mentioned by Augustine, Bede, Jerome, and Orosius. I manually found each relevant entry in PDF copies of the catalogues and copied the data of each entry into a big database. I also experimented with collecting manuscript metadata from the National Library of France’s online manuscript archive. I worked on building a program that could automatically gather this data from the archive for our database. To map all this metadata, Mateusz wrote a program that would process it and generate a heat map of manuscript production hubs in medieval Europe. Going forward, Mateusz and I will continue to make maps from the data we have already gathered. We will also continue to add more manuscript data from other databases, either manually or using web-scraping.

**Acknowledgements**

I’d like to thank the CESTA team for their support on this project. Our work was thanks to the help of Elaine Treharne, Amanda Wilson Bergado, Brian Kersey, Quinn Dombrowski, Scott Bailey, GP LeBourdais, and Erik Steiner. I’d also like to thank my project lead, Mateusz Fafinski, for the opportunity to work on this project this summer, guidance through many hurdles, and for the tools and encouragement to do digital humanities research of my own. I’m also grateful to all of my coworkers for making working at CESTA a fun time.

A code snippet showing the scraping script developed to automatically enrich the database with the information contained in the online catalogue of Bibliothèque Nationale de France.
Student Profile

Timothy Karoff

Timothy is a sophomore working on the Mapping Manuscripts project. He is originally from Massachusetts, and he’s considering majoring in either Science, Technology, and Society or Sociology. His interests include the politics of historical narratives, vegan snacks, the study of online communities, electronic music, giraffes, and Super Smash Bros.

A map showing the distribution of manuscripts containing letters of Jerome, an important 4th century Christian writer and translator. The map clusters data points in a way that allows for inclusion of non precise locations, thus showing areas were those manuscripts were of particular interest.
Poetic Media Lab
Amir Eschel, Cody Chun, Fyza Parviz, and Nelson Schumacher

The Poetic Media Lab is a digital humanities research collective housed at CESTA. Current projects include MaPaLiterature, a project to “visually map the circulation, international movement, and (im)mobility of Palestinian authors and texts,” and Poetic Thinking, a digital intervention, at the level of pedagogy, which seeks to facilitate creative forms of thinking. The lab also collaborates with research units at other institutions: currently, it is working with a team at Bucknell University on a digitization of Bernardino de Sahagún's Historia general de las cosas de Nueva España. Learn more at poeticmedia.stanford.edu.

Chinese Railroad Workers Project
Gordon Chang, Shelley Fisher Fishkin, Roland Hsu, and Erik Steiner

The Chinese Railroad Workers Project has been active throughout summer 2019. We have engaged scholars and the general public interested in the Chinese worker experience with our large corpus of new print and digital publications. As we begin our final year of activity, we have used the summer to post on our website the research findings and first person testimonies of our CESTA research interns. In the coming year, we will complete our work in partnership with CESTA and Stanford Libraries, to accession and curate our limited selection of rare and hard to access original historical materials, as well as the project papers, that we anticipate will be invaluable for future researchers. These materials, many of which originate in our work with CESTA RIs, will be available in perpetuity in the Stanford Library Chinese Railroad Workers Project collection. Learn more at chineserailroadworkers.stanford.edu.
Corpus Synodalium
Rowan Dorin, Thawsitt Naing, and Clara Romani

Throughout the later Middle Ages, bishops across Latin Christendom promulgated legislation to guide the clergy and instruct the faithful. Until recently, the extent and dispersion of the surviving sources has made impossible for scholars to explore them systematically. With CESTA’s support, the project team has now transcribed more than a thousand texts, gathered them into a flexible online database (powered by ARTFL’s Philologic4), and integrated them into the first-ever digital atlas of medieval Christian Europe. Going forward, we’re excited to use these new digital tools to explore the ways such local legislation challenged the centralizing efforts of the contemporary papacy. Learn more at corpus-synodalium.com.

Modernist Archives Publishing Project
Alice Staveley, Claire Battershill, Nicola Wilson, Helen Southworth, and Elizabeth Willson Gordon

The Modernist Archives Publishing Project (MAPP), a CESTA-supported DH project digitizing and enriching with metadata thousands of documents related to the history of twentieth century book publishing in the UK, USA, and Canada, concluded the data capture for a major subsidiary project within MAPP on the financial account books that record all sales of Hogarth Press publications between 1917 and 1941. CESTA’s stellar research interns over the past two years—Peter Morgan, Emily Elott, Victoria Ding, and Khuyen Nha Le—have spent countless hours transcribing the handwritten mid-century records of book sales of the Press famously founded in 1917 by Leonard and Virginia Woolf. Khuyen this past summer finished crucial work on an Order Book (the name of the MSS as they are catalogued at their home institution, the University of Reading, UK) we had thought ‘lost’ but which held vital records allowing us to complete the transcription of all Virginia Woolf’s sales. Now we can move forward with final data visualizations of the results, and begin writing up our findings for publication. Learn more at modernistarchives.com.
Mapping Ottoman Epirus
Ali Yaycioglu, Antonis Hadjikyriacou, Fatma Öncel, Selma Koroglu, and Erik Steiner

The Mapping Ottoman Epirus team worked on Ottoman-Turkish and Greek documents on Epirus and tried to solve the puzzle about how imperial and vernacular/regional archives coexisted and interacted in the Ottoman context. With the help of David Medeiros at the Stanford Geospatial Center, we also developed a regional study of slope and elevation. The next stage in the MapOE project will be to convey various questions about multilingualism, imperial and regional scribal cultures, and archives, and archiving into the digital platform. Learn more at mapoe.sites.stanford.edu.

Oral History Text Analysis Project
Estelle Freedman, Natalie Marine-Street, Katie McDonough, and Preston Carlson

Preston Carlson, Research Intern for the Oral History Text Analysis Project, made substantial progress this summer preparing and analyzing digitized transcripts of nearly 3,000 women’s oral history interviews to help understand the language used and meanings attached to sexual violence and sexual harassment. His tasks involved integrating transcripts and metadata from The HistoryMakers (a database of African-American leaders), running and modifying a tool to generate subcorpora based on keyword searches, and qualitatively coding interviews in NVivo.

Percent of HistoryMakers Women using terms
like tour. science. stories. project. text. manuscripts. notes. tree. map. cesta.