Unraveling The Complexity: A User-Centered Design Process For Narrative Visualization

Nil Tuzcu
nil_tuzcu@hks.harvard.edu
Harvard University
Cambridge, MA, USA

Annie White
annie_white@hks.harvard.edu
Harvard University
Cambridge, MA, USA

Brendan Leonard
brendan_leonard@hks.harvard.edu
Harvard University
Cambridge, MA, USA

Steven Geofrey
steven_geofrey@hks.harvard.edu
Harvard University
Cambridge, MA, USA

ABSTRACT

In this case study, we introduce a user-centered design process for developing Metroverse, a narrative visualization platform that communicates urban economic composition and growth opportunities for cities. The primary challenge in making Metroverse stems from the complexity of the underlying research and data, both of which need to be effectively communicated to a wide range of end-users with different backgrounds. To unravel the complexity of the research, and to design the platform, we followed a user-centered design process. Our design process brought together researchers, designers, and various end-users, who collectively guided the design of the narrative visualization. Engaging end-users in the early phases of the project allowed us to identify the valuable insights in the data and subsequently design effective visualizations that convey those insights. We believe findings from our process can provide a template for similar projects that require translating complex research data and methodologies into user-friendly story structures.

CCS CONCEPTS

• **Human-centered computing** → **Visualization**; **User centered design**; **User studies**.

KEYWORDS

data visualization, narrative visualization, storytelling

1 INTRODUCTION

Harvard Growth Lab's mission is to understand how economies grow and to translate these findings into policies, publications, and tools. For more than a decade, Harvard Growth Lab's flagship online tool, the Atlas of Economic Complexity[1], has been instrumental for policymakers in identifying a country's paths of growth and diversification. Increasingly, however, researchers and policymakers are seeking to understand not only how nations grow their economies but also how cities and towns can too. As local-level data availability has improved in recent years, the Growth Lab has tested its paradigms in cities to determine how growth happens at the sub-national level.

The Lab's new online tool, Metroverse[2], emerged from this desire to understand urban economic composition and growth opportunities in cities. Developed by the Growth Lab between 2020 and 2021, Metroverse allows users to explore the industrial composition of more than 1,000 cities worldwide. In doing so, a city's know-how is made "visible" through a series of interactive visualizations. Metroverse builds on this information to provide customized recommendations for growth, diversification, and comparison.

This paper introduces the steps taken to create Metroverse, including product discovery, user research, and design phases. By doing so, we seek to provide important insights into creating a user experience design for data exploration platforms that involve complex scientific research and multidisciplinary stakeholders.

2 PROBLEMS

The scientific methodologies behind Metroverse were developed over a decade and are based partly on the methodologies found in the Growth Lab's Atlas of Economic Complexity [4]. Metroverse data are derived from Dun and Bradstreet's World Base[3] and provide information for more than a hundred million establishments around the globe. The Growth Lab team[4] cleaned and aggregated these data into the database that powers Metroverse, with the final goal of describing a city's unique possibilities for economic growth.

Metroverse's research and data are scientific and combine complex facets from multiple disciplines of academic study, such as economics, policy studies, and urban development. Such complexity means that analysis, outcomes and insights, however useful, can nonetheless be inaccessible or poorly understood by a broad, public audience. Previous studies tackling similar issues suggest

---

[1] atlas.cid.harvard.edu
[2] metroverse.cid.harvard.edu
[3] dnb.com
[4] From here on, we will refer the Growth Lab Design and Development team as our team and Growth Lab Applied and Academic Research team as the research team.

Figure 1: Metroverse user interface. In this paper, we highlight the following parts: (1) The five-part narrative structure, (2) data visualization, (3) narrative text, (4) visualization options, and (5) comparison feature visual language.

that interactive visualization and storytelling techniques can effectively engage users to explore data and obtain insights, despite any inherent complexity[1, 3, 7]. We adopted similar methods in our product strategy and designed a narrative visualization that translates research and raw data into a user-friendly storytelling structure with functional data visualizations. However, when building Metroverse, our team encountered two fundamental problems: (1) How could we distill scientific, multidisciplinary research and data into usable insights suitable for policymakers and urban officials who may not normally engage with theoretical concepts? (2) How could we design a user experience (UX) that effectively communicated these insights to a diverse group of users?

To address these problems, we embarked on an iterative and collaborative process that included product discovery, user research, and design initiatives. The result was a narrative visualization platform that distilled Metroverse's underlying research and analysis into a coherent collection of data visualizations connected by a narrative structure that guided the user's journey.

3 PROCESS

In this section, we explain the three main steps in creating the Metroverse. First, product discovery included research and data discovery phases, involving two-way knowledge sharing between our team and the research team about data, research methodologies,

and best practices for storytelling with interactive data visualization. Second, at different stages of the project, our user research initiatives included persona building, user interviews, and user tests. Lastly, we employed wireframing and prototyping in the design process, which allowed us to create, test, and iterate ideas quickly, as well as collaborate with the research team. In this paper, we introduce the three phases of the project in a linear order. However, during the design and development of the project, all these phases were intertwined, forming a feedback loop. Figure 1 depicts the user interface of the Metroverse and below, we elaborate on each of these three steps.

3.1 Research and Data Discovery

Approaches and methods in product discovery are mostly derived from customer marketing studies, so they aim to maximize customer satisfaction and product profit [8]. However, our product is developed in an academic research community with the aim of translating research into a user-friendly tool. Therefore, we customized the product discovery methods based on our needs. Early Metroverse discovery included a broad set of cross-functional team members. Our team led discovery initiatives by applying design thinking to ideation, hypothesizing, sketching, data delivery, narrative construction, technical possibilities, and limitations. In turn, the researchers helped our team understand relevant methodologies,

data sources, and the ways in which a digital tool would complement existing and future projects. Since Metroverse research is complex and multi-layered, this collaboration helped minimize information distortion and dilution, a risk that could negatively impact robust product design. The entire Metroverse team met weekly and often in sub-groups numerous times per week. We discussed our progress and findings at roughly one-month intervals with our Principal Investigator.

3.1.1 Discovery Phase 1: Two-Way Knowledge Sharing. Knowledge sharing was our first discovery phase during which the research team supplied academic expertise while our team supplied design and technical expertise. The two teams also shared relevant materials, such as core datasets and existing visualizations. This ongoing interaction was essential to the product development process for three main reasons. First, it created a dialogue to foster understanding and alignment on core concepts, datasets, and terminology found in the scientific literature. Second, it introduced methods and techniques to frame scientific research for a hypothetical set of non-technical users. Third, it began to critically examine the reliability, scalability, and relevance of various research and data elements.

Knowledge-sharing activities mainly included online presentations, brainstorming sessions, reading, and informal conversation. For example, before our team could create a prototype of the Industry Space, one of the visualizations in the Metroverse that we explain later in this paper, we had first to understand its basic properties, such as what the Industry Space meant to depict and what informs the visual encoding of the visualization? The research team explained this information over several months of discussion and presentations, resulting in a shared understanding of the methodologies used to derive the Industry Space, its underlying data, and its many constraints.

This alignment positioned our team to explore the data and analysis Metroverse could potentially deliver. Ideas were framed as a set of simple queries categorized into several broad themes. For instance, after some initial discussion and analysis, our team determined that the dataset could accurately convey a city's composition across various dimensions, such as, what is my city's industrial composition and how many people are employed by a specific industry in my city? To draft the first set of queries, our team was motivated by what might be possible without being constrained by feasibility.

Once a long list of possible queries was drafted, our team would define a smaller subset of queries that could be developed within the project timeline and, more crucially, could be supported by the available data. These latter criteria would necessitate an intensive exploration of the existing dataset.

3.1.2 Discovery Phase 2: Data Exploration. Metroverse intends to provide a reliable representation of a city's current economic composition, its growth potential, and how it compares to other cities worldwide. Therefore, the feasibility of potential Metroverse queries had to be guided by four data dimensions.

(1) Quality - Are the data sound and reliable?
(2) Up-to-date - Are the data updated at regular intervals (i.e., monthly, quarterly, or annually)?

Figure 2: Defining a subset of queries for inclusion in Metroverse, based on which ones could be supported by various facets of the data.

(3) Scope - Are there data available for cities worldwide or just specific regions?
(4) Temporal coverage - Are the data available over a sufficiently long time period?

The data discovery phase also included an analysis of tradeoffs resulting from including different facets of data. For example, including queries that rely on firm subsidiary data would call for a substantial increase in the size of the Metroverse database. The increase in data would necessitate a compromise between application speed and reliability versus more expansive coverage. Data exploration also revealed that subsidiary data were inconsistent over time and only available for a small subset of cities. Ultimately, this information helped our team determine that subsidiary-based queries were insufficiently robust to include in the Metroverse prototype.

The data exploration phase of discovery allowed our team to interrogate each potential query against critical dimensions of data coverage and technical trade-offs. These activities helped to cull the original set of questions into a much smaller and more attainable framework.

3.1.3 Outcome: A Narrative Framework. Our discovery phases revealed that while there were many potential queries that our team wanted to include in the Metroverse prototype, only five of them were supported by sufficient data coverage and required a realistic amount of design and development time. Figure 2 depicts the process of this question selection. To transition these queries into a set of useful insights, our team decided to organize them into a narrative arc, as illustrated in Figure 1, part 1. In doing so, the intention was to ensure that, in isolation, each section could provide valuable insights to the user while also being embedded within a cohesive order that builds up the multilayered research framework from beginning to middle to end.

The research and data discovery phases were highly collaborative and iterative. It allowed our team to develop and share an understanding of the research, data, and methodologies underlying Metroverse. It would also help us transform this understanding into a set of insights that the prototype could realistically convey.
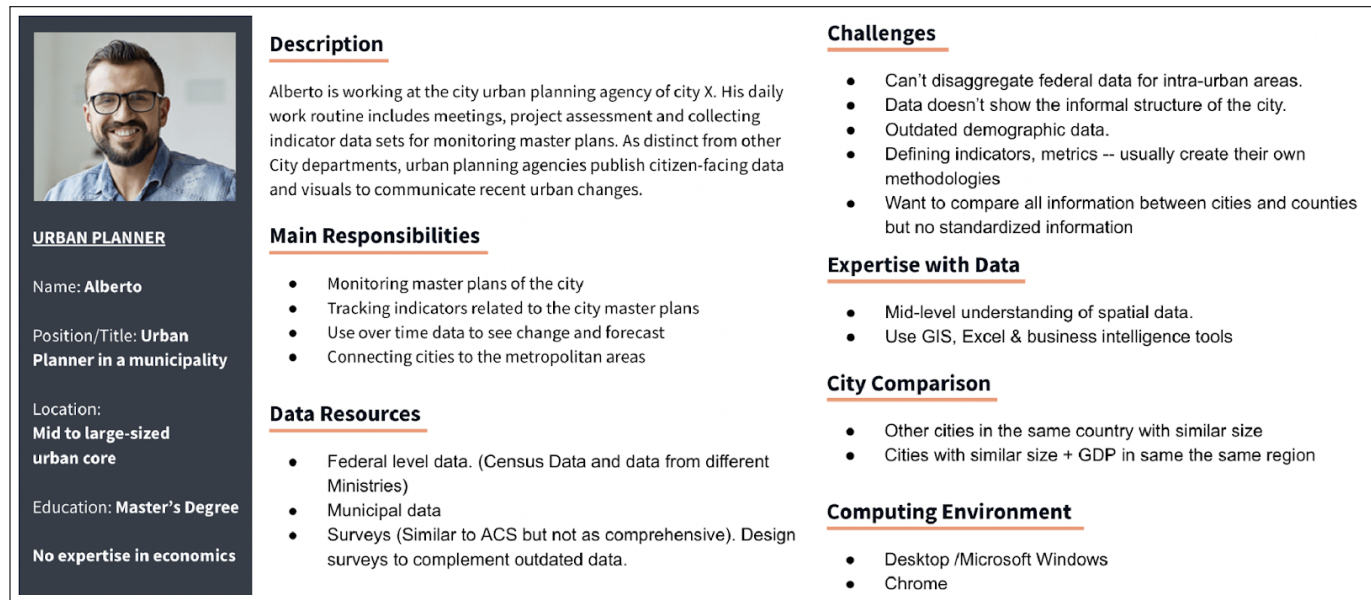
Figure 3: A sample persona profile card.

3.2 User Research

3.2.1 Forming The Pilot User Group.
The discovery phase of Metroverse produced internal understanding and alignment through knowledge sharing and data exploration. Combined, these phases resulted in forming a five-part narrative structure that would anchor the user interface and capture the underlying research methodologies. Establishing these five queries represented a vital inflection point in the process where we could begin to hypothesize, who outside the Growth Lab might use this information and how? How will we design a user experience that best communicates this information across a range of users?

To explore these questions, our team kicked off a user research phase to test and validate our assumptions about the needs and capabilities of a hypothetical set of users. To focus our user research, we recruited a group of domain experts that we believed might be typical users of the tool. The Metroverse Pilot User Group[5] comprised 15 senior-level officials from city offices worldwide. Since we designed Metroverse to be a global tool, we designed our Pilot User Group so that it was drawn from various income levels, populations, and industrial compositions. Roughly half the group worked in their mayor's office, while others held senior positions in affiliated urban groups.

3.2.2 Building Personas.
To validate our design and development decisions and better understand Metroverse users' goals, our team embarked on a process to create user personas. In the case of Metroverse, personas served two purposes: First, they allowed us to gain deep insights into our users, their behaviors, goals, daily work, and data literacy level. Second, they provided a common reference point for our team when planning and developing features for specific users.

Through surveys and structured interviews with the Pilot User Group members, we created five personas: urban planner, data manager, advanced researcher, professor, and economic developer. Each persona was detailed on profile cards, as shown in Figure 3. To determine the sections of the persona profile cards, we followed a backward design process. We first determined the user behavior data we would need to design Metroverse features. Then, we put together the interview questions that would provide the desired data. For example, in order to determine how much explanation, we should include in data visualizations, we investigated each user's data literacy level by asking them what data tools they use, how often they work with information graphics, and if they have other people in their team to take on data-related work.

User research and the resulting personas helped define the representative Metroverse users and their significant differences in motivation, capability, and behavior. These lessons helped guide our team's feedback sessions, design iterations, and design of a user experience and user interface that effectively communicate research insights.

3.2.3 User Testing.
With a set of user personas developed, our team could advance different designs and features into testing with the Pilot User Group members. Doing so allowed our team to learn whether proposed elements were relevant or usable and how they differed across personas.

Our user tests combined two methodologies: think-aloud and task-based. Two rounds of testing were conducted with each Pilot User Group member. Because the study took place during the Covid-19 pandemic, all tests were performed remotely: one via Zoom[6] and one on a remote testing platform called Lookback[7]. The remote test via the Lookback platform closely simulated the Zoom tests. The
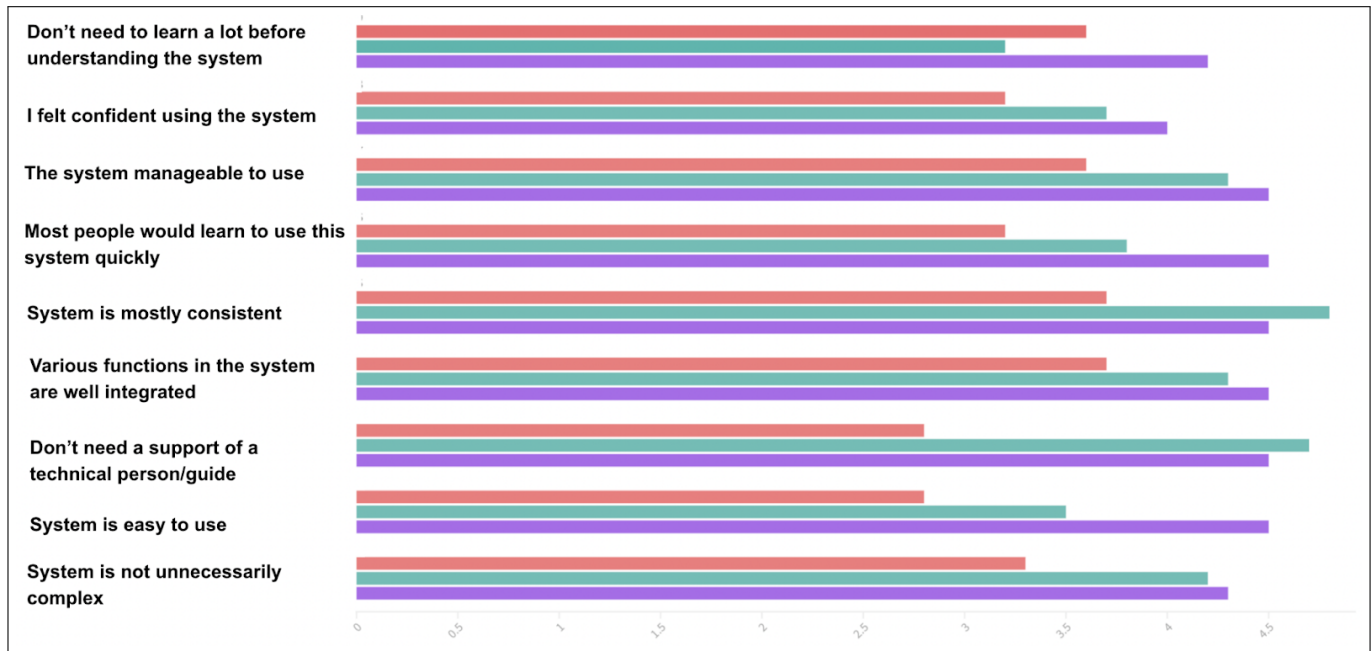
Figure 4: System Usability Scale results for the Metroverse and the other two tools by the Growth Lab.

platform showed users the tasks individually and recorded their screen, video, and audio. We then watched these recordings and took notes.

Besides collecting qualitative information from the users through observations and open-ended interviews, we also conducted quantitative evaluations to track our progress and see whether the usability of our tools was improving. To quantify a user's experience, many tools and methods are available such as A/B testing, task completion rates, and surveys [6]. Working as a small team and with limited time for user research, we decided to use the System Usability Scale (SUS) survey [2], asking all participants to complete it immediately after user test sessions. SUS provides a 10-item questionnaire about users' experience using the tool. Each of the 10 has five response options for respondents, ranging from strongly agree to strongly disagree. We calculated an overall score using SUS results and then used it as a benchmark to measure the relative success of our design. We found it to be a reliable tool for measuring usability.

In the case of Metroverse, the overall usability score was 4.4 out of 5.0. Figure 4 shows the tool's performance based on each usability metric. We then compared this score with the scores of other tools by our team. We discovered that Metroverse has been the most successful tool among the other tools we designed in terms of usability scores as seen in Figure 4. This was the first time our team had employed persona study and Pilot User Group interviews in the design process, which we believe was responsible for this high usability score.

3.3 Design Process

In our design process, prototyping and wireframing enabled us to explore alternative data visualization and user experience techniques for exposing the underlying research. The data visualizations were interwoven in a coherent story that guided users. Hence, storytelling was equally important in the Metroverse design process.

While wireframing was the primary tool used in creating the storytelling component and the overall user experience, prototyping was used to create data visualization design alternatives. Below, we present two case studies of each design track, wireframing and prototyping, to illustrate our overall methodology and approach to the design.

3.3.1 Wireframing Case Study: User Flow of the Metroverse. The Metroverse comprises five interactive data visualization pages that were brought together through a narrative design. Through static wireframes, we worked to find the most suitable narrative visualization technique to tell the story we wanted to convey.

In their paper, Narrative Visualization: Telling Stories with Data, Segel and Heer provided a taxonomy of narrative visualization. They analyzed 58 related projects based on principles such as interaction strategies, navigation tactics, and visual devices that support storytelling with data [7]. Their outcome highlighted the importance of having a balance between the author-driven narrative visualization, a linear path with no interactivity, and the reader-driven narrative visualization, a non-linear path with high-degree interactivity that gives the user control to customize their own story and, eventually, findings.

Based on this study and our previous experience in designing similar tools, we decided to visually show the research questions that Metroverse is built upon. The main navigation system is built
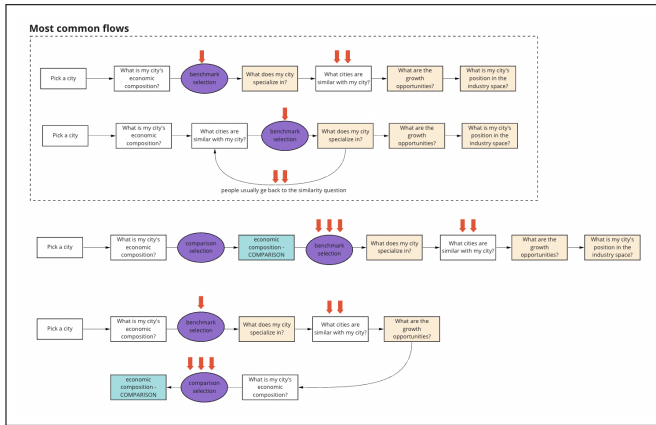
Figure 5: User flows from the user tests. Purple ovals represent the points when the user is prompted to make a city benchmark selection. Yellow and blue rectangles represent the visualizations dependent on the benchmark selection. Red arrows show the pain points where users are confused about the navigation.

around the five-part narrative structure on the arch-shaped navigation panel, as illustrated in Figure 1, part 1. This design decision provided a non-linear path in which users are free to click on any question and explore the related visualization. The data visualization and right-side narrative text (Figure 1, part 2, and part 3) on each page are almost non-interactive except for the tooltips. However, we designed an extensive Visualization Options panel (Figure 1, part 4) to allow users to customize the visualizations and discover different information. By doing so, we tried to have features from author- and reader-driven narrative visualization techniques.

During the research and data discovery phase, the research team emphasized the importance of introducing a comparison feature in which users can compare their selected city against another city or a group of cities. This comparison feature created a series of usability challenges in the user journey because users were periodically interrupted to select another city to use the comparison feature. In order to understand these usability challenges and detect exactly where users were interrupted, we conducted a series of user tests and visualized every tester's user journey through wireframes. We identified the points where users didn't understand the interface prompts or were confused about where to go and what to click on. These pain points are shown as red arrows in Figure 5. After discovering the user flow flaws, we went back to the wireframing phase and redesigned the new user experience and navigation features. As a result, we removed all interface prompts that were asking users to select peer cities. Instead, we designed a new visual language with colorful buttons and icons that indicate the comparison feature to the users. This can be seen in Figure 1, part 5.

3.3.2 Prototyping Case Study: The Industry Space. The industry space is one of the visualizations in the Metroverse. It is a formal representation in which related industries, represented as nodes in the visualization, are closer to each other. Industries are related when they require similar capabilities. For instance, in making cars,

a city requires similar capabilities to make trucks. These capabilities are, however, different from the ones that are needed for banking or insurance services. We used prototyping as a tool for Industry Space design discovery and stakeholder collaboration. We translated complex research insights as defined in the Industry Space research paper [5] into accessible, comprehensible, and explorable visual representations of the Industry Space in the Metroverse. As seen in Figure 6, through an iterative design process, we refined and simplified the Industry Space visualizations that can be understood and interpreted by non-expert users. We started to design Industry Space prototypes after the discovery phase of collecting and studying research questions, insights, and data from the research team. Building rapid, interactive, high-fidelity prototypes enabled an iterative process through feedback sessions with multiple stakeholders. Rapid prototyping allowed continuous feedback and improved the quality of input through constant exposure to new design ideas. Our prototypes were highly interactive because they presented multilayered information—the nature of research questions and findings that we worked with necessitated such complex information layering. In total, we worked on 25 Industry Space prototypes through 500 designer and developer hours, spanning 12 months. We presented each of these prototypes to the research team for their evaluation of whether the visualization communicates the research methodology and to the non-expert users for their evaluation of whether the visualization is understandable or not.

4 DISCUSSION AND LESSONS

Metroverse was launched on June 2, 2021. Since then, it has been used by researchers, policymakers, and government officials worldwide. We continue engaging with end-users through presentations, demo sessions, and interviews. This enables us to collect feedback and use cases to evaluate our processes and learn for future projects. Based on our evaluations thus far, we identified two main aspects that worked well during the project. First, our interviews showed that the Pilot User Group and persona studies significantly improved the user experience of the tool by helping to align hypothetical profiles with the actual end-user profiles. For example, a group of policy researchers has applied Metroverse to their project in South Africa, where they worked to improve the prosperity of cities. They stated that having multiple options to customize and interact with data visualizations led to new research questions and insights into their research. In another example, we heard from a professor who used the Metroverse in her class and stated that the tool helped her to teach students how to pull different data together to reach a conclusion about cities. Both examples served as early evidence that having a Pilot User Group and personas helps design a user experience that ultimately appeals to a diverse user group.

Second, we found that the iterative design process, through wireframing and prototyping, has been effective in distilling scientific, multidisciplinary research into user experience design. For example, our data visualization prototypes helped to facilitate conversations with the research team, acting as a "translator" between two different teams and helping us to find common ground for discussion and design decision-making.

Our outreach to end-users also highlighted some aspects of the project that did not work well. Particularly, we took two important
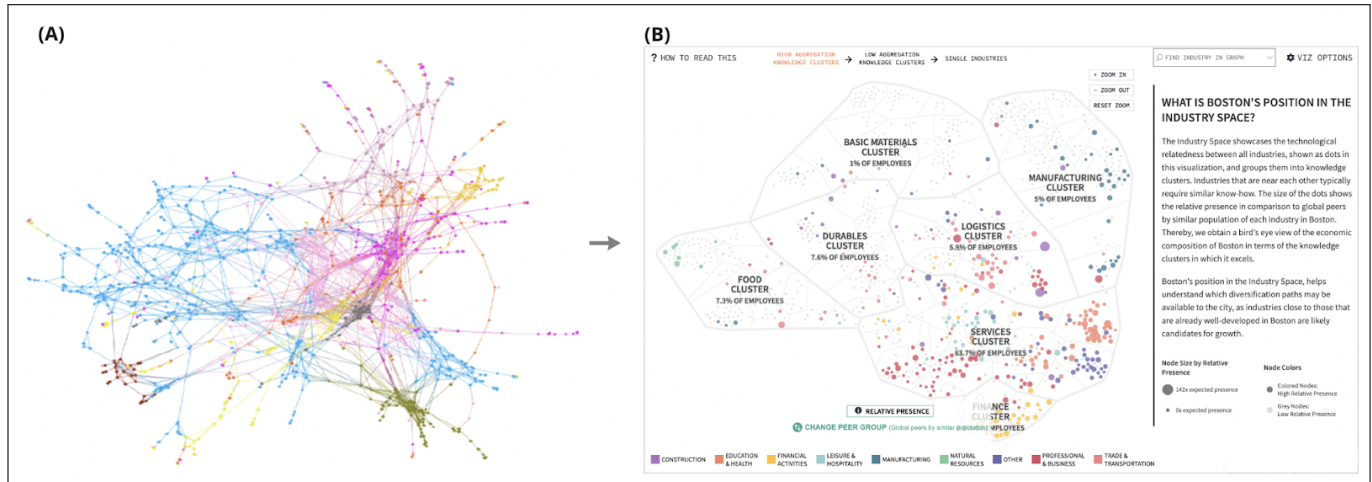
Figure 6: (A) Industry Space in a research paper. (B) Industry Space in the Metroverse.

lessons for future projects. First, we found that having learning materials for a new tool is as important as any feature of the tool. Our interviews revealed that current Metroverse users need detailed introductions on the research concepts and specific visualizations. Second, our observations and learnings from the users suggested that the tool needs an onboarding page that summarizes what the tool offers. For example, a couple of users expressed their confusion that they did not know what to do when they opened the tool, and needed basic information about a city, such as its population and GDP, before exploring its economy in greater detail. In sum, we learned that we need to develop a holistic approach to onboarding and teaching the users about our tool, and this needs to be simultaneous with the early design decisions.

5   CONCLUSION AND FUTURE WORK

In this case study, we described the experience of designing and developing the Metroverse prototype. By adapting various user-centered design methods such as product discovery, persona building, and prototyping to our needs, we distilled complex research and data into a user-friendly narrative visualization platform. We hope that the presented processes and lessons can help the designers and data visualization practitioners who work on similar problems.

As mentioned in the discussion section, one of the main takeaways of our end-user outreach was the lack of learning materials that accompany the tool. Currently, our team is working on prototyping various learning materials for a future iteration of the Metroverse. More user research studies are needed to determine what types of learning and onboarding materials would be the most useful for our diverse user group.

ACKNOWLEDGMENTS

REFERENCES

[1] Jeremy Boy, Francoise Detienne, and Jean-Daniel Fekete. 2015. Storytelling in information visualizations: Does it engage users to explore data?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. 1449–1458.
[2] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
[3] Nahum Gershon and Ward Page. 2001. What storytelling can do for information visualization. *Commun. ACM* 44, 8 (2001), 31–37.
[4] Ricardo Hausmann, César A Hidalgo, Sebastián Bustos, Michele Coscia, and Alexander Simoes. 2014. *The atlas of economic complexity: Mapping paths to prosperity.* Mit Press.
[5] Frank Neffke, Matté Hartog, Ron Boschma, and Martin Henning. 2018. Agents of structural change: The role of firms and entrepreneurs in regional diversification. *Economic Geography* 94, 1 (2018), 23–48.
[6] Jeff Sauro and James R Lewis. 2016. *Quantifying the user experience: Practical statistics for user research.* Morgan Kaufmann.
[7] Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148.
[8] Teresa Torres. 2021. *Continuous Discovery Habits: Discover Products that Create Customer Value and Business Value.* Product Talk LLC.