

Fostering Data Humanism with DataPortraits: Empowering People to Create a Personalized Visual Vocabulary

Nam Wook Kim*
Harvard University

Hyejin Im†
Tufts University

Nathalie Henry Riche‡
Microsoft Research

Krzysztof Gajos§
Harvard University

Hanspeter Pfister¶
Harvard University

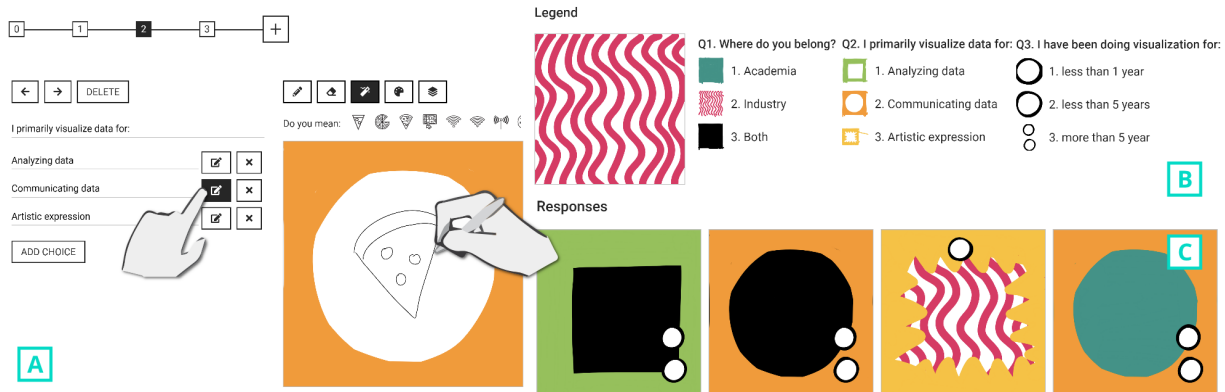


Figure 1: DataPortraits: A: the editor interface for creating a questionnaire and visual vocabulary. A user can draw a unique personalized visual for each option. B: An interactive legend aids the interpretation of visual encodings as necessary. C: Each form response generates a distinctive visualization based on the visual vocabulary.

ABSTRACT

Giorgia Lupi’s manifesto on data humanism [8] is advocating for leveraging data visualizations as catalysts to reflect upon ourselves and engage with others via data that matters in our lives. In this work, we present DataPortraits, an interface that enables people to design a personalized visual vocabulary to represent data. DataPortraits can support both individual and collaborative scenarios, empowering people to gather and represent personal data for self-reflection (e.g., create custom hand-drawn visualization of emotions and daily activities over time) or to generate visual artifacts that are shared in a group to foster engagement and discussion.

Index Terms: Human-centered computing—Visualization system and tools

1 INTRODUCTION

Data can capture a snapshot of the world and allow us to better understand ourselves and our communities. Giorgia Lupi, a renowned visualization artist, recently advocated for *data humanism*, a more personal approach to collecting, analyzing and visualizing data [8]. In contrast to traditional data systems, where visualization tools focus on an efficient analysis and presentation of large quantities of data, data humanism puts more emphasis on slowness, engagement with the data, as well as unique visuals with personality that are embracing context and subjectivity of the data.

Lupi’s manifesto echoes in a lot of research in the areas of personal informatics, self-tracking, and casual visualization [5], including an abundance of tools for collecting and visualizing personal

data [4]. Most of these tools, however, focus on data collection, specifically automated tracking, and use predefined presentations of the data. Users remain mostly passive and less engaged with the data as such. On the other hand, there has been a surge of construction tools allowing people to create custom visualizations of data [6, 11]. But these tools focus on authoring processes and tool expressiveness and have not been applied to a personal data context.

In this work, we are investigating the question of how to empower people with the ability to design their own visual vocabulary to represent qualitative and nuanced aspects of personal data. We developed DataPortraits, a work-in-progress prototype enabling users to create a questionnaire to ask questions about themselves and to design a personalized visual vocabulary to represent their collected data. DataPortraits is offering users full autonomy over the visual presentation of the data to create visualizations that are sensemaking tools in a casual context for users without advanced statistical literacy. DataPortraits makes it possible to design the visual vocabulary at the time of creating the questionnaire. In this way, the construction of visualizations is not an afterthought, but a primary activity that users actively engage in while thinking about the goals of their data collection. In this paper, we overview the current prototype and discuss future directions for improvements and user evaluation studies.

2 DATAPORTRAITS

The design of DataPortraits was mainly inspired by Giorgia Lupi’s recent work to generate visual data portraits of conference attendees to foster conversations and connections among them [9]. She created a questionnaire asking fun facts about attendees and developed a visual vocabulary that turns each answer into a set of shapes, colors, and symbols. After filling out the questionnaire, attendees received personalized badges with unique hand-drawn visualizations reflecting their views and personality based on the answers to the questionnaire. DataPortraits generalizes this innovative idea to empower users to create a visual vocabulary to represent data about ourselves, our peers, or our communities.

*e-mail: namwkim@seas.harvard.edu

†e-mail: hyejin.im@tufts.edu

‡e-mail: Nathalie.Henry@microsoft.com

§e-mail: namwkim@eecs.harvard.edu

¶e-mail: pfister@seas.harvard.edu

2.1 User Interface

The prototype interface consists of three main components similar to standard survey tools such as TypeForm or Google Forms: a questionnaire editor (Fig. 1A), the survey form (Fig. 1B), and a response viewer (Fig. 1C). We leverage this questionnaire metaphor as it is a versatile method for data collection and familiar to everyone.

Unlike the existing survey tools, the questionnaire editor comes with a canvas interface for constructing a visual mapping for each response in the multiple choice question. The canvas affords free-form sketching that is a flexible yet powerful design tool. To assist users with less experience in drawing, we integrated an auto-drawing feature that suggests a predefined icon based on a series of strokes (Fig. 1A [1]).

We take a layered approach to model the underlying visual vocabulary. In other words, a user can draw any visual at each layer in the hierarchy from a form to a question and each answer option. That is, drawings at the form-level serve as a background while drawings at the question-level shared by all answer options but specific to a single question.

Since all questions in a survey share the same canvas area, it is necessary to see if the visual mapping of the current question conflicts with those of other questions. To handle this, a user can toggle the visibility of visual mappings from other layers. In this way, a user can gauge how the visual encodings of the questions would interact in the final outcome.

The design practice of creating a visual vocabulary is a pro-active rather than a retro-active process. As a result, the authoring process is top-down, starting with visuals that are then mapped to data, while a typical visualization tool follows a bottom-up process, where data is mapped to visuals. The final visualization is not available until the user fills out the survey form. The visual vocabulary can be reused for more than one form response and revised iteratively without breaking the previous visualizations. This allows users to freely add additional questions and revise visual mappings as time goes on.

The response viewer shows all the form responses in the form of a visual data portrait. The responses are arranged in small multiples, allowing them to be compared and reveal potential patterns between them. A user can print and share each visual data portrait. To assist in interpreting the data portraits, a visual legend is automatically generated from the visual vocabulary.

2.2 Use Case Scenarios

We imagine DataPortraits to be used in both individual and collaborative scenarios. Individual users can use it for self-tracking purposes. Examples including tracking food consumption (e.g., water, veggies, junk food), bedroom routines (e.g., waking up early, going to bed late), resolutions (e.g., no smoking, no phone at dinner), and mood (e.g. happy, lonely). While there is a high data capture burden in this manual survey approach, in comparison to automated tracking, it provides more flexibility in tracking data and engages participants in data collection, thus enhancing self-awareness [7].

In this self-tracking context, a recent study shows that there is strong evidence that people prefer to use visuals that are personally meaningful to them to eventually build a positive emotional relationship to the data being presented [3]. For instance, bullet journalists [2] often use colors to indicate different moods, as well as object shapes that they simply like (e.g., animals) have some relevance to their tracking goals (e.g., vehicles and food).

Using data portraits for conference badges is an innovative collaborative scenario that has been demonstrated by Giorgia Lupi as mentioned earlier. In this case, each visual portrait represents each different individual. This scenario can be applied to different contexts and communities other than conferences (e.g., friends and families). By putting all the portraits into a collage can represent a community that can be displayed in a public space and available for a collective reflection [10].

3 DISCUSSION & FUTURE WORK

DataPortraits contributes to the field of personal visualization and personal informatics, in which a lot of research efforts have been being made to support enhanced personalization and customization. It follows the principle of data humanism by giving the power to people to add personal and contextual meaning to data using a visual vocabulary.

There is a lot of room for improving the current prototype. In particular, assistive features for data sketching would be beneficial for novices who may have fear of drawing on an empty canvas. Examples include providing automatic encoding of certain visual variables such as color and position, drawing templates (e.g., drawing bases like those in a coloring book), or inspirational artworks as done in the *Dear Data* studio in VIS'17.

We also plan to add additional questions to collect different types of data (e.g., numeric, Likert scale, checkboxes) and to provide questionnaire templates in different domains (e.g., health, productivity). These templates could come with predefined visual mappings that can be later customized. To reduce the burden in capturing data, integrating automated tracking capability would be a viable solution [7].

DataPortraits raises questions about the potential benefits for people to have complete control over how to represent their data. Does this slow and reflective practice provide a more mindful experience of making sense of personal data? More specifically, does it promote long-term engagement in a self-tracking scenario and are people more likely to comply with their goals? Does it encourage more conversations among people, if so, in what ways? We plan to address these questions through user evaluation studies.

ACKNOWLEDGMENTS

This work has been made possible through support from the Kwanjeong Educational Foundation.

REFERENCES

- [1] Autodraw. <https://www.autodraw.com/>. Accessed: 2018-06-16.
- [2] Bullet journal. <http://bulletjournal.com/>. Accessed: 2018-06-16.
- [3] A. Ayobi, T. Sonne, P. Marshall, and A. L. Cox. Flexible and mindful self-tracking: Design implications from paper bullet journals. In *CHI'18*, p. 28. ACM, 2018.
- [4] S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. Harrison, P. Klasnja, A. LaMarca, L. LeGrand, R. Libby, et al. Activity sensing in the wild: a field trial of ubifit garden. In *CHI'08*, pp. 1797–1806. ACM, 2008.
- [5] D. Huang, M. Tory, B. A. Aseniero, L. Bartram, S. Bateman, S. Carpendale, A. Tang, and R. Woodbury. Personal visualization and personal visual analytics. *IEEE TVCG*, 21(3):420–433, 2015.
- [6] N. W. Kim, E. Schweickart, Z. Liu, M. Dontcheva, W. Li, J. Popovic, and H. Pfister. Data-driven guides: Supporting expressive design for information graphics. *IEEE TVCG*, 23(1):491–500, 2017.
- [7] Y.-H. Kim, J. H. Jeon, B. Lee, E. K. Choe, and J. Seo. Omnitrack: A flexible self-tracking approach leveraging semi-automated tracking. *ACM IMWUT, Wearable and Ubiquitous Technologies*, 1(3):67, 2017.
- [8] G. Lupi. Data humanism, the revolution will be visualized. <http://www.printmag.com/information-design/data-humanism-future-of-data-visualization/>, Feb. 2017.
- [9] G. Lupi. How can we use personal data to create a connection between people? <http://giorgialupi.com/data-portraits-at-ted2017/>, 2017.
- [10] G. Lupi. A data wallpaper for story: soft and small data to capture people's vision for the future of work. <http://giorgialupi.com/work/#/collaborative-data-wallpaper-for-story/>, 2018.
- [11] H. Xia, N. Henry Riche, F. Chevalier, B. De Araujo, and D. Wigdor. Dataink: Direct and creative data-oriented drawing. In *CHI'18*, p. 223. ACM, 2018.