
Digital Collaborator: Augmenting Task Abstraction in Visualization Design with Artificial Intelligence

Aditeya Pandey*

Northeastern University
Boston, MA 02115, USA
pandey.ad@husky.neu.edu

Andrea G. Parker

Georgia Institute of Technology
Atlanta, GA 30308, USA
andrea@cc.gatech.edu

Yixuan Zhang*

Northeastern University
Boston, MA 02115, USA
zhang.yixua@northeastern.edu

Michelle A. Borkin

Northeastern University
Boston, MA 02115, USA
m.borkin@northeastern.edu

John A. Guerra-Gomez

Northeastern University
San Jose, CA 95138, USA
john.guerra@gmail.com

* Authors contributed equally

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Copyright held by the owner/author(s).
CHI'20., April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
<https://doi.org/10.1145/3334480.XXXXXXX>

Abstract

In the task abstraction phase of the visualization design process, including in “design studies”, a practitioner maps the observed domain goals to generalizable abstract tasks using visualization theory in order to better understand and address the user’s needs. We argue that this manual task abstraction process is prone to errors due to designer biases and a lack of domain background and knowledge. Under these circumstances, a collaborator can help validate and provide sanity checks to visualization practitioners during this important task abstraction stage. However, having a human collaborator is not always feasible and may be subject to the same biases and pitfalls. In this paper, we first describe the challenges associated with task abstraction. We then propose a conceptual *Digital Collaborator*—an artificial intelligence system that aims to help visualization practitioners by augmenting their ability to validate and reason about the output of task abstraction. We also discuss several practical design challenges of designing and implementing such systems.

Author Keywords

Information visualization; artificial intelligence; natural language processing; task abstraction.

CCS Concepts

•Human-centered computing → Visualization; •Computing methodologies → Artificial intelligence; NLP;

Task-Abstraction Frameworks

To highlight the variation in task-abstraction methodologies, we present some distinguishing characteristics of three common frameworks:

A Multi-Level Typology of Abstract Visualization Tasks [1]: A generic task abstraction framework that works well across disciplines and data-set types.

Task Taxonomy for Graph Visualization [4]: A descriptive framework for tasks in the field specific to graph visualizations. This taxonomy provides more descriptive identification of visualization goals than a generalized framework [8].

Hierarchical Task Abstraction (HTA) [11]: HTA highlights the importance of integrating context and leverages existing task abstraction frameworks in combination with a systematic analysis of user tasks, goals, and processes.

Introduction

Artificial intelligence (AI) has been used in the data information community to help improve design of visualizations [5, 10]. A visualization practitioner can get help from a variety of tools (e.g., Tableau, QlikView, SAS) [3] to select proper visual encodings. However this step must be carefully considered in the context of user goals and tasks. The visualization design process can be broadly divided into phases [6], including the design study process model [9], performing task analysis to understand domain problems, and task abstraction that aims to recast user goals from domain-specific languages to a generalized terminology for better understanding and readability [6]. Conducting task abstraction is an important but rigorous manual process that requires in-depth understanding of domain knowledge and familiarity with visualization literature [1, 4, 11]. For example, a biologist may be interested in results for tissue samples treated with LL-37 matching up with the ones without the peptide. A visualization researcher may translate this task to **compare** values between **two groups** [6]. However, to accurately perform task abstraction, a visualization practitioner must first choose the best abstraction framework and then the appropriate abstraction. A practitioner has to keep up with the ever-growing task abstraction literature [1, 4, 11] and ensure that their personal biases that might come from previous work experiences do not affect their ability to perform task abstraction.

As task abstraction is a manual and subjective phase of the visualization design process, we argue that it may be prone to human-judgment errors. For example, domain experts often serve as project collaborators to help visualization researchers and practitioners validate the task analysis and abstraction in human-centered studies. However, it is challenging to have collaborators' involvement in many situations. Furthermore, human collaborators are still prone to pitfalls like keeping pace with recent development of task-abstraction theories and priming biases. Therefore, we propose an AI-enabled *Digital Col-*

laborator (DC) that can serve as a feasible alternative to a human collaborator. We envision that a DC can assist visualization researchers by being up-to-date with task-abstraction frameworks to help identify the most appropriate framework for abstraction and can help validate the task analysis and abstraction process to identify judgment errors or biases. With this paper we hope to open a discussion on the advantages and challenges of building a DC for the visualization community.

Challenges of Performing Task Abstraction

We first discuss the main challenges associated with the task abstraction process.

A Wide Range of Task Abstraction Approaches: Visualization researchers have proposed various task abstraction approaches (e.g., [1, 4, 11]). On page 2 (side-bar), we discuss three common visualization task abstraction frameworks and explain how they differ. Adopting an appropriate task abstraction approach is pivotal for visualization design as it impacts the choice of visualization design and interaction idioms. However, selecting a proper task abstraction framework requires an extensive comparison of existing literature.

Interpretation of a Task Abstraction Framework: Task abstraction is a subjective evaluation of the domain experts' needs. Subjective assessments are prone to errors arising from variability in the practitioner's understanding of an abstraction framework or an innate bias such as recency bias where the task-abstraction may be influenced by recent work. Such abstraction biases can lead to a "domino" effect of errors that can only be objectively verified after prototyping [9]. Additionally, the analytic-task focused taxonomies require mastery of the terminology and definitions [4]. For example, in network abstraction, it is common to use the term *Topology* for properties related to the structure of the network. Topology is a

A Task Abstraction Process Example: Diabetes Management

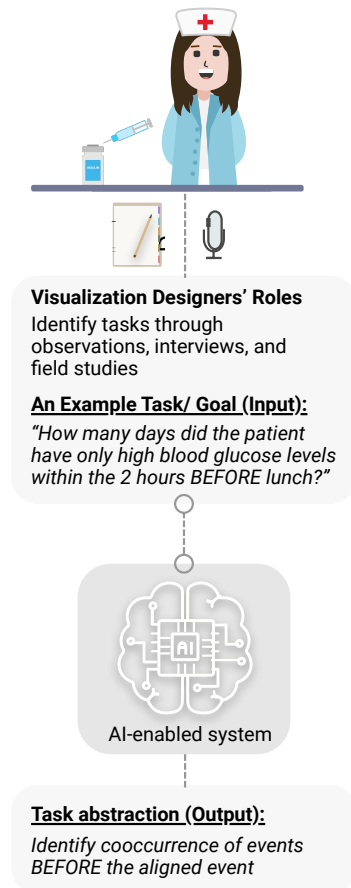


Figure 1: Example process to use AI to facilitate task abstractions in visualization research.

mathematical term, and practitioners coming from design backgrounds may be unfamiliar with its meaning.

Automate Task Abstraction using AI

We have discussed some challenges of performing task abstraction with human effort involved. Drawing inspiration from the idiom of an Intelligent Personal Assistant (IPA) [2], we propose a Digital Collaborator (DC)—a conceptual AI-enabled system to support task abstraction for visualization research. Figure 1 shows an example of how AI can be used to facilitate task abstraction.

Input and Output: Similar to the IPA system, our proposed DC will adopt a question-and-answer-based interface. The questions (*input*) will be domain goals identified by visualization practitioners through interviews and observations with domain experts. The DC should generate a translation of a domain goal to a generalized task description by applying an appropriate task-abstraction framework (*output*). To improve communication transparency, the DC should aim to provide the rationale for their output and a set of alternative translations.

AI System: We believe system goals should include identifying the right abstraction framework and recommending the appropriate analytical conversion of the tasks. Note that we do not intend to suggest replacing the human-centered approaches when conducting task analysis (e.g., field studies, interviews, and observations). Instead, we propose to leverage AI in designing systems to help ease the process of task abstraction.

Challenges of Our Proposed AI-enabled System

In order to develop an AI to help automate task abstraction we acknowledge that there will be challenges to design such an AI-enabled DC system.

Framework Characterization: Task-Abstraction frameworks are well established. However, there is little guidance on how

to select the “right” framework. Therefore, the first challenge of building a DC will be to develop parameters to distinguish between these abstraction frameworks.

Training Data: For automating the task-abstraction process, we need to train machine learning models with task data and their labeled outputs. One way to acquire training data is by parsing domain goals and their abstractions from existing literature. Smart data crawling tools may facilitate the process of extracting tasks from research papers with little manual effort. However, even after deploying web-crawlers, there might be problems with data quality. For instance, there might be conflicting abstractions where similar tasks have different abstractions. To counter the problem, we can think of human-in-the-loop methodology where visualization researchers working on the project can address quality issues.

Recommendation Validation: Task abstraction involves subjective evaluation and characterization of domain problems. Practitioners may disagree with the suggested results generated by the DC. Therefore, an open question is how to instill confidence in visualization practitioners to consider the suggested results before discarding them. For example, extending the recommendation list with confidence scores may increase transparency. Future research should examine design recommendations that can boost confidence in communicating results.

Equity Issues with AI: There has been an increasing body of research on equity issues in AI research, such as biased datasets and algorithm transparency [7]. An open research question is *how can an AI-based DC system promote equity?* There is a vital need for future research to examine how such systems can be made more accessible for a wide audience. For example, how a variety of voices and experiences can be captured using such systems? How nuanced aspects of these experiences contribute to different design requirements

and task abstractions, which will ultimately influence the design choices? Therefore, more research is needed to further explore these issues.

Conclusion

In this paper we propose a conceptual AI-enabled digital collaborator to assist in performing visualization task abstraction. We discuss the advantages as well as the challenges of designing such AI-enabled systems, including training data, designing for transparent communication, as well as equity issues with AI. Through this workshop paper, we want to initiate a discussion on the topic of how AI can assist task abstraction in visualization research and how to address these challenges.

Acknowledgements

This work was supported by NSF CISE CRII award no. 1657466.

REFERENCES

- [1] M. Brehmer and T. Munzner. 2013. A Multi-Level Typology of Abstract Visualization Tasks. *IEEE TVCG* 19, 12 (Dec 2013), 2376–2385. DOI: <http://dx.doi.org/10.1109/TVCG.2013.124>
- [2] Gabriela Czibula, Adriana-Mihaela Guran, Istvan Gergely Czibula, and Grigoreta Sofia Cojocar. 2009. IPA-An intelligent personal assistant agent for task performance support. In *2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing*. IEEE, 31–34.
- [3] Michael Diamond and Angela Mattia. 2017. Data visualization: An exploratory study into the software tools used by businesses. *Journal of Instructional Pedagogies* 18 (2017).
- [4] Bongshin Lee, Catherine Plaisant, Cynthia Sims Parr, Jean-Daniel Fekete, and Nathalie Henry. 2006. Task Taxonomy for Graph Visualization (*BELIV '06*). Association for Computing Machinery, New York, NY, USA, 1–5. DOI: <http://dx.doi.org/10.1145/1168149.1168168>
- [5] MLUI2019. 2020. MLUI 2019: Machine Learning from User Interactions for Visualization and Analytics. Website. (02 February 2020). <https://learningfromusersworkshop.github.io/>.
- [6] T. Munzner. 2009. A Nested Model for Visualization Design and Validation. *IEEE TVCG* 15, 6 (Nov 2009), 921–928. DOI: <http://dx.doi.org/10.1109/TVCG.2009.111>
- [7] Osonde A Osoba and William Welser IV. 2017. *An intelligence in our image: The risks of bias and errors in artificial intelligence*. Rand Corporation.
- [8] Aditeya Pandey, Harsh Shukla, Geoffrey S Young, Lei Qin, Amir A Zamani, Liangge Hsu, Raymond Huang, Cody Dunne, and Michelle A Borkin. 2020. CerebroVis: Designing an Abstract yet Spatially Contextualized Cerebral Artery Network Visualization. *IEEE TVCG* 26, 1 (Jan 2020), 938–948. DOI: <http://dx.doi.org/10.1109/TVCG.2019.2934402>
- [9] M. Sedlmair, M. Meyer, and T. Munzner. 2012. Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE TVCG* 18, 12 (Dec 2012), 2431–2440. DOI: <http://dx.doi.org/10.1109/TVCG.2012.213>
- [10] Manasi Vartak, Silu Huang, Tarique Siddiqui, Samuel Madden, and Aditya Parameswaran. 2017. Towards visualization recommendation systems. *ACM SIGMOD Record* 45, 4 (2017), 34–39.
- [11] Yixuan Zhang, Kartik Chanana, and Cody Dunne. 2019. IDMVis: Temporal Event Sequence Visualization for Type 1 Diabetes Treatment Decision Support. *IEEE TVCG* 25, 1 (Jan 2019), 512–522. DOI: <http://dx.doi.org/10.1109/TVCG.2018.2865076>