VisVisual: A Toolkit for Teaching and Learning Data Visualization

Chaoli Wang
University of Notre Dame

***Abstract*—This article describes the motivation, design, and evaluation of the VisVisual toolkit to engage students in learning essential visualization concepts, algorithms, and techniques. The toolkit includes four independent components: VolumeVisual, FlowVisual, GraphVisual, and TreeVisual, covering scalar and vector data visualization in scientific visualization and graph and tree layouts in information visualization. Complementary to the toolkit design is resource development, aiming to help instructors integrate VisVisual into their curriculum.**

ALTHOUGH the field of visualization, including scientific visualization (SciVis) and information visualization (InfoVis), has been developed for three decades, we still lack pedagogical tools for visualization education. Researchers have long investigated visualization education for years. For example, an early work [2] presented a software repository for education and research in InfoVis, but it was more a software library instead of a pedagogical tool. Many Eyes [11] provided a public website for users to upload data, create interactive visualizations, and initiate discussions. However, the website was more concerned with practicing and sharing visualization by the general public instead of grasping the underlying visualization algorithms. With built-in support for InfoVis techniques, the InfoVis Toolkit [4] and D3 (Data-Driven Documents) [3] are software framework and domain-specific language that help researchers and developers write their visualization applications. More recently, educational psychology concepts were applied to design four types of online guides for InfoVis [10]. These works focus on InfoVis but not SciVis. Unlike InfoVis, which mostly deals with abstract data and visual representation in various 2D forms (e.g., scatterplots, parallel coordinates, treemaps, graphs, and diagrams), SciVis deals with spatial data and 3D rendering (e.g., volume visualization, flow visualization), and therefore, can potentially benefit more from using an interactive pedagogical tool. However, little work has been done in this regard.

In a pilot experiment, my research group developed FlowVisual for learning flow visualization concepts. The tool was used in the Data Visualization classes at Michigan Technological University (CS 5090) and the University of Notre

 Published by the IEEE Computer Society

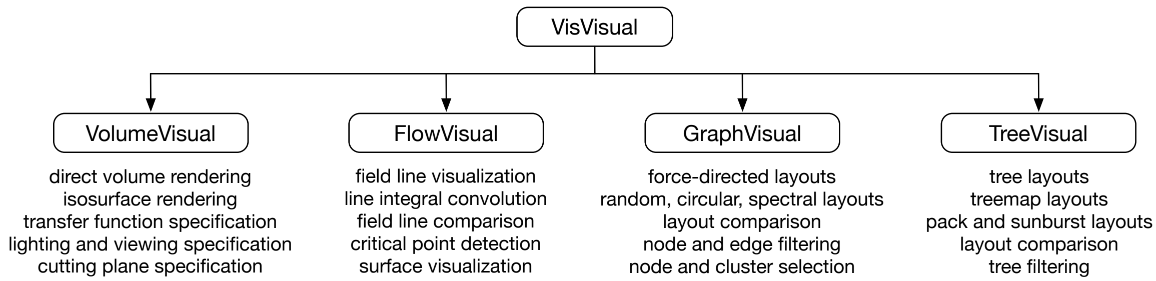

Figure 1. The VisVisual toolkit consists of VolumeVisual, FlowVisual, GraphVisual, and TreeVisual. Their major functions are illustrated.

Dame (CSE 40838). FlowVisual was demoed in class soon after a related concept or algorithm was explained. Students explored the tool during and after the class to reinforce their learning. FlowVisual also integrates a quiz module for student assessment. Its effectiveness in helping students understand flow visualization concepts was evaluated through a user study. Analysis of the user study data indicated that (1) regardless of the background of the students, FlowVisual can make a difference by actively engaging students in learning flow visualization concepts; and (2) a comprehensive pedagogical tool for data visualization is warranted as most students agreed that they would like to see more tools like FlowVisual helping them learn other visualization concepts and algorithms. These findings motivate us to expand the scope from FlowVisual to VisVisual, including other data visualization teaching and learning tools.

WHY VISVISUAL?

There exist visualization tools such as ParaView, VisIt, Graphviz, and Gephi. However, these packages are either rather complex packages (ParaView, VisIt) that encompass diverse visualization techniques or somewhat specialized libraries (Graphviz, Gephi) that target a particular means of visualization (e.g., graph visualization). They are mainly geared toward researchers and developers rather than beginning students. It could take a significant effort for beginners to grasp the necessary intricacies to use the tool and produce meaningful results, which inevitably deviates the learning focus from *understanding visualization concepts* to *mastering a complex package*. Therefore, it comes as no surprise for students to complain that it requires learning new, unfamiliar software and that this is asking too much of students [9]. Instructors should be aware of the complexity of visualization and its significance when developing tools to minimize learning difficulties.

The goal of an educational experience is for students to understand a certain principle and apply that information to solve problems in various situations. On the instructor side, designing cognitively engaging activities as part of the curriculum can increase the likelihood of learning from visualizations [8]. However, integrating a visualization toolkit into curriculum design poses a formidable challenge, persuading instructors away from adopting any visualization tools.

For effective learning of key SciVis and InfoVis algorithms and techniques, my research group designs and develops VisVisual, a toolkit, along with a set of resources, that integrates the tool into the curriculum. VisVisual can be a powerful means for active learning, contributing to teaching demonstrations, hands-on laboratory activities, technology use, and learning assessment. Figure 1 shows an overview of the VisVisual toolkit and its major functions. The primarily targeted setting for VisVisual is upper-level undergraduate courses in Data Visualization with students from diverse fields. Individual tools can be used separately for other courses as well. Besides, anyone interested in learning data visualization could benefit from using this toolkit and related materials for self-learning.

GUIDING PRINCIPLES

We advocate the following guiding principles for developing and evaluating VisVisual:

Targeting beginning learners. We aim to explain key concepts, illustrate basic algorithms,

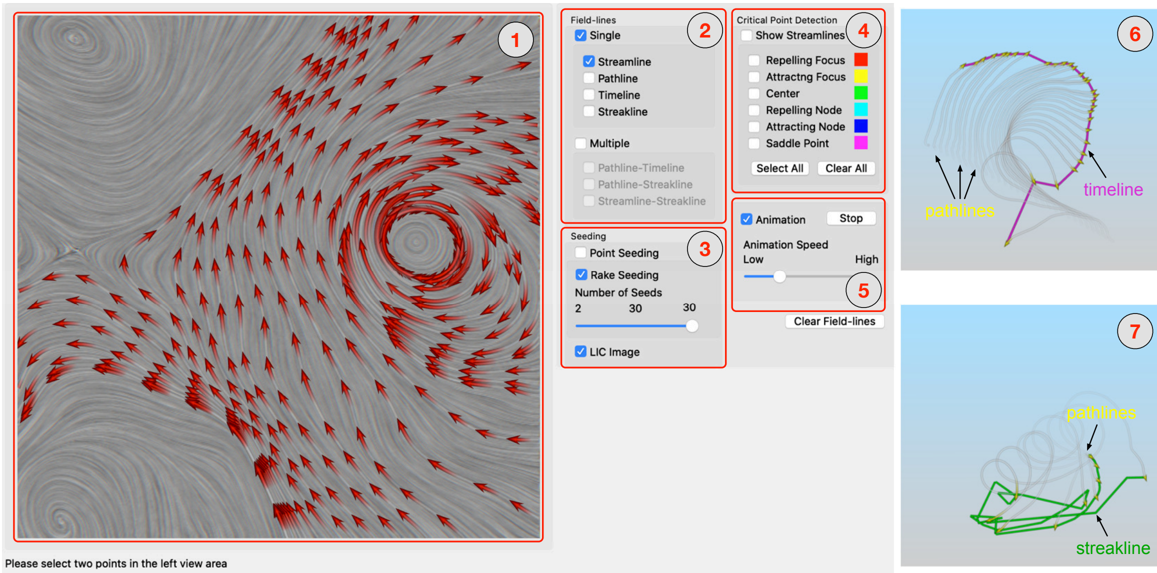

Figure 2. The FlowVisual desktop version shows rake seeding of streamlines in 2D and visualized animatedly (red arrows) and field line comparison. ① visualization panel. ② field line controls. ③ seeding controls. ④ critical point controls. ⑤ animation controls. ⑥ pathline-timeline comparison. ⑦ pathline-streakline comparison.

enable visual comparison, and facilitate critical understanding by creating a light-weighted toolkit for beginning users. As such, we include algorithms that beginners may find difficulty understanding without the aid of such a tool.

Selective and representative. In addition to straightforward forms of visualization as the baselines, we include non-trivial ones so that users can indeed gain insights. Therefore, we implement some of the most representative or widely used SciVis and InfoVis algorithms and techniques.

Simple is beautiful. Our visual design should be clean yet modern. Only parameters directly related to the understanding of targeted visualization algorithms will be available on the interface for interaction.

Interactivity is the key. VisVisual goes beyond what animations can offer by allowing users to interact with the software tool. They can change parameters, make selections, rotate the view, and compare results interactively.

Modular and web-friendly. We follow a modular design, making it easy to include other visualization techniques not yet planned. For easy deployment, we choose to use web-based programming languages such as WebGL and D3.

Integrating the curriculum. We provide curriculum materials to complement the VisVisual toolkit, easing instructors' burden to design teaching materials from scratch and broadening the toolkit's dissemination.

Community engagement. We aim to use and evaluate VisVisual at multiple institutions to engage students of diverse backgrounds and reach out to K-12 teachers and students from neighboring school districts for broader dissemination.

Evaluation is essential. VisVisual encompasses a comprehensive set of questions and tasks to evaluate the learning effectiveness. We design user studies to investigate the effectiveness of learner understanding.

TOOL COMPONENTS AND CURRICULUM MATERIAL

FlowVisual. The development of FlowVisual [13], [12] was carried out in two phases and over three years (2012-2015). We designed and implemented a desktop version to illustrate basic vector field concepts in 2D in the first phase (refer to Figures 2). The desktop version runs on various operating systems (Windows, Linux, and macOS). In the second phase, we developed an app version to illustrate flow visualization concepts in 3D. The app version was freely accessible in Apple's App Store. However, it was decommissioned in 2021 due to its low popularity

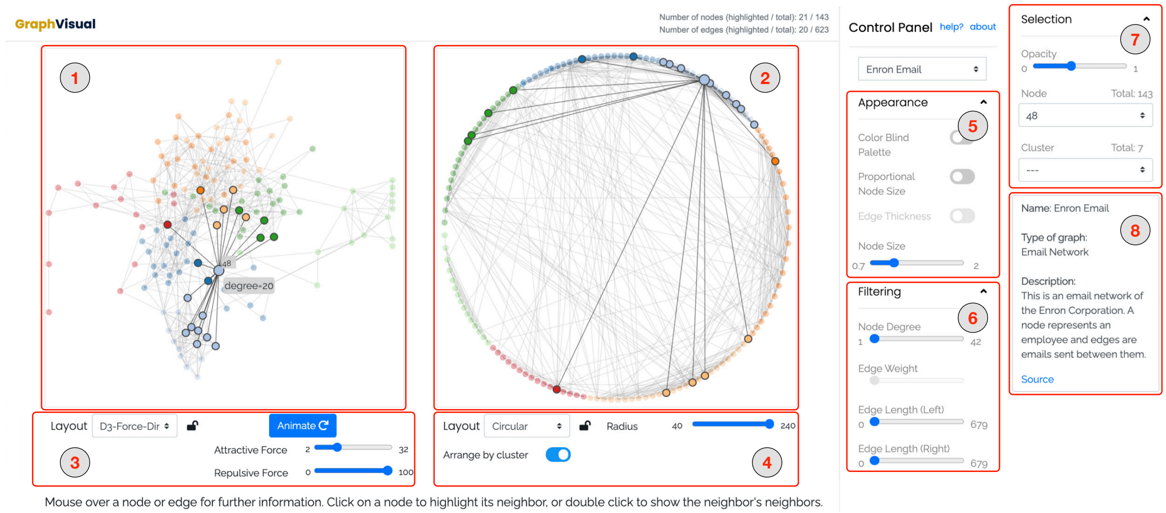

Figure 3. The web interface of GraphVisual shows a side-by-side comparison of the same graph dataset under two different layouts. ① and ② graph layout panels. ③ and ④ graph layout controls. ⑤ appearance controls. ⑥ filtering controls. ⑦ selection controls. ⑧ dataset information display.

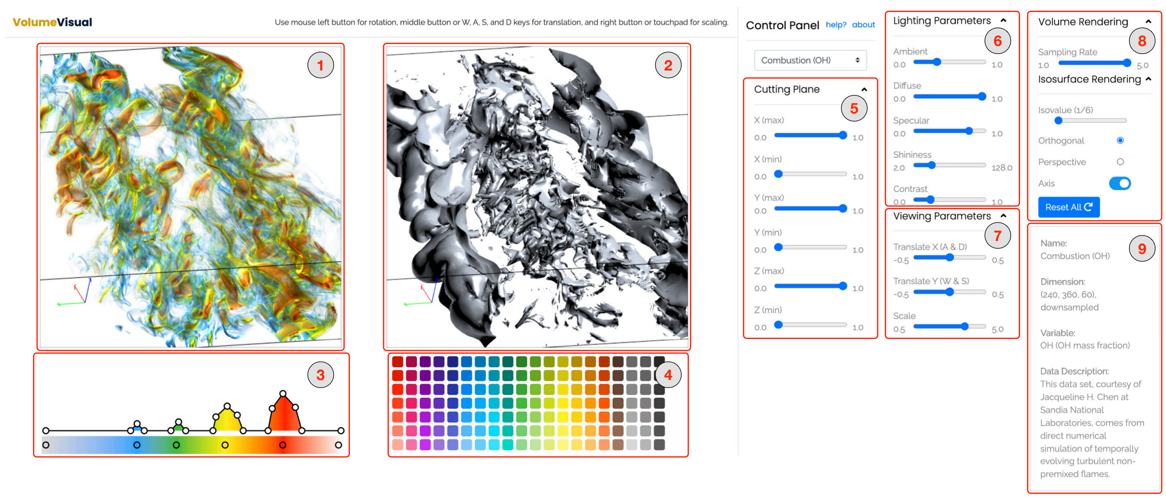

Figure 4. The web interface of VolumeVisual shows DVR and IR of the same dataset in the synchronized view. ① DVR panel. ② IR panel. ③ opacity transfer function editor. ④ color picker. ⑤ cutting plane controls. ⑥ lighting parameter controls. ⑦ viewing parameter controls. ⑧ DVR and IR controls. ⑨ dataset information display.

compared with its desktop counterpart.

We released FlowVisual and a tutorial online, which provides other instructors with a useful aid for updating their curriculum and teaching practice. Since the web page's launch in June 2013, the page has been accessed 10,000+ times within a few years, and the tool has been downloaded 1,000+ times. The pilot experiment of FlowVisual set up the groundwork for a successful NSF grant to develop the other three tools for teaching and learning data visualization. All these three tools would be web-based.

GraphVisual. The development of GraphVisual [5] was completed over one year (2018-2019). The primary goal of GraphVisual is to help users understand popular layout algorithms via visual comparisons. Therefore, the interface (refer to Figure 3) includes two side-by-side graph layout panels, displaying the same dataset under two different layouts. We released the GraphVisual

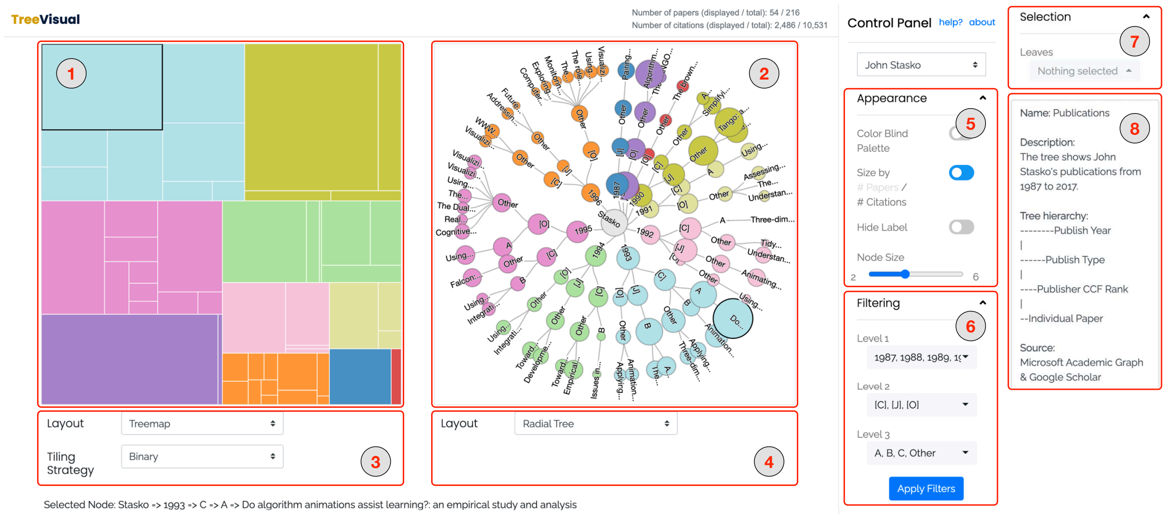

Figure 5. The web interface of TreeVisual shows a side-by-side comparison of the same tree dataset under two different layouts. ① and ② tree layout panels. ③ and ④ tree layout controls. ⑤ appearance controls. ⑥ filtering controls. ⑦ selection controls. ⑧ dataset information display.

tool in February 2020, which has been accessed 500+ times since then.

VolumeVisual. The development of VolumeVisual [1] was completed over one year (2019-2020). VolumeVisual is designed to help users understand essential volume visualization techniques, including direct volume rendering (DVR) and isosurface rendering (IR). We displayed synchronized renderings through side-by-side rendering panels so that users could better understand the similarities and differences between DVR and IR. Such an example is shown in Figure 4 ① and ②. We released the VolumeVisual tool in April 2021.

TreeVisual. The development of TreeVisual [7] was completed over two years (2020-2022). The main goal of TreeVisual is to help users understand the similarities and differences between different hierarchical layouts through visual comparison. Similar to GraphVisual, TreeVisual includes two side-by-side panels displaying the same hierarchical dataset under two different tree layouts (refer to Figure 5 ① and ②). We released the TreeVisual tool in May 2022.

Quiz design. Each tool of VisVisual comes with two modes: *study mode* and *quiz mode*. The study mode demonstrates the tool's capabilities as described above, intended for instructors in classroom teaching and students for self-exploration to reinforce the learning. The quiz mode includes a repository of questions integrated into the respective tool, and instructors can administer it to gauge student learning outcomes. The questions test students' general knowledge about the given topics and their ability to perform dataset-specific tasks. The questions may ask students to make single or multiple choices, provide written answers or perform visualization tasks leveraging the existing interface. Students answer questions or perform tasks in sequence akin to a standardized computer-based test. The online quiz can optionally record the time spent on each question.

Curriculum material. In addition to tool development, we developed and released curriculum material to help instructors cover selected visualization topics. The provision of curriculum material can decrease the barrier to use among other instructors and increase possible external adoption. The curriculum material includes a teaching plan, sample slides, introduction document, online tutorial, tutorial video, study guide, and survey questions. The VisVisual toolkit and related materials can be accessed at https://sites.nd.edu/chaoli-wang/visvisual/.

With VisVisual, the suggested teaching practice is as follows: First, the instructor lectures on the given visualization topic (slide presentation, example demos). Then, students form groups to discuss the pros and cons of each algorithm or technique and brainstorm the remaining chal-

Table 1. Means and standard deviations of VisVisual quiz results.

	FlowVisual Desktop	FlowVisual App	GraphVisual	VolumeVisual	TreeVisual
# Students	22	21	23	18	26
# Questions	13	15	22	21	27
Avg Score	0.79	0.82	0.82	0.88	0.93
Std Dev	0.15	0.16	0.10	0.13	0.24

Table 2. Means and standard deviations of FlowVisual survey results.

	Flow Field	Streamline	Pathline	Streakline	Timeline	Animation	Critical Point	UI Color	UI Size
	FlowVisual Desktop (22 ratings)								
Avg Score	4.68	4.33	4.53	4.08	4.57	4.42	4.49	4.53	4.20
Std Dev	0.50	0.83	0.68	0.92	0.62	0.77	0.60	0.68	0.67
	FlowVisual App (28 ratings)								
Avg Score	4.42	4.44	4.52	4.49	4.73	4.55	4.44	4.45	4.49
Std Dev	0.69	0.71	0.72	0.82	0.45	0.70	0.61	0.63	0.52

Table 3. Means and standard deviations of GraphVisual, VolumeVisual, and TreeVisual survey results.

	Layout	Dataset	UI	General	Dataset	UI	General	Layout	Dataset	UI	General
	GraphVisual (21 ratings)				VolumeVisual (20 ratings)			TreeVisual (26 ratings)			
Avg Score	4.37	4.16	4.63	4.31	4.36	4.40	4.91	4.47	4.75	4.46	4.71
Std Dev	0.63	0.66	0.59	0.78	0.49	0.35	0.19	0.79	0.43	0.80	0.64

lenges for visualization. The instructor demonstrates the tool (study component) in the next class. After class, students go over the study guide and explore the tool as an ungraded homework assignment. Students take the in-class quiz (quiz component) and complete an online survey a week later. Optionally, the instructor reports student survey summary and quiz performance to the class.

USER STUDY AND EXTERNAL EVALUATION

User study design. For each VisVisual tool, we designed a user study to assess the learning effectiveness. We recruited students taking the Data Visualization classes and additional students in the user study. Both undergraduate and graduate students were involved. These students major in computer science and engineering, aerospace and mechanical engineering, electrical engineering, and applied and computational mathematics and statistics. The students were first briefed with a review of visualization concepts and the corresponding tool (study mode). They were given a week to get familiar with the content and explore the tool. They then proceeded with the study, including an online quiz (quiz mode) and a survey questionnaire.

The quiz ranges from 13 (FlowVisual Desktop) to 27 (TreeVisual) questions. Quiz questions are comprehensive, including questions not related to the tool itself (e.g., "*Which techniques are the most often used for optimizing the efficiency of direct volume rendering?*" in VolumeVisual) and questions related to the tool, which could be dataset-related (e.g., "*Assuming that the edge length in the drawing represents the length of a fiber tract, which two nuclei are the farthest apart?*" in GraphVisual using the Cat Brain dataset) or dataset-unrelated (e.g., "*What is the benefit of Radial Tree (Collapsible) compared to Radial Tree?*" in TreeVisual). The datasets used in the quiz are new, which the students did not explore in the study mode. Often, the students must interact with the tool to get the correct answers to tool-related questions. During the quiz, the students were only allowed to ask questions regarding tool use. Quiz questions are presented one by one. Once a question is answered, the students cannot go back for review or correction. The quiz was not timed, but most students could complete it within one hour.

Finally, the students completed the survey. The survey ranges from 20 (VolumeVisual) to 28 (FlowVisual App) rating questions with a five-point Likert scale from "strongly disagree" (1) to "strongly agree" (5). It also has an open text question soliciting general feedback.

User study results. To analyze the user study data, we computed the mean and standard deviation of each question or student group and applied statistical tests (e.g., MANOVA). Table 1 summarizes student performances on quiz questions where the score ranges from 0 to 1. Tables 2 and 3 show the summary of student ratings on survey questions by categories (nine for FlowVisual, four for GraphVisual and TreeVisual, and three for VolumeVisual). The details of user study and result analysis with student group breakdowns (e.g., by major, level, topic familiarity, taking the class or not) were reported in [13], [12], [5], [1], [7]. The quiz results show that the tool effectively facilitates the learning of visualization concepts. The rating scores indicate that the students' reactions were very positive, and their ratings were similar despite their discipline and education differences. Therefore, regardless of student background, the tool is effective and helpful for classroom teaching and self-learning. The students' comments indicate that they are satisfied with the tool, including the design, interface, and interaction.

External evaluation. External evaluators have also assessed individual tools of VisVisual. Three professors from other institutions who conduct visualization teaching and research were involved in the evaluation. They went over the tool (including study and quiz modes) and support materials and completed a professional survey independently administered by a third party. The evaluation shows that they were in solid agreement across different categories of questions, including topics, dataset selection, user interface and interaction, quiz design, user experience, and overall impression and recommendation. They also concurred that they would likely use the VisVisual tools in a future class and recommend the toolkit to other colleagues. They would like to see VisVisual supporting users explore their own datasets in terms of improvement. When students construct their own answers with the datasets supplied by themselves and develop personally relevant cues for memory, they are more likely to remember that information or how to apply similar datasets.

CONCLUDING REMARKS

Visualization has become vital for analyzing data generated from diverse applications across many fields. As more and more institutions pay attention to research and education in data science and human-centered computing, expanding their curricula to include a new Data Visualization course becomes a growing trend. This trend calls for high-quality curriculum materials that help teach and learn visualization knowledge. We present VisVisual to fill the void by designing a suite of tools and curriculum material to help under-resourced institutions or underprepared instructors integrate visualization content into teaching more quickly. The toolkit can be used extensively in a Data Visualization course or selectively in other related courses that only cover a particular topic, such as Computational Fluid Dynamics (FlowVisual), Medical Imaging (VolumeVisual), and Visual Literacy or Network Science (GraphVisual and TreeVisual). The instructors can only use the tool to demonstrate a concept or incorporate it fully into teaching, learning, and performance assessment.

My research group will maintain the VisVisual toolkit and consider enabling users to selectively explore their own datasets. Before the COVID-19 pandemic, we offered a seminar on data visualization to around 20 K-12 teachers from the local school districts in early 2020. This two-day Notre Dame Teachers as Scholars (TAS) seminar features talks, group discussions, hands-out experiments, and an introduction to graph visualization using GraphVisual. The seminar was well-received as the post-survey shows that the participating teachers were satisfied with the format and content. Unfortunately, the subsequent outreach activities planned were canceled due to the pandemic. We want to resume more community engagement activities as the situation improves for broader dissemination.

The pedagogy of data visualization is becoming increasingly crucial as visualization tools and techniques flourish. Many instructors realize the value of developing an interactive software tool for teaching data visualization, yet the time it takes to create such a tool becomes the primary deterrent. The VisVisual toolkit will meet this need, making it less time-consuming and more convenient for instructors to "*use visualization to teach visualization*." We hope our VisVisual effort will inspire the community to pay due attention to this crucial but often-overlooked topic of visualization education. While there have been illuminating discussions on the need for visualization education and what to teach in a data visualization class, our work will contribute to further discussion on how to best support teaching data visualization.

ACKNOWLEDGMENT

This research was supported in part by the U.S. National Science Foundation through grant DUE-1833129. The author would like to thank all developers, collaborators, and evaluators who contributed to the VisVisual project and the anonymous reviewers for their constructive comments.

REFERENCES

1. X. Bao, J. Han, and C. Wang. VolumeVisual: Design and evaluation of an educational software tool for teaching and learning volume visualization. In *Proceedings of American Society for Engineering Education Annual Conference*, pages 32.259.1-32.259.14, 2021.
2. K. Borner and Y. Zhou. A software repository for education and research in information visualization. In *Proceedings of International Conference on Information Visualisation*, pages 257–262, 2001.
3. M. Bostock, V. Ogievetsky, and J. Heer. D^3 data-driven documents. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2301–2309, 2011.
4. J.-D. Fekete. The InfoVis toolkit. In *Proceedings of IEEE Symposium on Information Visualization*, pages 167–174, 2004.
5. M. Imre, W. Chang, S. Wang, C. P. Trinter, and C. Wang. GraphVisual: Design and evaluation of a web-based visualization tool for teaching and learning graph visualization. In *Proceedings of American Society for Engineering Education Annual Conference*, pages 28.501.1–28.501.15, 2020.
6. T. L. Naps, G. Rössling, V. Almstrum, W. Dann, R. Fleischer, C. Hundhausen, A. Korhonen, L. Malmi, M. McNally, S. Rodger, and J. Á. Velázquez-Iturbide. Exploring the role of visualization and engagement in computer science education. *ACM SIGCSE Bulletin*, 35(2):131–152, 2002.
7. B. O'Handley, Y. Wu, H. Duan, and C. Wang. TreeVisual: Design and evaluation of a web-based visualization tool for teaching and learning tree visualization. In *Proceedings of American Society for Engineering Education Annual Conference*, 2022, Accepted.
8. D. N. Rapp. Mental models: Theoretical issues for visualization in science education. In J. K. Gilbert, editor, *Visualization in Science Education*, pages 43–60. Springer, 2007.
9. M. Stieff, R. C. Bateman, and D. H. Uttal. Teaching and learning with three-dimensional representations. In J. K. Gilbert, editor, *Visualization in Science Education*, pages 93–120. Springer, 2007.
10. Y. Tanahashi, N. Leaf, and K.-L. Ma. A study on designing effective introductory materials for information visualization. *Computer Graphics Forum*, 35(7):117–126, 2016.
11. F. B. Viegas, M. Wattenberg, F. van Ham, J. Kriss, and M. McKeon. Many Eyes: A site for visualization at Internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128, 2007.
12. M. Wang, J. Tao, J. Ma, Y. Shen, and C. Wang. FlowVisual: A visualization app for teaching and understanding 3D flow field concepts. In *Proceedings of IS&T Conference on Visualization and Data Analysis*, pages 476–1–476–10, 2016.
13. M. Wang, J. Tao, C. Wang, C.-K. Shene, and S. H. Kim. FlowVisual: Design and evaluation of a visualization tool for teaching 2D flow field concepts. In *Proceedings of American Society for Engineering Education Annual Conference*, pages 23.609.1–23.609.20, 2013.

Chaoli Wang is an associate professor in the Department of Computer Science and Engineering at the University of Notre Dame. Contact him at chaoli.wang@nd.edu.