Interactive Data Visualization: A Petri Dish for Human-Computer Complementarity

Professional Statement, Arvind Satyanarayan, MIT EECS & CSAIL

> *"To think in interaction with a computer in the same way that you think with a colleague whose competence supplements your own will require much tighter coupling between [human] and machine."*—J. C. R. Licklider, Man-Computer Symbiosis, March 1960.

How can we harness advances in machine automation to promote human agency? This question has long-motivated human-computer interaction (HCI) dating back to Licklider's vision of having the strengths of computers complement the limitations of humans and vice-versa. The rapid advances of artificial intelligence (AI) have lent new urgency to this question as generative AI, which increasingly demonstrates competence at tasks previously thought to require human creativity, threatens to deskill and/or displace human labor. **My research investigates this question through the lens of interactive data visualization, a domain that exemplifies the promises and tensions of balancing automation and agency.** Society has embraced visualization—from scientific discovery to business intelligence, journalism, and public policy—as a critical medium for recording, analyzing, and communicating data. While automation plays an important role in this domain—for instance, by optimizing data processing or recommending perceptually effective designs—a visualization is only *insightful* because it allows a person to apply their perceptual and cognitive abilities to make observations about data, and then interpret these observations for their given task, data domain, and sociocultural context.

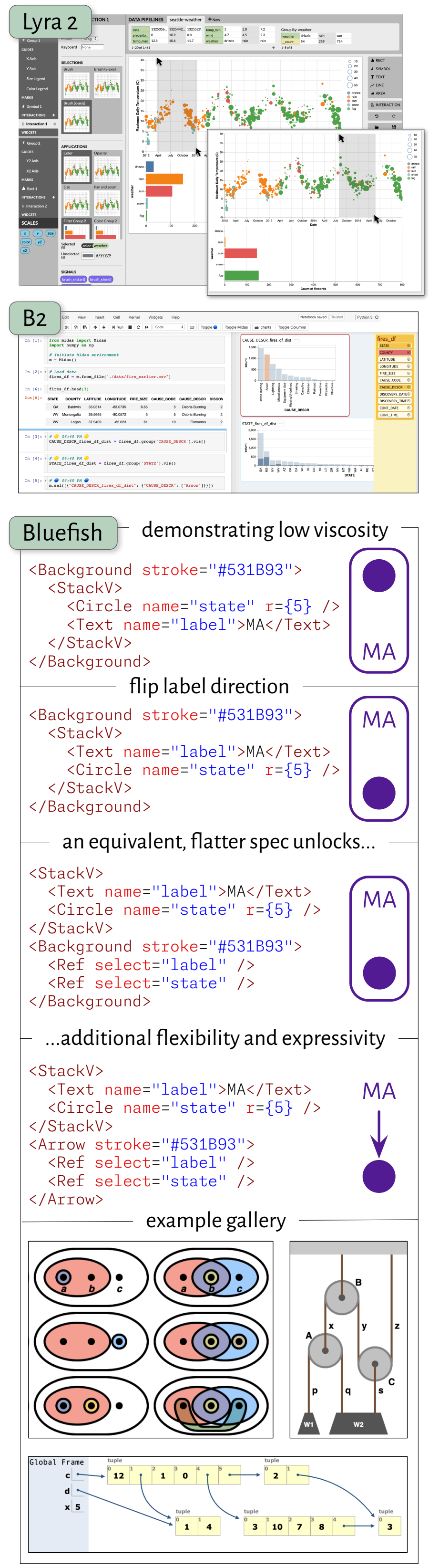
My research group studies complementarity via the computational representations that mediate interactions between humans and computers**. We design *domain-specific languages* for interactive visualization and user interfaces**, and explore how these languages can serve as *shared representations*: artifacts that users manipulate as part of a (graphical or textual) interface, and that automated systems reason about to accelerate a user's workflow. In parallel, to help users inspect automated methods, **we develop visual analysis techniques for comparing neural networks' *learned representations* with human-designed counterparts**, and explore how (mis)alignments are opportunities for uncovering novel abstractions of a domain's semantics. Finally, **we bridge human-designed and machine-learned representations to make visualizations accessible to blind and low-vision people**. By reflecting on these threads of work, I have also developed a **novel conceptual model of complementarity grounded in agency** that offers a roadmap for future research.

My group's research outcomes have been broadly disseminated and widely adopted. Our visualization languages, Vega and Vega-Lite, are used in industry (e.g., at Apple, Google, Microsoft, etc.), in journalism (e.g., at the LA Times), on Wikipedia, by the Jupyter data science community [35], and by researchers across diverse scientific disciplines (e.g., astrophysics [39], biology [28], electrochemistry [15], etc.). Moreover, the visualization research community is adopting our work on accessibility, including our accessible visualization toolkit Olli, to serve as the foundation for empirical research [10,29,40], benchmark evaluations [6,9], and systems-building [7,27,34]. An interactive article we published about our ethnography of COVID-19 visualizations has been read by over 35,000 people after repeatedly going viral on social media, and led to **invited presentations to the US Department of Health and Human Services and the UK Prime Minister's Office.** Finally, our work has been recognized with Best Paper awards and honorable mentions across a range of research communities (IEEE VIS, EuroVis, ACM CHI, ACM IUI, ACL), as well as the NSF CAREER award, Google Research Scholar Award, IEEE VGTC Significant New Researcher Award, and an Alfred P. Sloan Fellowship.

DESIGNING DECLARATIVE DOMAIN-SPECIFIC LANGUAGES TO GROW USERS' EXPRESSIVE POWER

The first thread of my research focuses on users' *expressive power*, targeting tasks that either require enormous effort or are altogether out-of-reach. I develop domain-specific languages (DSLs): textual interfaces with primitives grounded in the semantics of a particular domain. Critically, following an approach of complementarity, I design these languages to be *declarative*: users compose these primitives to specify *what* they would like to produce, and the underlying language runtime automatically synthesizes low-level details to determine *how* the necessary execution occurs. Thus, I achieve expressive gains by both designing for cognitive efficacy (e.g., ease-of-use) and by exploring novel inferences that language runtimes and graphical systems can perform to support users.

My students, collaborators, and I have conducted this work largely atop the declarative visualization stack from my PhD work [22,24–26] to address the task of authoring dynamic visualizations. At the lowest level of the stack, inspired by my Reactive Vega language [25,26], we developed **DIEL** [37]: a middleware framework that extends the benefits of declarative specification to scenarios where backing data is distributed across multiple locations, or where interactive events cannot be synchronously processed to completion. DIEL supports these use cases by **adapting two abstractions from distributed databases:** maintaining an *immutable log* of interaction events, and defining interactive state as *logical constraints* over these log entries. In contrast to prior approaches, **DIEL offers a unified data model for "live" and "stored" data**, and the DIEL runtime maintains an event loop to reevaluate data state as new events occur (e.g., in response to remote database calls).



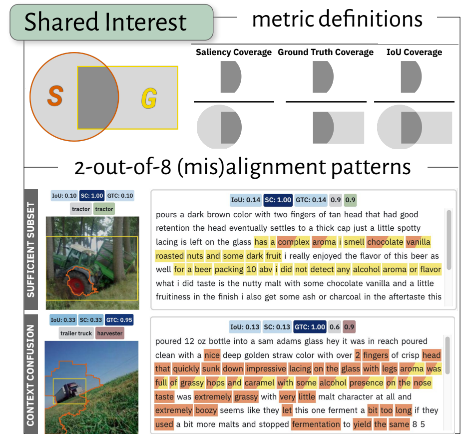
At the mid-level of the stack, we made a key observation: **interaction and animation are parallel concepts.** Where interaction transforms data or updates visual properties in response to *user input*, animation does the same in response to *timers*. This parallelism is especially important in immersive environments (i.e., augmented, mixed, and virtual reality) as interaction unfolds in more embodied ways (e.g., through gestures). Thus, animation becomes a crucial mechanism for maintaining *congruency*—that is, helping users map their embodied interactions to changes to the visualization. Armed with these insights, we developed **Animated Vega-Lite** [43], extending our Vega-Lite grammar [24] with a unified set of abstractions for authoring static, interactive, and animated 2D visualizations; and, **Deimos** [11], the first toolkit to unify these concepts for immersive 3D visualizations, defining congruency as a set of animated transitions between partially-specified interactive states.

Finally, at the highest level of the stack, we have explored how these DSLs can serve as *shared representations* by **developing novel systems that reason about interactive visualization state and design** in terms of Vega and Vega-Lite. **Lyra 2** [41] is a graphical interface for authoring interactive visualizations without coding. Direct manipulations (e.g., dragging data fields and dropping them onto the canvas) and demonstrations (e.g., performing mouse clicks or drags on the designed visualization) are heuristically parsed to generate corresponding statements in Vega or Vega-Lite. Similarly, **B2** [38] tightens the feedback loop between coding and interactive analysis in computational notebooks by automatically synthesizing Vega-Lite visualizations based on data frame provenance, and by reifying interaction states as a semantically-meaningful log of data queries in a code cell.

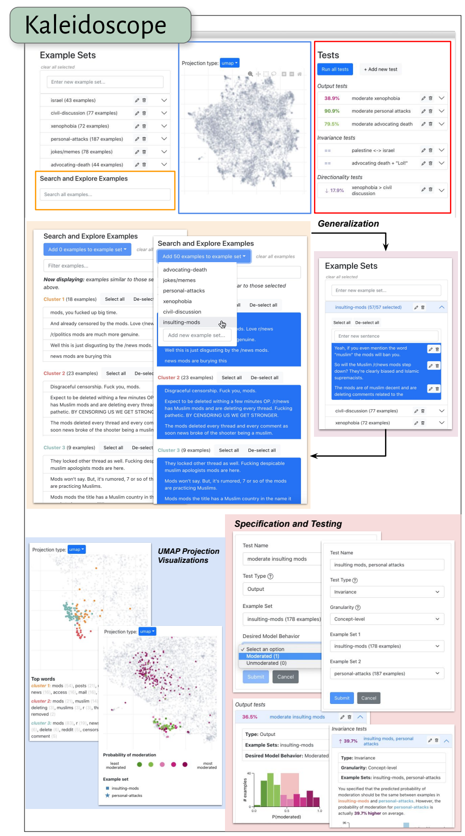
Through this work, we have been able to identify that a **high expressive ceiling is not sufficient for increasing users' sense of agency**. Rather, high expressivity must be coupled with **low viscosity: an interface language should allow users to nimbly move through alternate statements** (e.g., switching between client-side or client-server data distributions with DIEL, or mixing static, interactive, or animated modalities with Animated Vega-Lite). We have since investigated how these twin principles generalize to domains besides interactive visualization: with **Bluefish** [20] for authoring *graphical diagrams* such as those found in mathematics, computer science, and physics; and via **Varv** [5] for developing *malleable user interfaces* (UIs) that end-users can reprogram at runtime. To enable expressive and fluid composition in their respective domains, both toolkits take modern UI frameworks (e.g., React) as a useful point of departure. Bluefish relaxes the definition of a UI component such that a parent container need not fully specify the size and layout of its children, and computes a *relational* scenegraph—a compound graph structure that captures both hierarchical and adjacency relationships. Visual elements can, thus, participate in spatial arrangements that crosscut the nested hierarchies of UIs. And, through a series of high-level declarative statements rather than low-level imperative commands, **authors can smoothly flatten a specification to unlock spaces of alternate designs** (as the figure shows). Varv, as UI architecture pattern, can be thought of as (MC)V, or ModelController-View: it defines UIs in terms of *concepts*, declarative encapsulations of data and view-agnostic actions that define how the data can be manipulated; and, *templates*, which reify any given concept into myriad UI fragments. **This separation of concerns elevates interaction as the central focus of UI design, making it composable and extensible purely *through accretion* rather than modification**— akin to introducing new CSS stylesheets to override existing rules.

Alongside building toolkits—which helps identify design principles and usefully populates the design space for growing users' expressive power—my research also **empirically studies the cognitive [36] and sociotechnical implications [12] of greater expressive power**. For instance, at the height of the COVID-19 pandemic, my research group conducted a 6-month digital ethnography of social media discourse to discover **large groups of users marshaling data science and visualization in ways that exemplify citizen science**—including critically assessing data sources, debating the validity of various metrics, and designing polished visualizations—**but *in opposition to* public health guidance** [12]. Our findings expose urgent tensions in data visualization research: a focus on simplicity, clarity, and binary categories of data literacy/illiteracy, do not capture these groups' skillful manipulation of data. In ongoing work, we are further unpacking these tensions. By drawing on sociolinguistic theory, we show that such visualizations do more than simply encode data; they also function as *social* artifacts that mediate relationships and identities between individuals, groups, and institutions.

ALIGNING LEARNED REPRESENTATIONS TO ENABLE INTERPRETABILITY AND KNOWLEDGE DISCOVERY



Despite their astonishing capabilities, deploying neural networks is fraught with risk because it is difficult to understand how they operate, or determine how fair or equitable their output is. In response, researchers have begun to contribute techniques for explaining a model's output in terms of the features of its input data (e.g., the pixels of an image, words in a sentence, etc.). While valuable, **the lens of agency highlights why a purely technical approach isn't sufficient:** models are situated in specific sociocultural contexts and, as my students and I theorize [30] and demonstrate [31], people have rich formal and tacit knowledge as well as lived experience that must be brought to bear to interpret and adapt model outputs within these contexts. However, through formative interviews [3], we find **existing interpretability techniques poorly support users in this task**. Users interact with explanations in ad hoc ways which, at best, requires non-trivial effort to distill actionable insight and, at worst, undermines their confidence in the validity of their findings.



To facilitate more systematic interpretation of machine-learned representations, my students and I have developed novel visual analysis methods inspired by *Jaccard similarity*: a measure of how similar and/or diverse two sets are, calculated as their intersection-over-union. With the **Embedding Comparator** [3], we use Jaccard similarity to define the *reciprocal local neighborhood:* for every instance in a dataset, how many of its nearest neighbors are shared between a pair of models, and how many are unique to a specific model. Similarly, with **Shared Interest** [4], we compute the alignment between model saliency (a calculation of how important each input feature is to the final output) and human-annotated ground truth by **composing Jaccard similarity with proportionality measures analogous to precision and recall** (see figure). Across both systems, we use these metrics to organize users' exploration of models' latent spaces and accelerate investigation of specific data instances. For example, with Shared Interest, users can structure their analyses around eight (mis)alignment patterns that we identified as recurring across computer vision and natural language classification models (two patterns shown in the figure). In evaluations, we found that, by grounding interpretability interfaces in task-based metrics, **domain experts explored models' latent spaces more broadly and deeply, and reported greater confidence** in their results than using their existing Jupyter-based workflows [3,4].

Despite this success, neither system fully "closes-the-loop"—that is, once a user makes an observation, they are left to their own devices to take any follow-up action. To explore a more end-to-end workflow, my students and I developed **Kaleidoscope [32]: a system for testing model behavior through iterative interpretation.** Users start by identifying data instances that express a concept they would like to test, and then successively generalize the concept's definition with additional relevant instances that Kaleidoscope retrieves by computing a distance metric. Selected instances are then used to test model output or to test semantically meaningful shifts in model behavior. We instantiated Kaleidoscope to test a comment moderation model with 13 Reddit moderators, and found that **its iterative workflow not only preserved the affordances of systematic analysis** (e.g., helping participants define concepts more comprehensively), it also **helped draw out participants' implicit knowledge, and reason about context-specific tradeoffs**. For example, when testing the moderation of racist comments, the retrieved instances helped a participant identify a distinction between outrightly offensive comments and those that disguise racist sentiment behind lengthy arguments—the latter set proves challenging for even human moderators to evaluate. This finding echoes an observation made by a board-certified dermatologist who, when analyzing a melanoma-detection model with Shared Interest, alluded to the opportunity for deepening his domain knowledge: *"[the model] is seeing something we're not truly appreciating in the clinical image [...] Maybe there's really subtle changes we're not picking up that it is able to."*

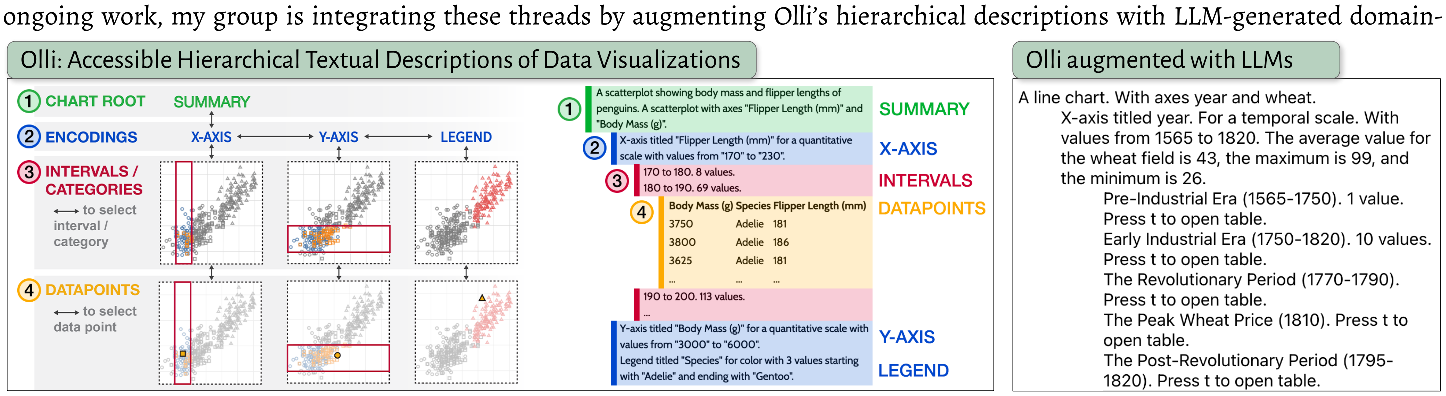
To facilitate such knowledge discovery through interpretability, my students, collaborators, and I have also **developed techniques for expressing learned representations in terms of the underlying domain semantics**. We have shown how an optimization process [16] transforms learned representations from abstract matrices into *semantic dictionaries* [17]: structures that map activations, or the amounts neurons fire, to symbolic representations of these neurons (e.g., iconic images for computer vision neurons). In doing so, we observe that **models learn abstractions for concepts in more nuanced ways than we have equivalent natural language to describe**—for instance, InceptionV1 has several "floppy dog ear" neurons that detect different levels of droopiness, length, and shape [17]. Alongside this "bottom-up" approach, my students and I have also developed a "top-down" approach we call **abstraction alignment** [2]: starting with a knowledge graph, we map a model's output classes to its leaves, and propagate confidence scores through this structure to determine how the conceptual relationships a model has learned correspond to human taxonomies. Here, **misalignments can signal opportunities to clarify or expand our knowledge.** For instance, in a case study with medical terms, abstraction alignment identifies discrepancies between how terms are meant to be used (i.e., as defined in guidelines maintained by the World Health Organization) and how they are used in practice; notably, our findings correspond to updates the WHO made between v9 and v10 of its guidelines.

BRIDGING DESIGNED AND LEARNED ABSTRACTIONS FOR ACCESSIBLE DATA REPRESENTATION

While the previous two threads of my research largely examine human-designed and machine-learned representations in isolation, the final thread of my research brings them together in a domain where striking the right balance between machine automation and human agency is particularly important and urgent: accessibility. As the COVID-19 pandemic demonstrated, visualization is a critical medium for analyzing and communicating data, but largely locks out blind and low-vision (BLV) people [13,14]. Recent advances in large language models (LLMs) suggest that AI has a socially beneficial role to play but, in a study we conducted [14], **BLV people expressed trepidation about how automated methods may foreclose opportunities to interpret data by and for themselves**. As one participant emphasized: *"I want to have the time and space to interpret the numbers for myself, before I read the analysis."*

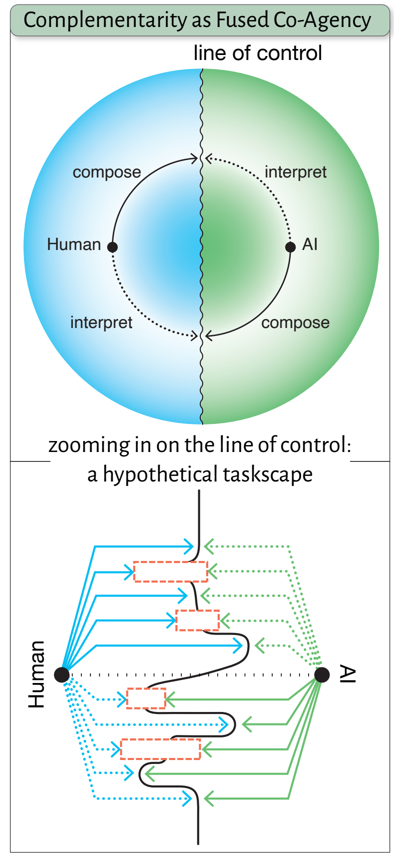
Through a co-design process with Daniel Hajas, a blind HCI researcher and science communicator at University College London, my students and I developed **a design space of richer experiences of visualizations with screen readers** [42]—a commonly used assistive technology that narrates on-screen content, but typically renders visualizations as either long lists of raw data values or as incomprehensible strings of *"image image image."* Thus, a key challenge is how to linearize the information depicted in a visualization while preserving the ability to read at multiple levels of granularity. Our design space suggests multiple serialization strategies and, through studies with BLV participants, we found **laying textual descriptions out in a tree structure** (as shown in the figure below) to be particularly promising. Branches correspond to different fields that are encoded in the visualization, and each level of the tree offers successively finer-grained resolution about the relevant data values. This structure can be navigated via the keyboard by moving up/down tree nodes, by moving along screen coordinates to replicate the experience of reading tactile graphics, or by jumping directly to specific points of interest. We have instantiated this approach as an **accessible visualization toolkit called Olli** [1], which ingests a declarative visualization specification (e.g., from Vega-Lite, Observable Plot, etc.) to generate the textual hierarchy. Descriptions are generated via tokenized string templates, which users can customize based on their context-specific needs (e.g., modulating verbosity) [8].

For more expressive textual descriptions of visualizations, my group has turned to LLMs. To understand what makes a textual description useful for BLV people, we collected and analyzed 2,000 visualization captions, identifying four levels of semantic content that they can convey [14]: chart construction (e.g., chart type, axis ranges); descriptive statistics (e.g., min, max); perceptual and cognitive features (e.g., complex trends and patterns); and, domain-specific context. Based on the preferences of 30 blind and 60 sighted participants, we focused on the first three levels and scaled our approach up to develop **VisText [33]: a benchmark for semantically-rich chart captioning, comprising a dataset of over 12,000 chart-caption pairs, and a baseline set of unimodal and multimodal models.** Besides scale and quality of data, VisText's key result is that models trained on *structured* representations of charts (e.g., scalable vector graphics, or SVG) outperform models trained on chart images and are just as viable as models trained on data tables. This result has important implications for wider deployment as SVG is the most common format for web-based visualizations and can be reverse-engineered from a rasterized chart image [18], whereas backing data tables can be difficult (if not impossible) to locate in-the-wild. In ongoing work, my group is integrating these threads by augmenting Olli's hierarchical descriptions with LLM-generated domain-



specific context (the fourth level of semantic content). For example, as the figure above shows, we use an LLM to group time-series data about wheat prices into significant periods such as *"The Pre-Industrial Era," "Early Industrial Era," "The Revolutionary Period,"* etc.

TOWARDS A THEORY OF HUMAN-COMPUTER COMPLEMENTARITY



By reflecting on my three lines of research, and through a deep collaboration with linguistic anthropologist Graham Jones, we observe that **existing AI and HCI paradigms position automation and agency as opposite ends of a spectrum.** AI measures intelligence in terms of competence, which makes it difficult to assess whether a system's response constrains human action, or to systematically imagine alternative, more empowering responses. HCI models interaction as an input-output loop, which largely treats the computer as an inanimate tool to be used by human agents. **Neither approach captures the richer interaction dynamics that generative AI increasingly enables**: its capacity for context-sensitive responses yields open-ended interaction flows where outcomes are more emergent and thus feel more *co-produced* than before, often in ways that surprise users, model developers, and designers.

In a recent white paper [23], Jones and I **redefine interaction and intelligence in terms of agency**: *compositional* agency, or the capacity to express a goal-directed intent; and *interpretive* agency, or the ability to evaluate the efficacy of this expression. Thus, **complementarity occurs by fusing the agencies of a human user and automated system together**—with different fusions being desirable for different tasks or contexts, or by different cultures. Critically, whereas prior approaches imagine a relatively stable balance of control between a human and computer that is fixed by the model developer or interface designer, **fused co-agency enables developers/designers to *delegate a region of control*** (the figure's dashed orange rectangles) **within which a user and computer can dynamically re-negotiate their roles over the course of an interaction**. In ongoing work, my students, collaborators, and I are operationalizing these ideas via an approach we call **programming with semi-formal representations** [19,21]: allowing a user to fluidly vary their compositional agency by mixing modalities and notational systems (e.g., code, handwritten math formula, a screenshot of a data table, etc.) to express program semantics that smoothly bridge multiple levels of abstraction, and having the computer flexibly respond (e.g., following directive and/or supportive communication styles).

TEACHING, MENTORING, AND SERVICE

At MIT, I co-instruct the undergraduate course *"6.1040: Software Design"* with Prof. Daniel Jackson, and developed a new graduate course *"6.859: Interactive Data Visualization"* which draws ~130 students from diverse disciplines (including computer science, architecture, design, urban planning, and business). Students have nominated me for MIT's **Teaching with Digital Technology Award multiple times**, and materials from 6.859 were upvoted to the front page of *Hacker News*. In collaboration with Profs. Sarah Williams and Catherine D'Ignazio (Urban Studies & Planning), I expanded 6.859 to a joint undergraduate/graduate course *"Interactive Visualization & Society."* Among many additions, this class features a weekly reading group where **~100 graduate students actively debate research papers on data ethics**. And, **final projects now address a topic of societal importance in collaboration with public policy organizations**. For example, last year, students worked with the World Food Program to visualize Central American migration into the United States. Class materials are being reused at several schools including Stanford University, Carnegie Mellon University, the University of Washington, and Université Paris-Saclay. My teaching has been recognized with the department's **Kolokotrones Education Award.**

I advise or have advised 6 PhD students and 2 Postdocs—5 of whom have started or secured faculty positions at MIT, Brown University, CU Boulder, the University of Utah, and UC San Diego. My PhD students and I have mentored over 30 junior researchers (including high school, undergraduate, and Master's students). These students have led or co-authored publications, and have contributed to our open-source projects. I prioritize cultivating a research group culture that is intellectually curious and stimulating, while also being supportive and inclusive. I have successfully advised several women and students of color and, based on student nominations, I was awarded the department's **Seth J. Teller Award for Excellence, Inclusion, and Diversity** and MIT's **Committed to Caring Award**.

I serve on organizing committees (IEEE VIS 2018–present), program committees (IEEE VIS, OpenVis Conf, Information+, ACM CHI, IUI, UIST), and in leadership (IEEE VIS 2026 General Chair, IEEE VGTC Publications Chair, Distill.pub co-editor). I devote significant effort to community building by organizing recurring BostonVIS events, a biennial PhD Summer School, and an annual workshop for visualization junior faculty. I organize the bi-weekly MIT CSAIL HCI seminar series which has fostered the HCI community at MIT and in greater Boston. To broaden participation in research, I serve on the MIT EECS Committee for Diversity, Equity, and Inclusion.

REFERENCES

1. Matt Blanco, Jonathan Zong, and Arvind Satyanarayan. 2022. Olli: An Extensible Visualization Library for Screen Reader Accessibility. In *IEEE VIS Posters*. Retrieved July 28, 2023 from https://mitvis.github.io/olli/
2. Angie Boggust, Hyemin Bang, Hendrik Strobelt, and Arvind Satyanarayan. 2024. Abstraction Alignment: Comparing Model and Human Conceptual Relationships. Retrieved July 18, 2024 from http://arxiv.org/abs/2407.12543
3. Angie Boggust, Brandon Carter, and Arvind Satyanarayan. 2022. Embedding Comparator: Visualizing Differences in Global Structure and Local Neighborhoods via Small Multiples. In *27th International Conference on Intelligent User Interfaces* (IUI '22), 746–766. https://doi.org/10.1145/3490099.3511122. **Best Paper Honorable Mention.**
3. Angie Boggust, Benjamin Hoover, Arvind Satyanarayan, and Hendrik Strobelt. 2022. Shared Interest: Measuring Human-AI Alignment to Identify Recurring Patterns in Model Behavior. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), 1–17. https://doi.org/10.1145/3491102.3501965. **Best Paper Honorable Mention.**
5. Marcel Borowski, Luke Murray, Rolf Bagge, Janus Bager Kristensen, Arvind Satyanarayan, and Clemens Nylandsted Klokmose. 2022. Varv: Reprogrammable Interactive Software as a Declarative Data Structure. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (CHI '22), 1–20. https://doi.org/10.1145/3491102.3502064
6. Sanjana Shivani Chintalapati, Jonathan Bragg, and Lucy Lu Wang. 2022. A Dataset of Alt Texts from HCI Publications: Analyses and Uses Towards Producing More Descriptive Alt Texts of Data Visualizations in Scientific Papers. In *Proceedings of the 24th International ACM SIGACCESS Conference on Computers and Accessibility*, 1–12. https://doi.org/10.1145/3517428.3544796
7. Joshua Gorniak, Yoon Kim, Donglai Wei, and Nam Wook Kim. 2024. VizAbility: Enhancing Chart Accessibility with LLM-based Conversational Interaction. https://doi.org/10.48550/arXiv.2310.09611
8. Shuli Jones, Isabella Pedraza Pineros, Daniel Hajas, Jonathan Zong, and Arvind Satyanarayan. 2024. "Customization is Key": Reconfigurable Textual Tokens for Accessible Data Visualizations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (CHI '24), 1–14. https://doi.org/10.1145/3613904.3641970
9. Shankar Kanthara, Rixie Tiffany Ko Leong, Xiang Lin, Ahmed Masry, Megh Thakkar, Enamul Hoque, and Shafiq Joty. 2022. Chart-to-Text: A Large-Scale Benchmark for Chart Summarization. *arXiv:2203.06486 [cs]*. Retrieved March 22, 2022 from http://arxiv.org/abs/2203.06486
10. N. W. Kim, G. Ataguba, S. C. Joyner, Chuangdian Zhao, and Hyejin Im. 2023. Beyond Alternative Text and tables: Comparative Analysis of Visualization Tools and Accessibility Methods. *Computer Graphics Forum* 42, 3: 323–335. https://doi.org/10.1111/cgf.14833
11. Benjamin Lee, Arvind Satyanarayan, Maxime Cordeil, Arnaud Prouzeau, Bernhard Jenny, and Tim Dwyer. 2023. Deimos: A Grammar of Dynamic Embodied Immersive Visualisation Morphs and Transitions. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3544548.3580754
12. Crystal Lee, Tanya Yang, Gabrielle D Inchoco, Graham M. Jones, and Arvind Satyanarayan. 2021. Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3411764.3445211. **Best Paper Honorable Mention.**
13. Alan Lundgard, Crystal Lee, and Arvind Satyanarayan. 2019. Sociotechnical Considerations for Accessible Visualization Design. In *IEEE Conference on Visualization and Visual Analytics (VIS)*.
14. Alan Lundgard and Arvind Satyanarayan. 2022. Accessible Visualization via Natural Language Descriptions: A Four-Level Model of Semantic Content. *IEEE Transactions on Visualization and Computer Graphics* 28, 1: 1073–1083. https://doi.org/10.1109/TVCG.2021.3114770
15. Matthew Murbach, Brian Gerwe, Neal Dawson-Elli, and Lok-kun Tsui. 2020. impedance.py: A Python package for electrochemical impedance analysis. *Journal of Open Source Software* 5, 52: 2349. https://doi.org/10.21105/joss.02349
16. Chris Olah, Alexander Mordvintsev, and Ludwig Schubert. 2017. Feature Visualization. *Distill* 2, 11: e7. https://doi.org/10.23915/distill.00007
17. Chris Olah, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye, and Alexander Mordvintsev. 2018. The Building Blocks of Interpretability. *Distill* 3, 3: e10. https://doi.org/10.23915/distill.00010
18. Jorge Poco and Jeffrey Heer. 2017. Reverse-Engineering Visualizations: Recovering Visual Encodings from Chart Images. *Computer Graphics Forum* 36, 3: 353–363. https://doi.org/10.1111/cgf.13193
19. Josh Pollock, Ian Arawjo, Caroline Berger, and Arvind Satyanarayan. 2024. Designing for Semi-formal Programming with Foundation Models. In *The 13th Annual Workshop at the Intersection of PL and HCI (PLATEAU 2024)*.
20. Josh Pollock, Catherine Mei, Grace Huang, Elliot Evans, Daniel Jackson, and Arvind Satyanarayan. 2024. Bluefish: Composing Diagrams with Declarative Relations. To appear in *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*. https://doi.org/10.48550/arXiv.2307.00146
21. Josh Pollock, Arvind Satyanarayan, and Daniel Jackson. 2023. Language Model Agents Enable Semi-Formal Programming.
22. Arvind Satyanarayan and Jeffrey Heer. 2014. Lyra: An Interactive Visualization Design Environment. *Computer Graphics Forum* 33, 3: 351–360. https://doi.org/10.1111/cgf.12391
23. Arvind Satyanarayan and Graham M. Jones. 2024. Intelligence as Agency: Evaluating the Capacity of Generative AI to Empower or Constrain Human Action. *An MIT Exploration of Generative AI*. https://doi.org/10.21428/e4baedd9.2d7598a2

22. Arvind Satyanarayan, Dominik Moritz, Kanit Wongsuphasawat, and Jeffrey Heer. 2017. Vega-Lite: A Grammar of Interactive Graphics. *IEEE Transactions on Visualization and Computer Graphics* 23, 1: 341–350. https://doi.org/10.1109/TVCG.2016.2599030. **Best Paper Award.**

25. Arvind Satyanarayan, Ryan Russell, Jane Hoffswell, and Jeffrey Heer. 2016. Reactive Vega: A Streaming Dataflow Architecture for Declarative Interactive Visualization. *IEEE Transactions on Visualization and Computer Graphics* 22, 1: 659–668. https://doi.org/10.1109/TVCG.2015.2467091

26. Arvind Satyanarayan, Kanit Wongsuphasawat, and Jeffrey Heer. 2014. Declarative Interaction Design for Data Visualization. In *Proceedings of the 27th annual ACM Symposium on User Interface Software and Technology* (UIST '14), 669–678. https://doi.org/10.1145/2642918.2647360

27. Nikhil Singh, Lucy Lu Wang, and Jonathan Bragg. 2024. FigurA11y: AI Assistance for Writing Scientific Alt Text. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*, 886–906. https://doi.org/10.1145/3640543.3645212

28. Tyler N. Starr, Allison J. Greaney, Sarah K. Hilton, Daniel Ellis, Katharine H. D. Crawford, Adam S. Dingens, Mary Jane Navarro, John E. Bowen, M. Alejandra Tortorici, Alexandra C. Walls, Neil P. King, David Veesler, and Jesse D. Bloom. 2020. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* 182, 5: 1295-1310.e20. https://doi.org/10.1016/j.cell.2020.08.012

29. Chase Stokes, Vidya Setlur, Bridget Cogley, Arvind Satyanarayan, and Marti A. Hearst. 2023. Striking a Balance: Reader Takeaways and Preferences when Integrating Text and Charts. *IEEE Transactions on Visualization and Computer Graphics* 29, 1: 1233–1243. https://doi.org/10.1109/TVCG.2022.3209383

30. Harini Suresh, Steven R. Gomez, Kevin K. Nam, and Arvind Satyanarayan. 2021. Beyond Expertise and Roles: A Framework to Characterize the Stakeholders of Interpretable Machine Learning and their Needs. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3411764.3445088

31. Harini Suresh, Kathleen M Lewis, John Guttag, and Arvind Satyanarayan. 2022. Intuitively Assessing ML Model Reliability through Example-Based Explanations and Editing Model Inputs. In *27th International Conference on Intelligent User Interfaces* (IUI '22), 767–781. https://doi.org/10.1145/3490099.3511160

32. Harini Suresh, Divya Shanmugam, Tiffany Chen, Annie G Bryan, Alexander D'Amour, John Guttag, and Arvind Satyanarayan. 2023. Kaleidoscope: Semantically-grounded, context-specific ML model evaluation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI '23), 1–13. https://doi.org/10.1145/3544548.3581482

31. Benny Tang, Angie Boggust, and Arvind Satyanarayan. 2023. VisText: A Benchmark for Semantically Rich Chart Captioning. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 7268–7298. Retrieved July 21, 2023 from https://aclanthology.org/2023.acl-long.401. **Outstanding Paper Award.**

34. John R Thompson, Jesse J Martinez, Alper Sarikaya, Edward Cutrell, and Bongshin Lee. 2023. Chart Reader: Accessible Visualization Experiences Designed with Screen Reader Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3544548.3581186

35. Jacob VanderPlas, Brian Granger, Jeffrey Heer, Dominik Moritz, Kanit Wongsuphasawat, Arvind Satyanarayan, Eitan Lees, Ilia Timofeev, Ben Welsh, and Scott Sievert. 2018. Altair: Interactive Statistical Visualizations for Python. *Journal of Open Source Software* 3, 32: 1057. https://doi.org/10.21105/joss.01057

34. Dylan Wootton, Amy Fox, Evan Peck, and Arvind Satyanarayan. 2025. Charting EDA: Characterizing Interactive Visualization Use in Computational Notebooks with a Mixed-Methods Formalism. To appear in *IEEE Transactions on Visualization and Computer Graphics* 31, 1.

37. Yifan Wu, Remco Chang, Joseph M. Hellerstein, Arvind Satyanarayan, and Eugene Wu. 2022. DIEL: Interactive Visualization Beyond the Here and Now. *IEEE Transactions on Visualization and Computer Graphics* 28, 1: 737–746. https://doi.org/10.1109/TVCG.2021.3114796

38. Yifan Wu, Joseph M. Hellerstein, and Arvind Satyanarayan. 2020. B2: Bridging Code and Interactive Visualization in Computational Notebooks. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 152–165. https://doi.org/10.1145/3379337.3415851

39. Guang-Yao Zhao, José L. Gómez, Antonio Fuentes, Thomas P. Krichbaum, Efthalia Traianou, Rocco Lico, Ilje Cho, Eduardo Ros, S. Komossa, Kazunori Akiyama, Keiichi Asada, Lindy Blackburn, Silke Britzen, Gabriele Bruni, Geoffrey B. Crew, Rohan Dahale, Lankeswar Dey, Roman Gold, Achamveedu Gopakumar, Sara Issaoun, Michael Janssen, Svetlana Jorstad, Jae-Young Kim, Jun Yi Koay, Yuri Y. Kovalev, Shoko Koyama, Andrei P. Lobanov, Laurent Loinard, Ru-Sen Lu, Sera Markoff, Alan P. Marscher, Iván Martí-Vidal, Yosuke Mizuno, Jongho Park, Tuomas Savolainen, and Teresa Toscano. 2022. Unraveling the Innermost Jet Structure of OJ 287 with the First GMVA + ALMA Observations. *The Astrophysical Journal* 932, 1: 72. https://doi.org/10.3847/1538-4357/ac6b9c

40. Hanxiu "Hazel" Zhu, Shelly Shiying Cheng, and Eugene Wu. 2022. How Do Captions Affect Visualization Reading? https://doi.org/10.48550/arXiv.2205.01263

41. Jonathan Zong, Dhiraj Barnwal, Rupayan Neogy, and Arvind Satyanarayan. 2021. Lyra 2: Designing Interactive Visualizations by Demonstration. *IEEE Transactions on Visualization and Computer Graphics* 27, 2: 304–314. https://doi.org/10.1109/TVCG.2020.3030367

42. Jonathan Zong, Crystal Lee, Alan Lundgard, JiWoong Jang, Daniel Hajas, and Arvind Satyanarayan. 2022. Rich Screen Reader Experiences for Accessible Data Visualization. *Computer Graphics Forum* 41, 3: 15–27. https://doi.org/10.1111/cgf.14519. **Best Paper Honorable Mention.**

43. Jonathan Zong, Josh Pollock, Dylan Wootton, and Arvind Satyanarayan. 2023. Animated Vega-Lite: Unifying Animation with a Grammar of Interactive Graphics. *IEEE Transactions on Visualization and Computer Graphics* 29, 1: 149–159. https://doi.org/10.1109/TVCG.2022.3209369