Computational Approaches for Understanding, Generating, and Adapting User Interfaces

YUE JIANG*, Max Planck Institute for Informatics, Germany
YUWEN LU*, University of Notre Dame, United States
JEFFREY NICHOLS, Apple Inc, United States
WOLFGANG STUERZLINGER, Simon Fraser University, Canada
CHUN YU, Tsinghua University, China
CHRISTOF LUTTEROTH, University of Bath, United Kingdom
YANG LI, Google Research, United States
RANJITHA KUMAR, University of Illinois Urbana-Champaign, United States
TOBY JIA-JUN LI, University of Notre Dame, United States

Computational approaches for user interfaces have been used in adapting interfaces for different modalities, usage scenarios and device form factors, understanding screen semantics for accessibility, task-automation, information extraction, and in assisting interface design. Recent advances in machine learning (ML) have drawn considerable attention across HCI and related fields such as computer vision and natural language processing, leading to new ML-based user interface approaches. Similarly, significant progress has been made with more traditional optimization- and planning-based approaches to accommodate the need for adapting UIs for screens with different sizes, orientations and aspect ratios, and in emerging domains such as VR/AR and 3D interfaces. The proposed workshop seeks to bring together researchers interested in all kinds of computational approaches for user interfaces across different sectors as a community, including those who develop algorithms and models and those who build applications, to discuss common issues including the need for resources, opportunities for new applications, design implications for human-AI interaction in this domain, and practical challenges such as user privacy.

CCS Concepts: • **Human-centered computing → User interface toolkits**.

Additional Key Words and Phrases: user interfaces, adative interfaces, interface semantics, interface generation, design mining

ACM Reference Format:
Yue Jiang, Yuwen Lu, Jeffrey Nichols, Wolfgang Stuerzlinger, Chun Yu, Christof Lutteroth, Yang Li, Ranjitha Kumar, and Toby Jia-Jun Li. 2022. Computational Approaches for Understanding, Generating, and Adapting User Interfaces. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22 Extended Abstracts), April 29-May 5, 2022, New Orleans, LA, USA*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3491101.3504030

*Contributed equally.

1 OVERVIEW

The use of computational approaches for understanding, generating, and adapting user interfaces has been an important long-time research topic in the CHI community [3, 33, 35]. In recent years, the need for adapting interfaces to support different modalities, usage scenarios, and form factors of devices [7, 12, 31, 32], the goal of understanding screen semantics for accessibility purposes [23, 36], the use of graphical interfaces as a resource for understanding their underlying tasks [19, 21, 22], the computational generative approach for assisting the design of user interfaces [9, 14], and the popularity of intelligent interface agents that understand and manipulate user interfaces based on users' behaviors [16, 29] have significantly expanded the horizon of this area, opening up new research opportunities while drawing significant attention and community interest to this topic.

Besides long-standing optimization-based (*e.g.*, [7, 11–13]) and decision-tree-based (*e.g.*, [26]) approaches that generate and adapt user interfaces to meet constraints, the availability of large user interface datasets such as RICO [4] also enabled a stream of work on data-driven learning-based approaches (*e.g.*, [19, 34, 36]). As a result, this research topic has become increasingly interdisciplinary with participation and contribution from researchers in ML, Computer Vision, and Natural Language Processing. Furthermore, with the improving capability in understanding interface semantics at scale and generating interface designs, we are also seeing increased interest in this topic from the broader CHI community.

In this one-day workshop, we seek to bring together researchers from different sub-disciplines of HCI, across the intersections of HCI and related fields (*e.g.*, ML, CV, NLP), at different stages of the pipeline from developing algorithms and models to developing applications, and across the industry-academia boundaries, to discuss the challenges and opportunities in computational approaches for understanding, generating, and adapting user interfaces. In particular, we plan to use the following guiding questions as sources of inspiration for initiating discussion:

- **Resources:** We have observed the impact of large data sets (*e.g.*, RICO [4]) on the advancement of research on this topic. What other resources can be useful for research in this domain? How should we approach pursuing such resources as a community?
- **Applications:** What new or improved applications can be enabled with the advances in computational approaches in user interfaces? What methods are useful for design, prototype, evaluate, and deploy such applications?
- **Role of Human:** What roles should human users play in applications that use computational interface approaches? How can we adapt existing principles, guidelines, and heuristics in human-AI interaction in this domain?
- **Practical Challenges:** Many practical challenges in *e.g.*, user privacy, legal implications, and data cleaning have been identified in past work in this area. What are some useful strategies to address these challenges?

2 BACKGROUND

2.1 Optimization-based Approaches

A stream of research in this area is optimization-based and focuses on GUI adaptation and customization to improve the user experience for different screen sizes and user requirements. To meet screen-size constraints, optimization-based systems automatically generate user interfaces by applying alternative widgets or alternative groupings (*e.g.*, SUPPLE [7]), optimizing layout parameters (*e.g.*, Arnauld [6]), or introducing disjunctive constraints (*e.g.*, ORC Layout [11]). These systems were also used to enable personalization to adapt to the needs of users and customize interfaces for people with disabilities [8]. With efficient layout solvers (*e.g.*, Cassowary [1], ORCSolver [13]), these approaches can generate adaptive user interfaces in real time at the time of interaction.

Generating adaptive interfaces is a long-standing problem. It is still challenging due to the increasing diversity of both devices and user needs. It is time-consuming and tedious for GUI designers to design different layouts for different devices potentially with different orientations. Thus, it would be useful to have efficient ways to design adaptive GUIs being able to apply to screens with different sizes, orientations, and aspect ratios. On the other hand, GUI users prefer to have personalized layout and content on their own devices. To better understand and customize user interfaces, reverse engineering has been used as a way to perform GUI customization. Prior reverse engineering approaches can identify interface elements and allow the user to add interactive enhancements to a GUI (*e.g.*, Façades [30] and Prefab [5]), infer constraints to describe a layout from UI exemplars (*e.g.*, InferUI [2]), recover higher-level constraint-based specifications [25], enable users to engage with web applications via different input modalities (*e.g.*, Genie [31]), generate potential constraint-based layout alternatives (*e.g.*, Scout [32]), and enable platform- and framework-independent customization for both GUI layouts and web layouts (*e.g.*, ReverseORC [12]). An important objective on this research topic that we wish to discuss in this work is to better understand both the layout structure and the design process of GUI designers, and further develop efficient approaches to design adaptive user interfaces.

2.2 Data-driven learning Approaches

Recently, the data-driven approach has drawn significant attention in the academic community and the industry to understand, generate, and adapt various types of user interfaces. This new wave of research was enabled by the availability of large GUI datasets like Rico [4], which includes the visual, textual, structural, and interactive design properties of more than 66k unique UI screens. The data-driven approach has been used in downstream tasks such as similarity-based GUI retrieval [4], sketch-based GUI retrieval (*e.g.*, Swire [10]), learning design semantics [24], generate accessibility metadata [36] and creating the semantics representation of GUIs (*e.g.*, Screen2Vec [19]). Recent document layout generation research also used deep learning approaches to generate documents and without manually defining constraints and templates (*e.g.*, LayoutGAN [15], Neural Design Network [14]). In addition, leveraging advances in natural language processing technologies, recent works introduced computational multi-modal interfaces enabling users to add functionalities that support voice input (*e.g.*, Geno [28]), specify data descriptions verbally (*e.g.*, APPINITE [17]), and resolve ambiguities with conversations and demonstrations (*e.g.*, PUMICE [20]).

The above line of work not only makes significant research contributions to modelling techniques for GUIs, but it also holds promise for designing and developing future data-driven applications with the new computational methods of GUI modelling and GUI semantics representation. For example, more effective GUI retrieval methods can enable designers to easily search for relevant existing app design interfaces as references for their own work [24]. Also, GUIs encapsulate rich underlying information of apps beyond graphical elements, *e.g.*, user task structures, constraints, processes, user preferences, and knowledge about app domains [18, 19]. Learning design semantics and generating accessibility metadata can help designers better understand interfaces from a new, non-graphic perspective. Auto-generation of interface layouts reduces the effort required to create multiple alternative interface layouts. This can free designers to focus on more intellectually demanding tasks and enable them to more easily develop parallel designs for testing [27]. Our workshop would be a great opportunity for researchers working on creating new GUI modelling techniques and those who build applications that benefit from these techniques to collaboratively discuss the potential application opportunities of GUI modelling techniques, and in this process, how to effectively utilize resources and what are potential risks to consider.

2.3 Combining the Two Approaches

While data-driven approaches have drawn significant attention in recent years and showed important advantages in many applications compared to optimization-based approaches, data-driven approaches tend to only work well in areas where high-quality domain-specific data is available. Besides, designing GUIs is a step-by-step process. To create better design prototypes, it is important to provide designers with fine-granularity control for their design while providing choices that GUI designers really need to make. Optimization-based approaches can give designers more control over their design, while data-driven approaches are better at generating different final results and suggestions. Our workshop would encourage researchers to think deeply about the strengths and weaknesses of both types of approaches and discuss the potential combination of both.

3 THE GOAL OF THE WORKSHOP

The main goal of this workshop is to build a community for those who are interested in developing and using computational approaches for understanding, generating, and adapting GUIs. By connecting those who work on optimization-based and data-driven approaches, those who work on algorithms and applications, those who work only on GUIs and those who work on interfaces that use different or multiple modalities, those in the CHI community and those from adjacent academic communities, academic researchers and industry practitioners, we aim to create a research agenda that maximizes the academic, practical, and social impact our community is able to make.

Through the discussion at the workshop, we hope to address the challenges lying in front of us as a community, *e.g.*, what resources can we benefit most from and potential risks to users' privacy and intellectual property. Through discussions, we expect to bring more attention to relevant work happening at the intersection of HCI and other adjacent disciplines, and outside the CHI community. We hope this workshop can serve as a platform for us to continue growing as a community, as well as attracting more CHI community members that are interested in participating in this area.

4 ORGANIZERS

Yue Jiang is a Ph.D. student in the Department of Visual Computing and Artificial Intelligence supervised by Prof. Christian Theobalt at the Max Planck Institute for Informatics, Germany. Her main research interests are in HCI and Graphics with a focus on adaptive user interfaces and 3D human performance capture. Her recent work with Prof. Wolfgang Stuerzlinger and Prof. Christof Lutteroth focuses on adaptive GUI layout based on OR-Constraints (ORC).

Yuwen Lu is a Ph.D. student in the Department of Computer Science and Engineering at the University of Notre Dame, working on using data-driven approaches for understanding and generating user interfaces to support UX research and design work. Prior to joining Notre Dame, Yuwen received a Master's degree in Human-Computer Interaction from Carnegie Mellon University.

Jeffery Nichols is a Research Scientist in the AI/ML group at Apple working on intelligent user interfaces. Previously he was a Staff Research Scientist at Google working on the open-source Fuchsia operating system. His most important academic contribution recently was the creation of the RICO dataset [4]. He also worked on the PUC project [26], whose primary focus was creating a specification language that can define any device and an automatic user interface generator that can create control panels from this specification language.

Wolfgang Stuerzlinger is a Professor at the School of Interactive Arts + Technology at Simon Fraser University. His work aims to gain a deeper understanding of and to find innovative solutions for real-world problems. Current research projects include better 3D interaction techniques for Virtual and Augmented Reality applications, new human-in-the-loop systems for big data analysis, the characterization of the effects of technology limitations on human performance, investigations of human behaviors with occasionally failing technologies, user interfaces for versions, scenarios, and alternatives, and new Virtual/Augmented Reality hardware and software.

Chun Yu is an Associate Professor at Tsinghua University. He is keen to research computational models and AI algorithms that facilitate the interaction between human and computers. Current research directions include novel sensing and interaction techniques, accessibility and user interface modeling. His research outcome has been integrated into commercial products serving hundreds of millions of users on smart phones, such as a touch sensing algorithm, a software keyboard decoding algorithm, a smart keyboard, and a screen reader for visually impaired people.

Christof Lutteroth is a Reader in the Department of Computer Science at the University of Bath. His main research interests are in HCI with a focus on immersive technology, interaction methods, and user interface design. In particular, he has a long-standing interest in methods for user interface layout. He is the director of the REal and Virtual Environments Augmentation Labs (REVEAL), the HCI research centre at the University of Bath.

Yang Li is a Senior Staff Research Scientist at Google, and an affiliate faculty member at the University of Washington CSE, focusing on the area intersecting AI and HCI. He pioneered on-device interactive ML on Android by developing impactful product features such as next app prediction and Gesture Search. Yang has extensively published in top venues across both the HCI and ML fields, including CHI, UIST, ICML, ACL, EMNLP, CVPR, NeurIPS (NIPS), ICLR, and KDD, and has constantly served as area chairs or senior area (track) chairs across the fields. Yang is also an editor of the upcoming Springer book on "AI for HCI: A Modern Approach", which is the first thorough treatment of the topic.

Ranjitha Kumar is an Associate Professor of Computer Science at the University of Illinois at Urbana-Champaign (UIUC) and the Chief Research Scientist at UserTesting. At UIUC, she runs the Data Driven Design Group, where she and her students leverage data mining and machine learning to address the central challenge of creating good user experiences: tying design decisions to desired outcomes. Her research has won best paper awards/nominations at premier conferences in HCI, and is supported by grants from the NSF, Google, Amazon, and Adobe. She received her BS and PhD from the Computer Science Department at Stanford University, and co-founded Apropose, Inc., a data-driven design startup based on her dissertation work that was backed by Andreessen Horowitz and New Enterprise Associates.

Toby Jia-Jun Li is an Assistant Professor in the Department of Computer Science and Engineering at the University of Notre Dame and the Director of the SaNDwich Lab. Toby and his group use human-centered methods to design, build, and study human-AI collaborative systems. In the domain of this workshop, Toby has recently done work in building interactive task learning agents that learn from the user's demonstrations on GUIs and natural language instructions about GUIs [16, 20], graphs models for representing and grounding natural language instructions about GUIs [17], and semantic embedding techniques for modeling GUIs [19].

5 PRE-WORKSHOP PLANS

5.1 Before the Workshop

We will distribute the Call for Participation through HCI-related emailing lists. We will also advertise the workshop at upcoming HCI conferences, among key research groups, on social media, and through our professional networks.

To help candidates get familiar with the scope and the goals of the workshop, we will create a website to provide information about the workshop, as described in Section 6.1.

6 WORKSHOP STRUCTURE

The workshop will be a one-day workshop containing approximately 30 participants and the organizers. The workshop will consist of keynote talks, workshop paper presentations, and topic discussions. Each participant will contribute to the workshop with a position paper (4–6 pages in CHI EA format) including the participant's prior research, future plans, insights, or interests in the area. The submissions will be reviewed by the workshop organizers and a program committee. The selection of participants will be based on the quality and novelty of and the insights presented in their submissions. We will also try to balance different workshop topics and criteria to ensure a diverse group of workshop participants.

6.1 Workshop Format and Asynchronous Engagement

We expect the workshop to be *hybrid* with the majority of participants attending in person. Synchronous remote participation will be available for those who are unable to join us in person. Specifically, all talks will be streamed at the workshop and virtual discussion "breakout rooms" will be created for remote attendees. No special technical capacity will be necessary to support the workshop day beyond the typical equipment (*e.g.*, Wi-Fi, projector, microphone) at a conference center. The organizers can prepare *e.g.*, speakerphones and cameras to support remote participation if needed. Asynchronous engagement with workshop materials will also be supported on the workshop website.

A workshop website will be released upon the acceptance of the workshop. It will list the call of paper of the workshop, the program, the list of organizers and speakers, and the pre-prints of accepted workshop position papers. After the workshop, the website will also host recorded workshop talks (with consents from the authors) with discussion threads for each talk to support asynchronous engagement with workshop materials.

The plan for the workshop format and logistics is subject to change depending on the final plan for the CHI conference. If an in-person meeting of the workshop is not feasible due to the development of the pandemic situation or other unexpected circumstances, the workshop will be converted to a fully online event on Zoom. The workshop schedule will be adjusted to be two half-day meetings in order to accommodate attendees from different time zones.

During the workshop, participants will have the opportunity to interact with experts in the fields, and the workshop organizers will provide input into different areas that will be discussed during the workshop. We plan to invite two keynote speakers at the workshop. Each will present a talk in their areas of expertise for 30 minutes, followed by an extensive Q&A and discussion session.

Accepted workshop position papers will be sorted into sessions based on their topics, and each position paper will be presented for 5 minutes.

The participants will form small breakout groups for topic discussion. Several groups will be pre-created by the organizers on each of the key questions of the workshop (*e.g.*, resources, applications, the role of human). The participants may also propose and form new groups as needed. At the end of the workshop, each breakout group will report back their discussion results for the large group to discuss.

7 POST-WORKSHOP PLAN

After the CHI workshop, we plan to produce a report of the workshop outcome. The workshop papers and results will be available on the workshop website prior to and after the workshop, providing opportunities for a larger audience to get familiar with this area. We may seek opportunities for an edited book or a special issue in a selected journal (*e.g.*, ToCHI) where the participants will be encouraged to publish their work.

Time	Session
9:00 - 9:30	Introduction of workshop organizers, participants, topics, and goals
9:30 - 10:30	Keynote 1 by an invited speaker
10:30 - 11:00	Coffee break
11:00 - 12:00	Paper Presentation (6 papers)
12:00 - 13:30	Lunch
13:30 - 14:30	Keynote 2 by an invited speaker
14:30 - 15:30	Paper Presentation (6 papers)
15:30 - 16:00	Coffee break
16:00 - 17:00	Group discussion
17:00 - 17:30	Discussion group report back, wrap-up
17:30	Dinner (optional)

Table 1. Tentative agenda of the workshop

A central goal of this workshop is community building for researchers and practitioners in this area. After the workshop, we plan to create a platform for community members to continue the discussion and share resources. Potential options may include a periodical email newsletter, a public GitHub repository, or a Slack/Discord channel. Participants and organizers will discuss the next steps at the workshop.

8 CALL FOR PARTICIPATION

"Computational Approaches for Understanding, Generating, and Adapting User Interfaces" is a workshop at CHI2022. In this one-day workshop, we seek to bring together researchers from different sub-disciplines of HCI, across the intersections between HCI and adjacent fields (*e.g.*, ML, CV, NLP), at different stages of the pipeline from developing algorithms and models to developing applications, and across the industry and academia boundaries to discuss the challenges and opportunities in computational approaches for understanding, generating, and adapting user interfaces.

Researchers and practitioners are invited to submit a 4–6 page (excluding references) position paper in the CHI extended abstract format to participate in the workshop. Submissions can reflect on past work, in-progress projects, present challenges and approaches, identified opportunities, or critical opinions and arguments covering but not limited to the following topics:

- New models, algorithms, and techniques for understanding, generating, and adapting user interfaces.
- Applications, methodology, and theories motivated by the advances in computational approaches for understanding, generating, and adapting user interfaces.
- Practical challenges (*e.g.*, privacy) and resources (*e.g.*, dataset) relevant to the above topics.

The position papers should follow the instruction on the website https://sites.google.com/nd.edu/computational-uichi2 and be submitted via user.interface.workshop@gmail.com. Each submission will be reviewed by the workshop organizers and program committee members and selection will be based on the paper's quality, novelty, and fit of the topic, while aiming for a balance of different perspectives. Accepted papers will be optionally made available at the workshop website (with consent from the authors). At least one author of each accepted position paper must register and attend the workshop and register for at least one day of the conference. The workshop will be in a hybrid structure. We will broadcast the workshop live for remote participants and make the recordings available on the website after the workshop. The authors of each accepted position paper will have about 8 minutes for a live or pre-recorded presentation of their work and an additional 2-minute Q&A.

8.1 Key Dates

- Call for participation released: December 16, 2021
- Position paper submission deadline: February 24, 2022
- Notification of acceptance: March 15, 2022
- Workshop date: April 30 (Saturday) or May 1 (Sunday), 2022

REFERENCES

[1] Greg J. Badros, Alan Borning, and Peter J. Stuckey. 2001. The Cassowary Linear Arithmetic Constraint Solving Algorithm. *ACM Trans. Comput.-Hum. Interact* 8, 4 (2001), 267–306. https://doi.org/10.1145/504704.504705

[2] Pavol Bielik, Marc Fischer, and Martin Vechev. 2018. Robust Relational Layout Synthesis from Examples for Android. *Proc. ACM Program. Lang.* 2, OOPSLA, Article 156 (Oct. 2018), 29 pages. https://doi.org/10.1145/3276526

[3] Alan Borning and Robert Duisberg. 1986. Constraint-Based Tools for Building User Interfaces. *ACM Trans. Graph.* 5, 4 (Oct. 1986), 345–374. https://doi.org/10.1145/27623.29354

[4] Biplab Deka, Zifeng Huang, Chad Franzen, Joshua Hibschman, Daniel Afergan, Yang Li, Jeffrey Nichols, and Ranjitha Kumar. 2017. Rico: A Mobile App Dataset for Building Data-Driven Design Applications. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. Association for Computing Machinery, New York, NY, USA, 845–854. https://doi.org/10.1145/3126594.3126651

[5] Morgan Dixon and James Fogarty. 2010. Prefab: Implementing Advanced Behaviors Using Pixel-Based Reverse Engineering of Interface Structure. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1525–1534. https://doi.org/10.1145/1753326.1753554

[6] Krzysztof Gajos and Daniel Weld. 2005. Preference Elicitation for Interface Optimization. *UIST: Proceedings of the Annual ACM Symposium on User Interface Softaware and Technology*, 173–182. https://doi.org/10.1145/1095034.1095063

[7] Krzysztof Z. Gajos, Daniel S. Weld, and Jacob O. Wobbrock. 2010. Automatically Generating Personalized User Interfaces With Supple, In Proceedings of the 9th International Conference on Intelligent User Interfaces. *Artif. Intell* 174, 12-13, 910–950. https://doi.org/10.1016/j.artint.2010.05.005

[8] Krzysztof Z. Gajos, Jacob O. Wobbrock, and Daniel S. Weld. 2008. Improving the Performance of Motor-Impaired Users With Automatically-Generated, Ability-Based Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. ACM, 1257–1266. https://doi.org/10.1145/1357054.1357250

[9] Forrest Huang, John F. Canny, and Jeffrey Nichols. 2019. Swire: Sketch-Based User Interface Retrieval. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3290605.3300334

[10] Forrest Huang, John F. Canny, and Jeffrey Nichols. 2019. Swire: Sketch-Based User Interface Retrieval. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3290605.3300334

[11] Yue Jiang, Ruofei Du, Christof Lutteroth, and Wolfgang Stuerzlinger. 2019. ORC Layout: Adaptive GUI Layout with OR-Constraints. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, Article 413, 12 pages. https://doi.org/10.1145/3290605.3300643

[12] Yue Jiang, Wolfgang Stuerzlinger, and Christof Lutteroth. 2021. *ReverseORC: Reverse Engineering of Resizable User Interface Layouts with OR-Constraints*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445043

[13] Yue Jiang, Wolfgang Stuerzlinger, Matthias Zwicker, and Christof Lutteroth. 2020. ORCSolver: An Efficient Solver for Adaptive GUI Layout with OR-Constraints. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376610

[14] Hsin-Ying Lee, Lu Jiang, Irfan Essa, Phuong B Le, Haifeng Gong, Ming-Hsuan Yang, and Weilong Yang. 2019. Neural Design Network: Graphic Layout Generation with Constraints. *arXiv e-prints* (2019), arXiv–1912.

[15] Jianan Li, Jimei Yang, Aaron Hertzmann, Jianming Zhang, and Tingfa Xu. 2019. Layoutgan: Generating graphic layouts with wireframe discriminators. *arXiv preprint arXiv:1901.06767* (2019).

[16] Toby Jia-Jun Li, Amos Azaria, and Brad A. Myers. 2017. SUGILITE: Creating Multimodal Smartphone Automation by Demonstration. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6038–6049. https://doi.org/10.1145/3025453.3025483

[17] Toby Jia-Jun Li, Igor Labutov, Xiaohan Nancy Li, Xiaoyi Zhang, Wenze Shi, Wanling Ding, Tom M. Mitchell, and Brad A. Myers. 2018. APPINITE: A Multi-Modal Interface for Specifying Data Descriptions in Programming by Demonstration Using Natural Language Instructions. In *2018 IEEE*

Symposium on Visual Languages and Human-Centric Computing (VL/HCC). 105–114. https://doi.org/10.1109/VLHCC.2018.8506506

[18] Toby Jia-Jun Li, Tom Mitchell, and Brad Myers. 2020. Interactive Task Learning from GUI-Grounded Natural Language Instructions and Demonstrations. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Association for Computational Linguistics, Online, 215–223. https://doi.org/10.18653/v1/2020.acl-demos.25

[19] Toby Jia-Jun Li, Lindsay Popowski, Tom Mitchell, and Brad A Myers. 2021. Screen2Vec: Semantic Embedding of GUI Screens and GUI Components. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 578, 15 pages. https://doi.org/10.1145/3411764.3445049

[20] Toby Jia-Jun Li, Marissa Radensky, Justin Jia, Kirielle Singarajah, Tom M. Mitchell, and Brad A. Myers. 2019. PUMICE: A Multi-Modal Agent That Learns Concepts and Conditionals from Natural Language and Demonstrations. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 577–589. https://doi.org/10.1145/3332165.3347899

[21] Toby Jia-Jun Li and Oriana Riva. 2018. KITE: Building conversational bots from mobile apps. In *Proceedings of the 16th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys 2018)*. ACM.

[22] Yang Li, Jiacong He, Xin Zhou, Yuan Zhang, and Jason Baldridge. 2020. Mapping Natural Language Instructions to Mobile UI Action Sequences. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. ACL, Online, 8198–8210. https://doi.org/10.18653/v1/2020.acl-main.729

[23] Yang Li, Gang Li, Luheng He, Jingjie Zheng, Hong Li, and Zhiwei Guan. 2020. Widget Captioning: Generating Natural Language Description for Mobile User Interface Elements. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. ACL, Online, 5495–5510. https://doi.org/10.18653/v1/2020.emnlp-main.443

[24] Thomas F. Liu, Mark Craft, Jason Situ, Ersin Yumer, Radomir Mech, and Ranjitha Kumar. 2018. Learning Design Semantics for Mobile Apps. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 569–579. https://doi.org/10.1145/3242587.3242650

[25] Christof Lutteroth. 2008. Automated Reverse Engineering of Hard-Coded GUI Layouts. In *Proceedings of the Ninth Conference on Australasian User Interface - Volume 76* (Wollongong, Australia) *(AUIC '08)*. Australian Computer Society, Inc., AUS, 65–73. https://doi.org/10.5555/1378337.1378350

[26] Jeffrey Nichols, Brad A. Myers, Michael Higgins, Joseph Hughes, Thomas K. Harris, Roni Rosenfeld, and Mathilde Pignol. 2002. Generating Remote Control Interfaces for Complex Appliances. In *Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology* (Paris, France) *(UIST '02)*. Association for Computing Machinery, New York, NY, USA, 161–170. https://doi.org/10.1145/571985.572008

[27] Jakob Nielsen and Jan Maurits Faber. 1996. Improving system usability through parallel design. *Computer* 29, 2 (1996), 29–35.

[28] Ritam Jyoti Sarmah, Yunpeng Ding, Di Wang, Cheuk Yin Phipson Lee, Toby Jia-Jun Li, and Xiang 'Anthony' Chen. 2020. Geno: A Developer Tool for Authoring Multimodal Interaction on Existing Web Applications. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 1169–1181. https://doi.org/10.1145/3379337.3415848

[29] Alborz Rezazadeh Sereshkeh, Gary Leung, Krish Perumal, Caleb Phillips, Minfan Zhang, Afsaneh Fazly, and Iqbal Mohomed. 2020. VASTA: A Vision and Language-Assisted Smartphone Task Automation System. In *Proceedings of the 25th International Conference on Intelligent User Interfaces* (Cagliari, Italy) *(IUI '20)*. Association for Computing Machinery, New York, NY, USA, 22–32. https://doi.org/10.1145/3377325.3377515

[30] Wolfgang Stuerzlinger, Olivier Chapuis, Dusty Phillips, and Nicolas Roussel. 2006. User Interface Façades: Towards Fully Adaptable User Interfaces. *UIST '06: ACM Symposium on User Interface Software and Technology* (10 2006). https://doi.org/10.1145/1166253.1166301

[31] Amanda Swearngin, Amy J. Ko, and James Fogarty. 2017. Genie: Input Retargeting on the Web through Command Reverse Engineering. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 4703–4714. https://doi.org/10.1145/3025453.3025506

[32] Amanda Swearngin, Chenglong Wang, Alannah Oleson, James Fogarty, and Amy J. Ko. 2020. *Scout: Rapid Exploration of Interface Layout Alternatives through High-Level Design Constraints*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376593

[33] Pedro Szekely, Ping Luo, and Robert Neches. 1993. Beyond Interface Builders: Model-Based Interface Tools. In *Proceedings of the INTERCHI '93 Conference on Human Factors in Computing Systems* (Amsterdam, The Netherlands) *(INTERCHI '93)*. IOS Press, NLD, 383–390.

[34] Bryan Wang, Gang Li, Xin Zhou, Zhourong Chen, Tovi Grossman, and Yang Li. 2021. Screen2Words: Automatic Mobile UI Summarization with Multimodal Learning. (2021).

[35] Brad Vander Zanden and Brad A. Myers. 1990. Automatic, Look-and-Feel Independent Dialog Creation for Graphical User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '90)*. Association for Computing Machinery, New York, NY, USA, 27–34. https://doi.org/10.1145/97243.97248

[36] Xiaoyi Zhang, Lilian de Greef, Amanda Swearngin, Samuel White, Kyle Murray, Lisa Yu, Qi Shan, Jeffrey Nichols, Jason Wu, Chris Fleizach, Aaron Everitt, and Jeffrey P Bigham. 2021. *Screen Recognition: Creating Accessibility Metadata for Mobile Applications from Pixels*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445186