The Future of Computational Approaches for Understanding and Adapting User Interfaces

Yue Jiang
yue.jiang@aalto.fi
Aalto University
Finland

Yuwen Lu
ylu23@nd.edu
University of Notre Dame
Notre Dame, United States

Christof Lutteroth
c.lutteroth@bath.ac.uk
University of Bath
United Kingdom

Toby Jia-Jun Li
toby.j.li@nd.edu
University of Notre Dame
Notre Dame, United States

Jeffrey Nichols
jwnichols@apple.com
Apple Inc
United States

Wolfgang Stuerzlinger
w.s@sfu.ca
Simon Fraser University
Canada

ABSTRACT

Building on the success of the first workshop on understanding, generating, and adapting user interfaces at CHI2022, this workshop will advance this research area further by looking at existing results and exploring new research directions. Computational approaches for user interfaces have been used in adapting interfaces for different devices, modalities, and user preferences. Recent work has made significant progress in understanding and adapting user interfaces with traditional constraint/rule-based optimization and machine learning-based data-driven approaches; however, these two approaches remain separate. Combining the two approaches has great potential to advance the area but remains under-explored and challenging. Other contributions, such as datasets for potential applications, novel representations of user interfaces, the analysis of human traces, and models with multi-modalities, will also open up future research options. The proposed workshop seeks to bring together researchers interested in computational approaches for user interfaces to discuss the needs and opportunities for future user interface algorithms, models, and applications.

ACM Reference Format:
Yue Jiang, Yuwen Lu, Christof Lutteroth, Toby Jia-Jun Li, Jeffrey Nichols, and Wolfgang Stuerzlinger. 2023. The Future of Computational Approaches for Understanding and Adapting User Interfaces. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3544549.3573805

1 OVERVIEW

Computational approaches for understanding and adapting user interfaces have attracted significant attention in the HCI community [4, 36, 38]. Recent approaches have explored the generation of user interfaces for assisting the process of designing them [10, 15, 19, 35], the adaptation to different devices and user preferences [8, 9, 11], methods to understand underlying tasks and user behaviors [20, 24, 25, 32, 34], reverse engineering for accessibility and user interface understanding [6, 13, 26, 39], and novel representations of user interfaces [11, 22].

Optimization-based approaches generate and adapt user interfaces to satisfy constraints [8, 11, 13, 14, 30]; data-driven approaches employ deep learning networks to improve the capability in understanding interface semantics at scale and generating interface designs [22, 37, 39]). As both these directions progress, the research topic of user interface design has become increasingly interdisciplinary, with participation and contribution from researchers working on optimization algorithms, data-driven models, software engineering, accessibility, and various user interface applications. We are seeing increased interest in this topic from the broader community.

In the previous workshop at CHI2022 [12], we explored optimization-based approaches and data-driven approaches separately and focused on user interface applications. In this workshop, we aim to focus on closing the gap between these two streams of approaches and fulfill the fundamental needs for future user interface research.

- **Combination:** It is crucial to provide designers with fine-granular control for their design while providing choices that UI designers need to make. Optimization-based approaches can give designers more control over their design; data-driven approaches are better at generating different final results and suggestions. Thus, analyzing the strengths and weaknesses of both kinds of approaches and exploring their combination shows significant potential to advance the topic.
- **Datasets:** Data-driven approaches depend highly on datasets. Large datasets, such as RICO [5], have triggered a stream of research on this topic. However, the RICO dataset has limitations and researchers have consequently worked on cleaning up or augmenting it to achieve better model performance [16, 17]. In addition, the shifting trends in the design industry have resulted in major interface updates. As a result, new meaningful datasets can form the basis of future research. Given such datasets, foundation models can also bring significant opportunities to build AI-infused applications.

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3573805

- **Representations:** Powerful representations of user interfaces are essential for both understanding existing layouts and generating new ones. UIBERT [2] and Screen2Vec [22] demonstrate the promise of UI representations. ORC Layout [11] represents GUI layouts with an extended constraint system with OR-constraints. Novel representations often have the potential to extend the scope of the applications and surpass the limitations of older representations.
- **Human Traces:** Most existing works focused on understanding static user interfaces. It is also essential to analyze the dynamic human interaction traces with interfaces, towards understanding user intentions.
- **Domain Extension:** Most 2D interface approaches can be directly applied to 3D. If we consume UIs in mixed reality, how we generate user interfaces will need to change drastically. It is still challenging to understand the connection between physical objects and virtual interfaces, and to further optimize the virtual interfaces to adapt to user preferences and cognitive load.

2 BACKGROUND

2.1 Optimization-based and Data-driven learning Approaches

Generating adaptive interfaces remains challenging due to the increasing diversity of devices and user preferences. Optimization-based approaches focus on UI adaptation and customization to improve the user experience for different devices and user requirements. They automatically generate optimized user interfaces that satisfy device properties and layout element constraints, while minimizing user effort (*e.g.*, SUPPLE [8], Arnauld [7], ORC Layout [11]). Efficient Layout solvers, such as Cassowary [1] and ORCSolver [14], enable real-time adaptive UI generation. Reverse engineering approaches are a key component to understand existing user interfaces and to customize them. Prior research identified interface elements and enabled users to add advanced interactions [6, 33], inferred constraints from input UI exemplars [3], reconstructed higher-level constraint-based specifications [28], enabled webpage adaptation via different input modalities [34], and enabled platform- and framework-independent customization for both graphical UIs and webpages [13]. Data-driven approaches have also been used to understand UIs for accessibility [39], perform UI retrieval [5, 10], learn design semantics [27], generate semantic representations from UIs [22], create documents without manually defining constraints and templates (, *e.g.*, LayoutGAN [18], Neural Design Network [15]), and for combining programming-by-demonstration and natural language processing technologies to understand and author multimodal UIs [21, 23, 31]).

Designing UIs is an iterative process; thus, designers would like to have fine-granularity control for their design given constructive suggestions. Optimization-based approaches can give designers more control over their design; data-driven approaches are better at generating different final results and suggestions. Our workshop will encourage researchers to think deeply about the strengths and weaknesses of both types of approaches and discuss the potential combination of both.

2.2 Datasets and Representations

A crucial limitation of data-driven approaches is that they only work well in areas where high-quality domain-specific data is available. The current stream of research on data-driven approaches was motivated mainly by the availability of the large GUI dataset RICO [5]. It includes more than 72k unique UI screens and triggered data-driven applications for UI design retrieval, layout generation, code generation for UIs, user interaction modeling, and user perception prediction. A few more datasets have been released after RICO. Most recent datasets are extensions or cleaned versions of RICO [2, 16]. Moran et al. published a novel dataset called CLARITY [29] containing functional UI descriptions. Additional effort is still required to provide researchers with new datasets to extend the domain of this area.

An appropriate, powerful representation of a user interface is essential for understanding existing and generating new layouts. For example, UIBERT [2] learned generic feature representations for UI components. Screen2Vec [22] created semantic representation of UIs. ORC Layout [11] represents UIs as an extended constraint system with OR-constraints. Novel representations have the potential to extend the scope of future research.

3 THE GOAL OF THE WORKSHOP

The main goal of this workshop is to encourage the community to think about potential opportunities for research on user interfaces. We aim to create a research agenda that maximizes the academic, practical, and social impact our community can make. We will encourage people from the CHI community, from adjacent academic communities, and industry practitioners to participate.

Through the discussion at the workshop, we expect to bring more attention to relevant work at the intersection of HCI, other adjacent disciplines (e.g. Machine Learning and Software Engineering), and industry practitioners. We hope this workshop can serve as a platform to continue growing as a community and attract more members interested to participate.

4 ORGANIZERS

Yue Jiang is a Ph.D. student in Intelligent Systems by Prof. Antti Oulasvirta at Aalto University, Finland. Her main research interests are in adaptive user interfaces and eye tracking. Her recent work with Prof. Wolfgang Stuerzlinger and Prof. Christof Lutteroth focuses on adaptive GUI layouts 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. Before joining Notre Dame, Yuwen received a Master's degree in Human-Computer Interaction from Carnegie Mellon University.

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 center at the University of Bath.

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 [20, 23], graphs models for representing and grounding natural language instructions about GUIs [21], and semantic embedding techniques for modeling GUIs [22].

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 [5]. He also worked on the PUC project [30], 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.

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 research groups, on social media, and through our professional networks.

To help candidates get familiar with the workshop's scope and goals, we will create a website to provide information about the workshop.

6 WORKSHOP STRUCTURE

The workshop will be a one-day workshop containing approximately 30 participants and the organizers. It will consist of two 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). The workshop organizers and a program committee will review the submissions. We will select participants based on the quality and novelty of 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.

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

6.1 Workshop Format and Asynchronous Engagement

We expect the workshop to be *hybrid* with most participants attending in person. Synchronous remote participation will be available for those unable to join us in person. We will stream all talks at the workshop and will create virtual discussion "breakout rooms" for remote attendees. No special technical capacity will be necessary to support the workshop day beyond the standard equipment (*e.g.*, WiFi, projector, microphone) at the conference center. The organizers can prepare *e.g.*, speakerphones and cameras, to support remote participation if needed. We will also support synchronous engagement with workshop materials on the workshop website.

We will release the workshop website upon the acceptance of the workshop. It will list the call for papers of the workshop, the program information, 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 author consent) with discussion threads for each talk to support asynchronous engagement with workshop materials. The workshop format and logistics plan are subject to change depending on the final plans for the CHI conference itself.

During the workshop, participants will have the opportunity to interact with experts in the fields. The workshop organizers will provide input into different areas to discuss during the workshop. We plan to invite two keynote speakers to the workshop. Each will present a talk in their areas of expertise for 30 minutes, followed by an extensive Q&A and discussion session. We will sort accepted workshop position papers 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-defined by the organizers to address each of the critical questions of the workshop (*e.g.*, the combination of optimization-based and data-driven approaches, datasets, and representations). The participants may also propose and form new groups as needed. At the end of the workshop, each

breakout group will report their discussion results for the larger group to discuss.

7 POST-WORKSHOP PLAN

After the CHI workshop, we plan to produce a report on the workshop outcome. The workshop papers and results will be available on the website before 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.

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

"The Future of Computational Approaches for Understanding and Adapting User Interfaces" is a workshop at CHI2023. 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, SE) at different stages of the pipeline from developing algorithms and models to developing applications, and across industry and academia boundaries to discuss the opportunities and needs of future computational approaches for adaptive user interfaces.

We invite researchers and practitioners 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:

- **Combination:** It is crucial to provide designers with fine-granularity control for their design while providing choices that UI designers need to make. Optimization-based approaches can give designers more control over their design; data-driven approaches are better at generating different final results and suggestions. Thus, analyzing the strengths and weaknesses of both kinds of approaches and exploring their combination shows significant potential to advance the topic.
- **Datasets:** Data-driven approaches highly depend on datasets. Large datasets, such as the RICO dataset, have triggered a stream of research on this topic. New meaningful datasets can be the basis of future research. Given datasets, foundation models can also bring significant opportunities to build AI-infused applications.
- **Representations:** The representations of user interfaces are essential for understanding and generating layouts. Novel UI representations have the potential to extend the scope of applications and surpass the limitations of existing representations. For example, UIBERT and Screen2Vec show the promise of UI representations. Similarly, ORC Layout represents UI layouts as a constraint system with OR-constraints which enables new UI resize behaviors.
- **Human Traces:** Most existing works focused on understanding static user interfaces. To understand user intentions during UI usage, analyzing dynamic human UI usage traces is critical.
- **Domain Extension:** 2D interface approaches can of course be extended to 3D. Yet, if we consume UIs in mixed reality, how we generate user interfaces will change drastically. It is still challenging to understand how physical objects and virtual interfaces connect, and further how to optimize virtual interfaces to adapt to user preferences and cognitive load.

Participants should follow the instruction on the website https://sites.google.com/nd.edu/computational-uichi23 and submit the position papers via user.interface.workshop@gmail.com. Workshop organizers and program committee members will review submissions. We will select submissions based on quality, novelty, and topic fit while aiming for a balance of different perspectives. Accepted papers will optionally be available on the workshop website (with the author's consent). 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 use 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 followed by an additional 2-minute Q&A.

8.1 Estimated Key Dates

- Call for participation released: December 15, 2022
- Position paper submission deadline: February 23, 2023
- Notification of acceptance: March 15, 2023
- Workshop date: April 23 or April 28, 2023

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] Chongyang Bai, Xiaoxue Zang, Ying Xu, Srinivas Sunkara, Abhinav Rastogi, Jindong Chen, et al. 2021. Uibert: Learning generic multimodal representations for ui understanding. *arXiv preprint arXiv:2107.13731* (2021).

[3] 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

[4] 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

[5] 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

[6] 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

[7] 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

[8] Krzysztof Gajos and Daniel S. Weld. 2004. SUPPLE: Automatically Generating User Interfaces. In *Proceedings of the 9th International Conference on Intelligent*

User Interfaces (Funchal, Madeira, Portugal) *(IUI '04)*. Association for Computing Machinery, New York, NY, USA, 93–100. https://doi.org/10.1145/964442.964461

[9] 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

[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, 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*. 1–6.

[13] Yue Jiang, Wolfgang Stuerzlinger, and Christof Lutteroth. 2021. ReverseORC: Reverse Engineering of Resizable User Interface Layouts with OR-Constraints. 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 316, 18 pages. https://doi.org/10.1145/3411764.3445043

[14] 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

[15] Hsin-Ying Lee, Lu Jiang, Irfan Essa, Phuong B. Le, Haifeng Gong, Ming-Hsuan Yang, and Weilong Yang. 2020. Neural Design Network: Graphic Layout Generation with Constraints. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III* (Glasgow, United Kingdom). Springer-Verlag, Berlin, Heidelberg, 491–506. https://doi.org/10.1007/978-3-030-58580-8_29

[16] Luis A. Leiva, Asutosh Hota, and Antti Oulasvirta. 2020. Enrico: A High-quality Dataset for Topic Modeling of Mobile UI Designs. In *Proc. MobileHCI Adjunct.*

[17] Gang Li, Gilles Baechler, Manuel Tragut, and Yang Li. 2022. Learning to Denoise Raw Mobile UI Layouts for Improving Datasets at Scale. In *CHI Conference on Human Factors in Computing Systems*. 1–13.

[18] 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).

[19] Jianan Li, Jimei Yang, Jianming Zhang, Chang Liu, Christina Wang, and Tingfa Xu. 2021. Attribute-Conditioned Layout GAN for Automatic Graphic Design. *IEEE Transactions on Visualization and Computer Graphics* 27, 10 (oct 2021), 4039–4048. https://doi.org/10.1109/TVCG.2020.2999335

[20] 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

[21] 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

[22] 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

[23] 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

[24] 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.

[25] 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

[26] 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

[27] 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

[28] 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

[29] Kevin Moran, Ali Yachnes, George Purnell, Junayed Mahmud, Michele Tufano, Carlos Bernal Cardenas, Denys Poshyvanyk, and Zach H'Doubler. 2022. An Empirical Investigation into the Use of Image Captioning for Automated Software Documentation. In *2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*. 514–525. https://doi.org/10.1109/SANER53432.2022.00069

[30] 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

[31] 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

[32] 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

[33] 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

[34] 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

[35] 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

[36] 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.

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

[38] 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

[39] 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