ISSN 2186-7437

NII Shonan Meeting Report

No. 213

Augmented Multimodal Interaction for Synchronous Presentation, Collaboration, and Education with Remote Audiences

Matthew Brehmer
Maxime Cordeil
Christophe Hurter
Takayuki Itoh

June 24–27, 2024

National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-Ku, Tokyo, Japan

Augmented Multimodal Interaction for Synchronous Presentation, Collaboration, and Education with Remote Audiences

Organizers:
Matthew Brehmer (University of Waterloo, Canada)
Maxime Cordeil (Queensland University, Australia)
Christophe Hurter (ENAC / University of Toulouse, France)
Takayuki Itoh (Ochanomizu University, Japan)

June 24–27, 2024

Abstract

NII Shonan Meeting #213 in June 2024 was a four-day seminar dedicated to identifying emerging challenges in data-rich multimodal remote collaboration. We gathered 32 researchers interested on multimodal, synchronous, and remote or hybrid forms of communication and collaboration; 27 of these researchers attended the seminar in person, while the remaining five attended a subset of sessions remotely. This topic lies at the intersection of Data Visualization / Information Visualization (InfoVis), Human-Computer Interaction (HCI), and Computer-Supported Collaborative Work (CSCW), and overlaps thematically with several prior seminars and workshops, including several that were co-organized by the authors. The meeting included short presentations by each attendee, topic-based breakout sessions, and group discussion sessions. Beyond this report, one expected outcome of this seminar is a survey / position paper outlining the aforementioned challenges and establishing a shared research agenda. We identified 19 challenges across four categories: (1) Technology, Tools, and Techniques; (2) People; (3) Artificial Intelligence (AI); and (4) Evaluation.

Introduction and Background

Throughout the last two decades, remote synchronous communication tools have been in regular use across workplaces and educational contexts. However, it was the global COVID-19 pandemic and the shift to remote and hybrid work / education that made these tools the dominant means by which people communicate, collaborate, and teach at a distance. Today, it is common for people to meet using teleconference tools such asZoom [121], Cisco Webex Meetings [22], Slack Huddles [101], Google Meet [39], and others. Typically, these tools afford multimodal communication including multi-party video and audio conferencing, screen sharing, breakout rooms, polls, reactions, and side-channel text chat functionality. Often these tools are used in conjunction with collaborative productivity tools, and collaboration platforms organized by channels and threads, such as Slack [100] or Microsoft Teams [80], forming synchronous episodes within a larger timeline of asynchronous communication.

However, despite the multimodal nature of these communication platforms, the experience is often a poor substitute for co-located communication, particularly when presenting complex and/or dynamic multimedia content via screen-sharing [8]. When doing so, the presenter is often relegated to a secondary thumbnail video frame, and only they can interact with the shared content using mouse and keyboard controls. In contrast, consider co-located communication scenarios such as those in meeting rooms or lecture halls, where all participants can use their physical presence and body language to interact with and point to the multimedia content being discussed. In particular, embodied cognition research [77] suggests that nonverbal hand gestures are essential for comprehending complex or abstract content, such in mathematics education, [1] in engineering and design [13], and in business decision-making [23].

Some telecommunication tools have recently introduced ways to restore the missing embodied presence of a presenter as they share multimedia content such as slides, diagrams, data visualization, and interactive interfaces. For instance, Cisco Webex Meetings [22] and Microsoft Teams [80] offer functionality to segment the presenter's outline from their webcam video and composite them in front of screen-shared content. Virtual camera applications have also become popular, including mmhmm[81] and OBS Studio [84]; these tools allow for considerable flexibility with respect to video compositing, and are compatible with most teleconferencing applications. However, only a single presenter can interact with shared content, and they must do so using standard mouse and keyboard interfaces.

Meanwhile, advancements in extended reality (XR) suggest approaches to multimodal communication that bypass standard desktop environments.For instance, Flow Immersive [35] is an application that allows people to present complex 3D data visualization content in mobile augmented reality (AR) or within an immersive virtual reality (VR) environment. In general, VR meeting spaces are an emerging trend in enterprise settings, in which all participants join a meeting as an avatar in an immersive 3D conference room. While the potential of XR for remote multimodal communication is promising, it also exhibits several limitations. The first issue is the lack of general access to affordable and comfortable hardware devices, including depth sensors and head-mounted displays; moreover, many XR applications require multi-device coordination with hand-held pointing devices or simultaneous touch-screen interaction. A

second issue is a relatively higher amount of fatigue induced by XR applications incorporating head-mounted displays. A third issue is the difficulty of maintaining side-channel chat conversations in an XR environment. Lastly, sharing and interacting with complex and dynamic multimedia content in XR remains to be tedious and error-prone.

Recently, an exciting alternative approach to remotely presenting rich multimedia content with remote audiences has emerged: the combination of publicly-available computer vision and speech recognition models with commodity webcam and microphones has the potential to bring the immersive experience of XR to remote communication experiences without abandoning a familiar desktop environment. This combination allows for real-time video compositing and background segmentation, pose and gesture recognition, and voice commands, thereby giving presenters multiple ways to interact with shared multimedia content. For brevity, we will refer to this approach as augmented video interaction.

Our proposed seminar is an opportunity to gather those who are similarly captivated by the potential of augmented video interaction for remote communication. We have already seen applications in this space for presenting business intelligence content in enterprise scenarios [6, 43], for presenting STEM topics in online education[73], for personalized product marketing [73], and for interacting with large cultural collections [92], such as those in gallery and museum archives.

This seminar is at the crossroads of human computer-interaction, computer-supported collaborative work, data visualization, and online education. The research community can take advantage of the recent technological improvements to forecast future usage scenarios, and this seminar aims to discuss and develop technical approaches and design guidelines for achieving effective multimodal remote communication and collaboration. The scope of our proposed seminar can also be summarized by (Figure 1).

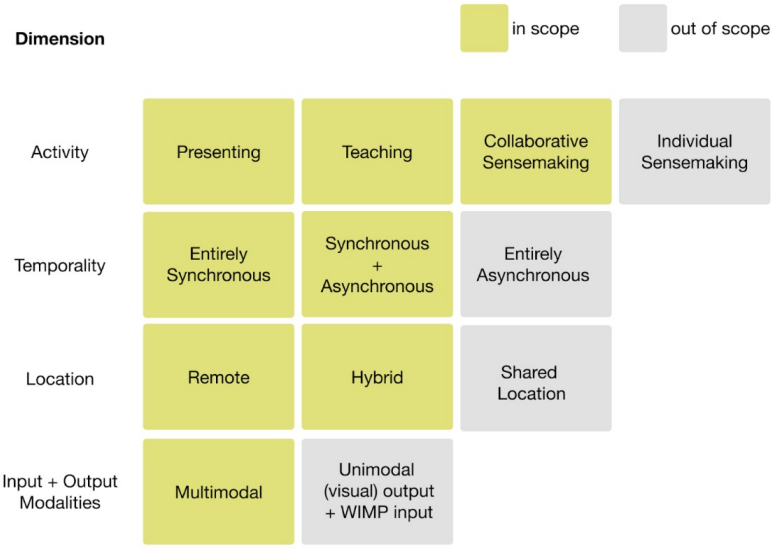

Figure 1: This diagram summarizes the scope of our proposed workshop on augmented multimodal interaction for synchronous presentation, collaboration, and education with remote audiences

Related Prior Workshops and Seminars

Our seminar overlapped thematically with several prior workshops, as indicated in Figure 2. While co-organizer M. Cordeil was a co-organizer of prior immersive analytics workshops as well as related Dagstuhl and Shonan seminars, our seminar was not restricted to immersive techniques applied to communication and collaboration.

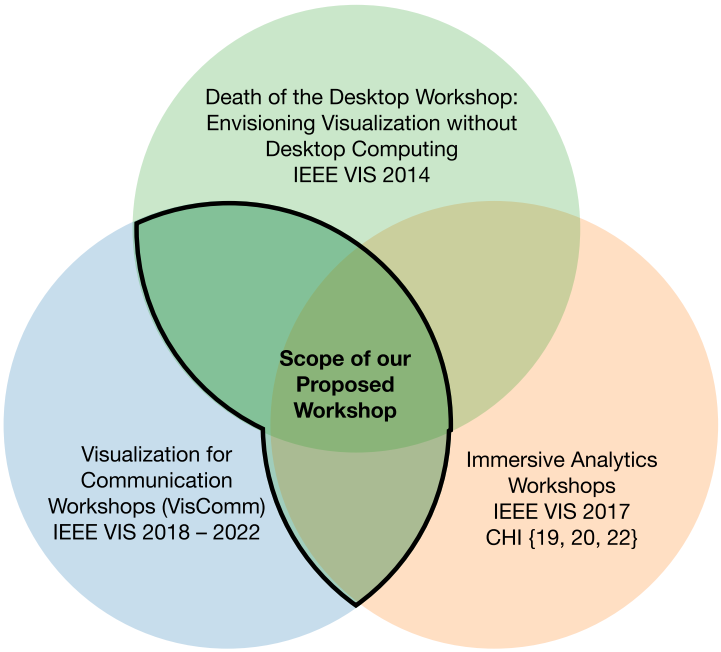

Figure 2: The area enclosed by the black contour indicates the thematic overlap between our seminar and prior events.

This seminar is a continuation and extension of the themes discussed at the MERCADO workshop at IEEE VIS 2023 (Multimodal Experiences for Remote Communication Around Data Online, see Figure 3), coordinated by the same team of organizers as this seminar; an archival version of the IEEE VIS workshop proposal can be found on arXiv [7]. This half-day workshop took place on Sunday, October 22 2023 in Melbourne at IEEE VIS 2023 and consisted of two 75-minute workshop sessions. The sessions included an invited keynote talk by Tom Bartindale of Monash University (Australia), six paper presentations [40, 58, 63, 71, 102, 116], and a panel discussion with Tom Bartindale, Andrea Batch (Bureau of Economic Analysis, USA), and Andy Cotgreave (Tableau, UK). The panel discussion touched upon eight themes and several sub-topics:

- **The role of a presenter**: Active / passive / others (shared roles-transitioning), ...
- **Design considerations**: Physical metaphor of the physical place, good or bad, hybrid teaching, social interaction: what new paradigms?, ...
- **Situation awareness**: Engagement, feedback, remote technologies feedback, motion detection, virtual room benchmarks (organiser view), ...

Figure 3: Highlights from the MERCADO workshop at IEEE VIS 2023

- **Automation and scaffolding**: Scripting, subtitle - translation (other language), play pause (time stretching), game / questionnaires, ...
- **AI and automatization**: e.g., IBM Debater, PICOVOICE: Voice Assistance-speech commands systems, assist disabled users with novel technologies, ...
- **Benefits / pros**: Shared physical space, one to one space, direct audience attention, same shared experience, turn everyone into Hans Rosling, No extra time to build presentations, ...
- **Limitations / cons**: Device limitation of social interactions, failure of interaction, midas effect, usage of raycasting techniques (accuracy), 2D/3D size Assessment, practice to make it work, multimodality (eye tracker, voice, gesture, physical switch), translation with specific or technical language, Visual literacy (lies for viz and lies with presentations), ...
- **The future of MERCADO**: How to assess the different techniques, EEG analytic to gain user KPI, Trust into the system, ...

Seminar Organizers

Matthew Brehmer is an Assistant Professor with the Assistant Professor in the David. R. Cheriton School of Computer Science at the University of Waterloo (Canada). He specializes in new experiences for interpersonal communication with and around data (e.g., [8, 6, 43]). Prior to joining The University of Waterloo, he was a Lead Research Staff with Tableau Research in Seattle (USA) and a postdoctoral researcher at Microsoft Research in Redmond (USA). He completed his PhD research on information visualization at the University of British Columbia. He was a co-organizer of the first workshop on visualization on mobile devices at CHI 2018 [66], and a co-organizer of the VisInPractice [111] event at IEEE VIS between 2018 and 2021. In 2022, he was elected to the VIS Executive Committee (VEC) and appointed to the IEEE Visualization and Computer Graphics Technical Community (VGTC) Executive Committee. **Website**: mattbrehmer.ca.

Maxime Cordeil is a Senior Lecturer at the University of Queensland, Australia. Dr. Cordeil has been recognised Australia's top researcher in computer graphics (2021, 2022). His research focuses on human-computer interaction, data visualisation and analytics. He has published over 60 journal and conference in top venues such as ACM CHI, IEEE VIS or IEEE VR. Dr. Cordeil is a key international member of the Immersive Analytics community of researchers, and has organised several workshops on the topic of Immersive Analytics ("IA Workshop series" at VIS 2017, CHI 2018, CHI 2019, CHI 2020, and CHI 2022). The activities of the IA community focuses on designing and evaluating the future graphical user interfaces for data analysis in Virtual / Augmented Reality. **Website**: sites.google.com/view/cordeil.

Christophe Hurter is a Professor working at the University of Toulouse, France, leading the Interactive Data Visualization group (DataVis) of the French Civil Aviation University (ENAC). His research covers explainable A.I. (XAI), big data manipulation and visualization (InfoVis), immersive analytics, and human-computer interaction (HCI). He investigates the design of scalable visual interfaces and the development of pixel-based techniques. He is an associate researcher at the research center for the French Military Air Force Test Center (CReA, Base militaire de Salon de Provence) and at the Brain and Cognition Research Center (CerCo, Hospital University Center of Toulouse). He published 2 books, 4 book chapters, 20 patents, 25 journal papers, more than 100 per reviewed international research papers. **Website**: recherche.enac.fr/ hurter.

Takayuki Itoh is a full professor of the department of information sciences in Ochanomizu University, Japan since 2011, and the director of the center for artificial intelligence and data science of the university since 2019. He was a researcher at Tokyo Research Laboratory of IBM Japan during 1992 to 2005. He has been an associate professor in Ochanomizu University since 2005, and a full professor since 2011. He is the general chair of Graph Drawing 2022, the general chair of IEEE Pacific Visualization 2018, short paper co-chair of IEEE VIS 2023, and organizing members of other many international conferences. His representative studies include fast isosurface generation [52, 57], hierarchical data visualization [56, 55], network visualization [54, 51, 74] and multidimensional data visualization [53]. **Website**: itolab.is.ocha.ac.jp/ itot.

Seminar Goals and Research Questions

The goal of our seminar was to identify emerging challenges in data-rich multimodal remote collaboration. It also served as a forum to identify new application scenarios and expand upon existing ones, and as a forum to test new interaction techniques applicable across these scenarios. We anticipate that the results of this meeting may include contributions to several academic communities including those affiliated with the IEEE VGTC (VIS, PacificVis, ISMAR, Eurovis) and ACM SIGCHI (CHI, CSCW, UIST, ISS).

We invited participants to submit, challenges, ideas, and reflections relating to the following research questions:

- Considering recent work exploring the design space of synchronous remote collaboration around data via a shared WIMP interface [82, 96], how can these techniques be extended to incorporate additional modalities?

- Considering techniques for co-located collaborative work around data, whether using conventional desktop workstations (e.g., [76]) or immersive augmented reality head-mounted displays (e.g., [11]), how can these techniques be expanded to support hybrid or remote collaboration and communication?

- Considering presentation techniques employed by television news broadcasters for presenting technical or data-rich stories [27] (e.g., weather, finance, sports), how can these techniques be applied (while keeping production costs low) and expanded to support multi-party multimodal interaction around data at a distance?

- Considering presentation and video compositing techniques employed by livestreamers [21, 118] (e.g., Twitch, YouTube, Facebook Live) and recorded video content creators (e.g., YouTube, TikTok), how can these techniques be applied during synchronous communication and collaboration around data, as well as in conjunction with multimodal interaction (e.g., pose / gesture input, voice prompts, proxemic interaction)?

- Considering the techniques by which individuals display and interact with representations of data in XR (AR / VR, e.g., [16, 67]) how can we extend or adapt these techniques? In other words, how can techniques initially designed with expensive or exclusive hardware be adapted to low-cost, accessible commodity input and output devices? Similarly, how could augmented video techniques designed for use with depth sensors [95] and pointing devices [89] be similarly adapted?

- Currently, most teleconference applications assume a single speaker / presenter, with other participants in audience roles. How can we support multi-party augmented video interaction, such as with Grønbæk et al.'s MirrorBlender project [42]? Alternatively, how can we support both *'sage on the stage'* and *'guide on the side'* style communication [61], differentiating an orator from a discussion facilitator. Similarly, how can we support formal, linear, and scripted presentations as well as informal, unscripted, interruption-prone, and collaborative discussions? Using Brehmer and Kosara's musical performance analogy [8], the former experience is likened to a *'concert recital* while the latter one is likened to a *'jam sessions'* [8]).

- Overall, what are the dimensions of the design space for multimodal and synchronous communication and collaboration around data? Where does existing work fit within this design space and which parts remain underexplored?

As of writing, we plan to report the emerging challenges that we identified in a top quality outlet research venue, such as the ACM SIGCHI conferences (CHI, UIST, ISS, CSCW) or those associated with the IEEE Visualization and Graphics Technical Community (VGTC, including VIS and PacificVis).

List of Attendees

Figure 4: The in-person attendees of Shonan Meeting #213: *Back row (L to R)*: Tim Dwyer, Zhu-Tian Chen, Christophe Hurter, Arnaud Prouzeau, Harald Reiterer, Jian Zhao, Wolfgang Büschel, Mahmood Jasim, Matthew Brehmer, Jonathan Schwabish, Maxime Cordeil, Ryo Suzuki, Andrew Cunningham, Lyn Bartram, Bongshin Lee, Brian Smith, Yasuyuki Sumi, Alark Joshi, Masahiko Itoh, Bektur Ryskeldiev, Kiyoshi Kiyokawa. *Front row*: Takayuki Itoh, Samuel Huron, David Saffo, Hideaki Kuzuoka, Anthony Tang, Gabriela Molina León.

Organizing team:

- **Matthew Brehmer**, University of Waterloo (Canada)
- **Maxime Cordeil**, Queensland University (Australia)
- **Christophe Hurter**, ENAC / University of Toulouse (France)
- **Takayuki Itoh**, Ochanomizu University (Japan)

*Invited attendees; * = joined remotely:*

- **Lyn Bartram**, Simon Fraser University (Canada)
- **Wolfgang Büschel**, TUD Dresden University of Technology (Germany)
- **Sheelagh Carpendale**, Simon Fraser University (Canada)*
- **Zhu-Tian Chen**, University of Minnesota (USA)
- **Andrew Cunningham**, University of South Australia (Australia)

- **Tim Dwyer**, Monash University (Australia)
- **Mar Gonzalez Franco**, Google (USA)*
- **Eric Gonzalez**, Google (USA)*
- **Samuel Huron**, Institut Polytechnique de Paris (France)
- **Petra Isenberg**, INRIA / Université Paris-Saclay (France)*
- **Masahiko Itoh**, Hokkaido Information University (Japan)
- **Mahmood Jasim**, Louisiana State University (USA)
- **Alark Joshi**, University of San Francisco (USA)
- **Kiyoshi Kiyokawa**, Nara Institute of Science and Technology (Japan)
- **Hideaki Kuzuoka**, University of Tokyo (Japan)
- **Bongshin Lee**, Yonsei University (Republic of Korea)
- **Gabriela Molina León**, University of Bremen (Germany)
- **Arnaud Prouzeau**, Inria (France)
- **Harald Reiterer**, University of Konstanz (Germany)
- **Bektur Ryskeldiev**, Mercari R4D (Japan)
- **David Saffo**, JPMorgan Chase & Co. (USA)
- **Jonathan Schwabish**, Urban Institute (USA)
- **Brian Smith**, Columbia University (USA)
- **Yasuyuki Sumi**, Future University Hakodate (Japan)
- **Ryo Suzuki**, University of Colorado, Boulder (USA)
- **Anthony Tang**, Singapore Management University (Singapore)
- **Yalong Yang**, Georgia Tech (USA)*
- **Jian Zhao**, University of Waterloo (Canada)

Seminar Schedule

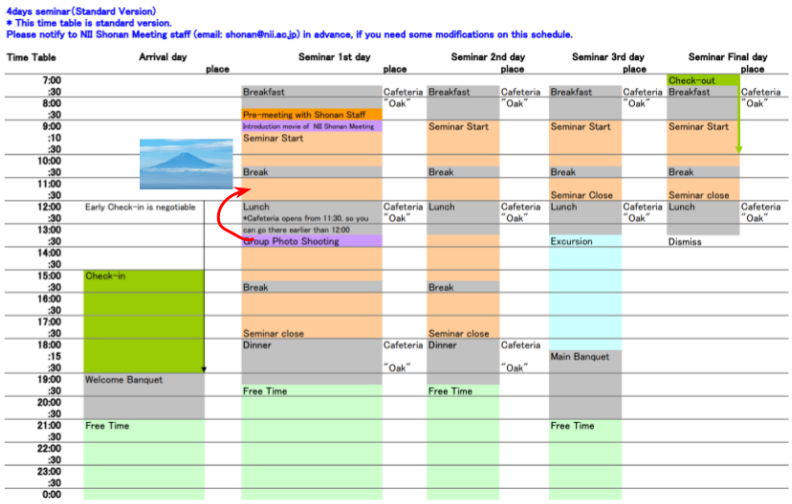

Figure 5: The schedule of Shonan Meeting #213

Check-in Day / June 23 (Sunday):
- Welcome banquet (19:00-20:30)

Day 1 / June 24 (Monday):
- Introduction and overview (9:00-9:50)
- Attendee presentations (9:50-11:30, 16:00-16:20)
- Group photo (11:30-12:00)
- Breakout session topic mapping and scheduling (13:00-14:45)
- Breakout sessions (14:45-15:30, 16:20-18:00)

Day 2 / June 25 (Tuesday):
- Attendee presentations (9:00-9:40, 11:00-11:20, 13:30-14:00)
- Breakout sessions (9:40-10:30, 14:30-15:30, 16:00-18:00)
- Breakout groups report (11:20-12:00)

Day 3 / June 26 (Wednesday):
- Breakout sessions (9:00-9:45)
- Breakout groups report (9:45-11:00)
- Unconference session (11:15-12:00)
- Excursion and main qanquet (13:30-21:00)

Day 4 / June 27 (Thursday):
- Breakout sessions (9:00-10:30)
- Final discussion (11:00-12:00)

Attendee Presentations

Each attendee gave a five-minute presentation, following a template provided by the organizers. This template included a brief biography and a list of broad research interests, a summary of their recent work related to the topic of the seminar (Figure 6), and a brief reflection on their goals or hopes for attending the seminar. We summarized the last of this in Figure 7.

Apart from the four organizers, who presented in alphabetical order, the order of attendee presentations was random, with accommodations for remote attendees to present in a time that was convenient for their schedule.

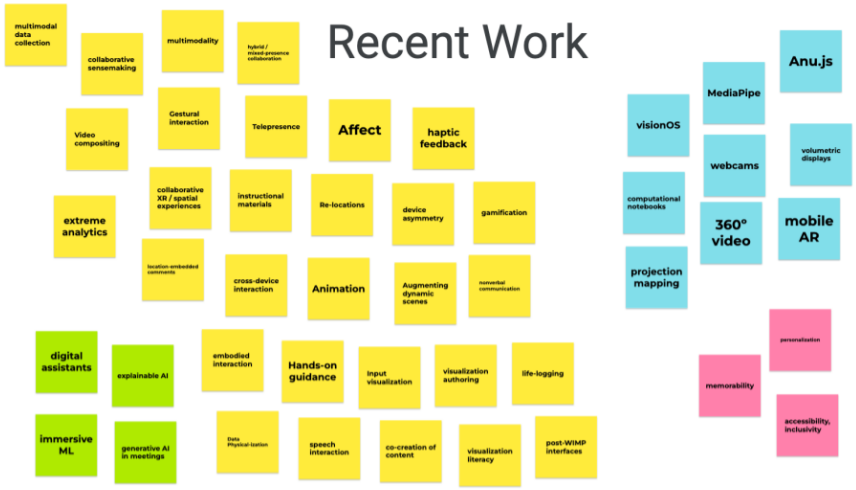

Figure 6: A high-level classification of attendees' recent work related to this seminar, with yellow denoting HCI research keywords, green denoting AI keywords, pink denoting constructs of evaluation, and blue denoting specific technologies.

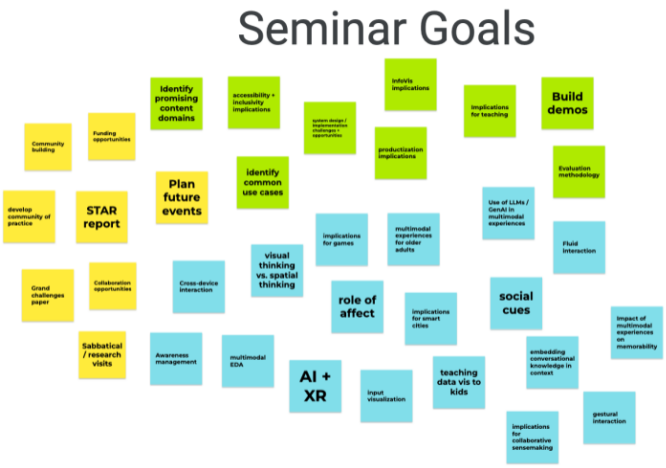

Figure 7: A high-level classification of attendees' goals and hopes for attending this seminar, with yellow denoting future research deliverables or events, green denoting immediate goals for meeting activities, and blue denoting desire topics of discussion.

Matthew Brehmer, University of Waterloo (formerly, Tableau)

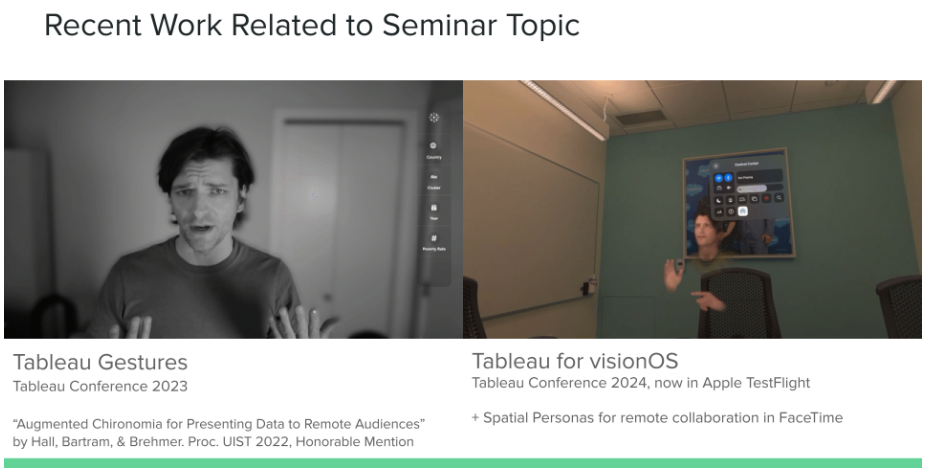

Figure 8: Matthew Brehmer's recent work related to the seminar.

Brehmer introduced research interests including communication & collaboration around data (especially in enterprise / organizational contexts) [8], encompassing presentation and storytelling. He is also interested in InfoVis beyond charts, and information *ubiety* [12].

His recent work related to seminar topic (Figure 8) includes Tableau Gestures [43] (and its extension with voice commands [102]), Tableau for visionOS [98], and VisConductor [30]

His goals and hopes for this seminar include community building, extending MERCADO @ VIS 2023 [7], initiating a Grand Challenges / Vision / State-of-the-Art paper, starting project collaborations, identifying funding opportunities, co-advising students, and develop a consensus on how to evaluate and compare these experiences.

Maxime Cordeil, Queensland University

Figure 9: Maxime Cordeil's recent work related to the seminar.

Cordeil introduced research interests including Visualisation beyond the desktop, Interactive AI for data exploration, Collaboration, large/big data, Network, and MD-data.

His recent work (Figure 9) related to seminar topic includes ImmersiveIML [24] and Hanstreamer[63].

His goals and hopes for this seminar included connecting with others on the topic, setting up a research agenda for augmented presentations, a grand challenge paper, and Have fun (building demos).

Christophe Hurter, ENAC / University of Toulouse

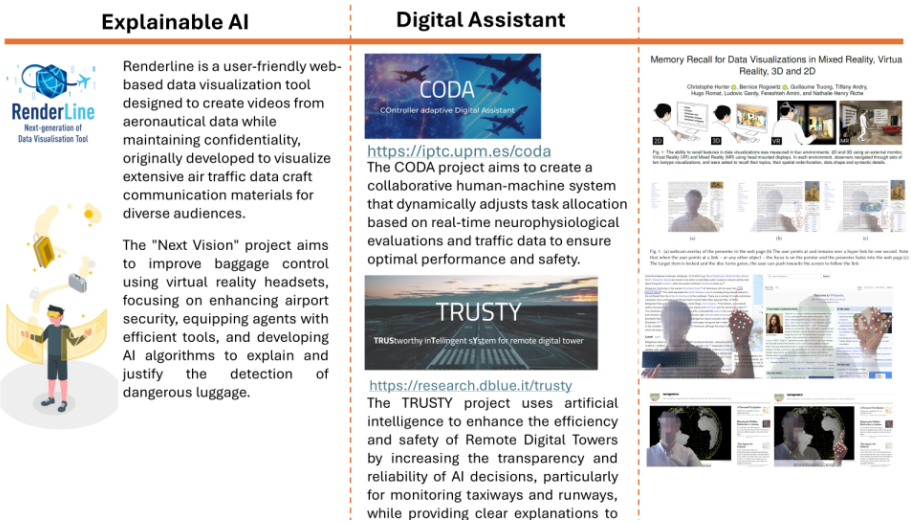

Figure 10: Christophe Hurter's recent work related to the seminar.

Hurter introduced research interests including explainable AI, Big data, Information visualization, Immersive analytics, and Human-Computer interaction.

His recent work (Figure 10) related to seminar topic includes work on explainable AI and Digital assistants, and experimental work on information recall in mixed reality [48].

His goals and hopes for this seminar included building an up-to-date map of the state of the art, understanding current trends and forecasting future directions and technical challenges for leveraging teaching capabilities, exploring novel technologies, and leveraging memorability for students to enhance understanding of complex data processing or physical phenomena.

Takayuki Itoh, Ochanomizu University

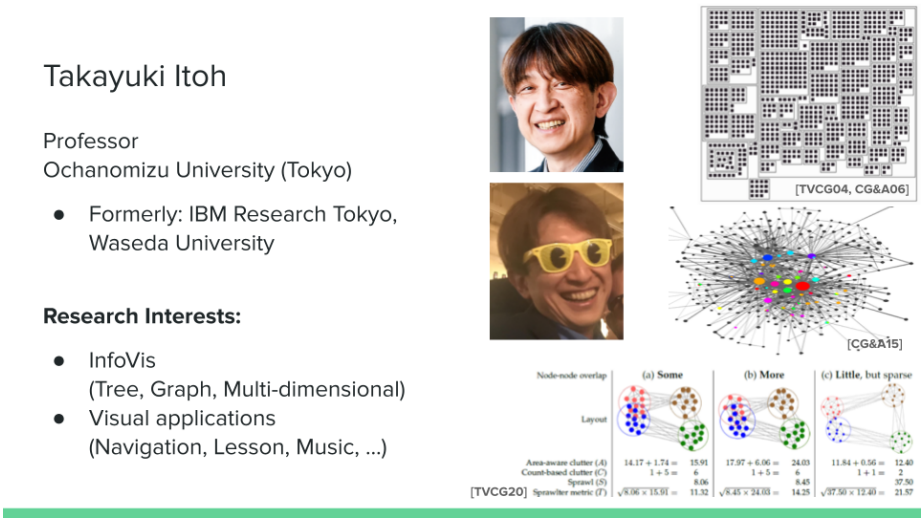

Figure 11: Takayuki Itoh's recent work related to the seminar.

Itoh introduced research interests including InfoVis (Tree, Graph, Multi-dimensional) and Visual applications (Navigation, Lesson, Music).

His recent work related to seminar topic included navigation with location-embedded comments [85] and Visualization for lesson of dance and karuta [62].

His goals and hopes for this seminar included finding new topics from the InfoVis side, opportunities for joint projects and student visits, as well as planning publications and next events.

Lyn Bartram, Simon Fraser University

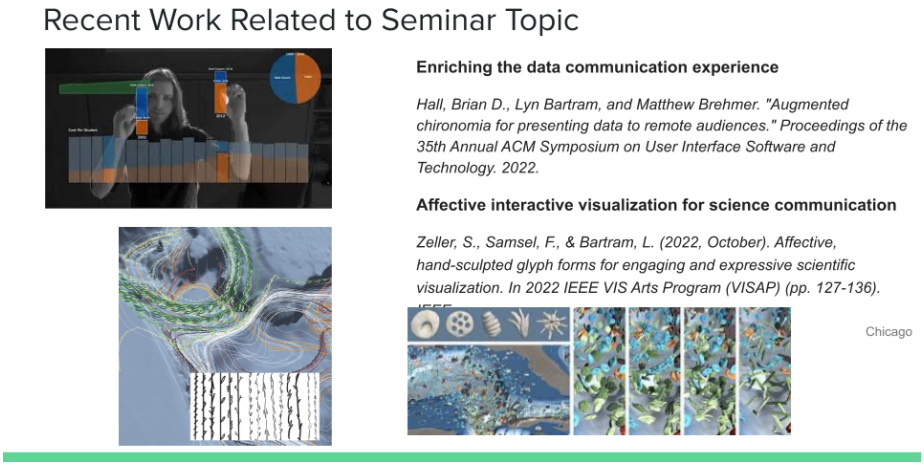

Figure 12: Lyn Bartram's recent work related to the seminar.

Bartram introduced research interests including perception and visual theory, the role of affect in data thinking and visualization design, artistic knowledge and practice for richer visual language, how motion and movement carry both rich data and meaning, sensemaking beyond analytics, and cognitive systems engineering.

Her recent work (Figure 12) related to seminar topic included enriching the data communication experience [43] and affective interactive visualization for science communication [117].

Her goals and hopes for this seminar included more insight into enriching the visual language of data through gesture, opening up the research space of affective data + human interpretation, and exploring new applications for collaborative sensemaking.

Wolfgang Büschel, TUD Dresden University of Technology

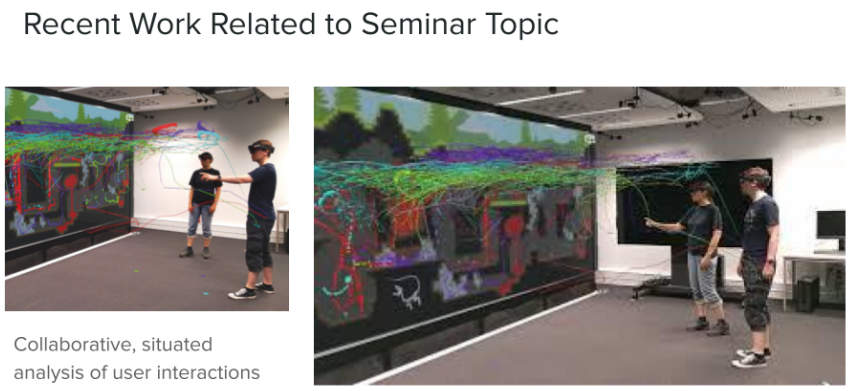

Figure 13: Wolfgang Büschel's recent work related to the seminar.

Büschel introduced research interests including mixed reality, natural 3D interaction, and 3D data visualization.

Wolfgang Büschel presented his research interests at the recent workshop (Figure 13). His work primarily focuses on Mixed Reality and Natural 3D Interaction, with a keen interest in Immersive and Situated Data Visualization. He shared insights from his latest projects, which align closely with the seminar's theme, highlighting his contributions to Collaborative and Situated Analysis of User Interactions. MIRIA [10] is a toolkit that supports this type of in-situ data analysis to explicitly couple the analysis to the data's spatial origin. Building on this toolkit is a series of visualizations to investigate human movement in situ [75]. Finally, he presented novel, multi-modal interaction with transparent tablets to support interaction with visualizations and annotations in mixed-reality environments [64]. Büschel also examines Multimodal Interaction within Mixed Reality environments and explores the dynamics of Remote and Mixed-Presence Collaboration. Furthermore, he expressed his goals for the seminar: to engage with fellow participants, brainstorm on potential paper topics, and foster partnerships that could lead to future collaborative efforts.

Sheelagh Carpendale, Simon Fraser University

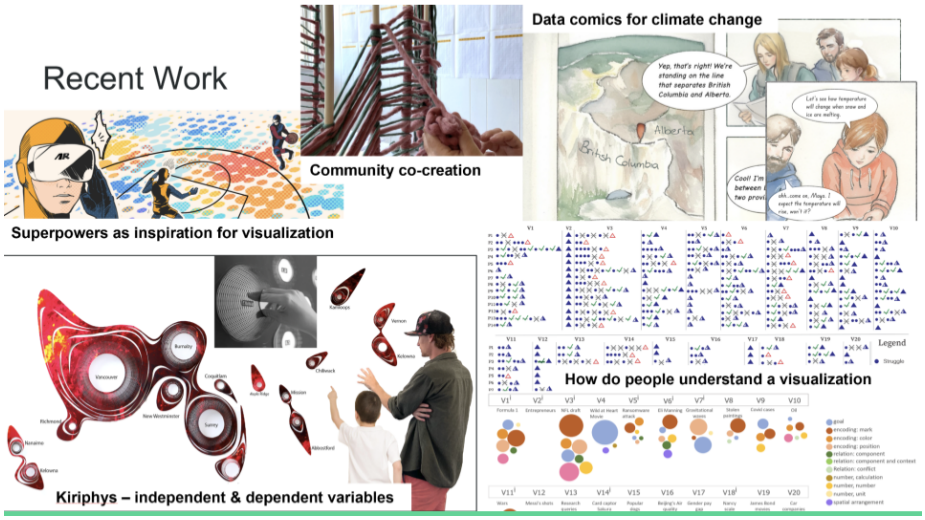

Figure 14: Sheelagh Carpendale's recent work related to the seminar.

Carpendale introduced research interests including Data Visualization and Data Physicalization, the co-design of datavis with the creative community, data comics for climate change, and taking a fresh look at interaction for comprehension and collaboration.

Her recent work (Figure 14) related to seminar topic included superpowers as inspiration for visualization[112], community co-creation, data comics for climate change [72], and Kiriphys [25].

Her goals and hopes for this seminar included making connections and re-connections with peers, new and / or rekindled collaborations, and taking a fresh look at interaction for comprehension and collaboration.

Zhu-Tian Chen, University of Minnesota

Figure 15: Zhu-Tian Chen's recent work related to the seminar.

Zhu-Tian introduced research interests including augmenting human intelligence in daily activities through interactive visual representations and human-centered XR + AI.

He introduced recent work (Figure 15) related to the seminar topic including augment dynamic scenes [19, 15, 17, 18], immersive visualization, and embodied interactions for data analysis in VR.

His goals and hopes for this seminar included making some friends, discussing / learning about the difference between Visual Thinking vs Spatial Thinking (e.g., differences in cognition and perception?), learning from other folks' great works, starting project collaborations, and identify funding opportunities.

Andrew Cunningham, University of South Australia

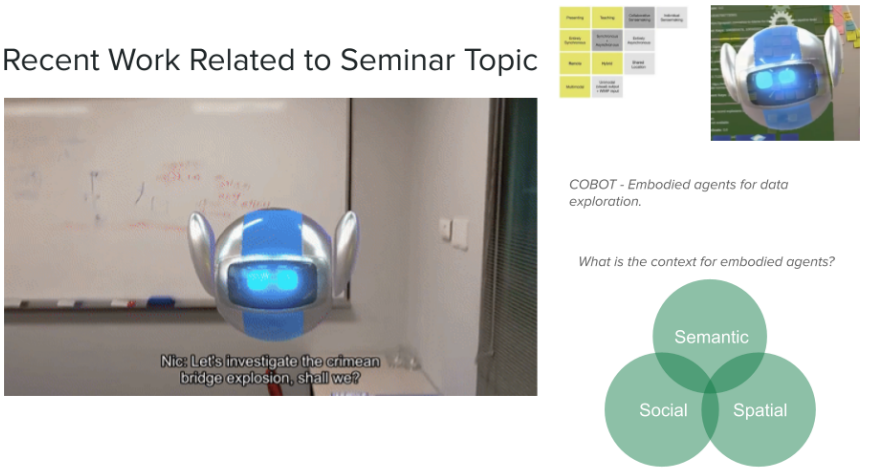

Figure 16: Andrew Cunningham's recent work related to the seminar.

Cunningham introduced research interests including immersive visualisation and analytics applied to various domains, novel interaction techniques, and immersive data storytelling.

His recent work (Figure 16) related to seminar topic includes collaborative sensemaking in hybrid environments [79], LogAR (XR Core Logging for geologists), and COBOT (embodied agents for data exploration).

His goals and hopes for this seminar include developing international collaborations, inviting people to come visit South Australia, developing a "grand challenges" paper to propel collaborative sensemaking forward in the community, and to eat great food together.

Tim Dwyer, Monash University

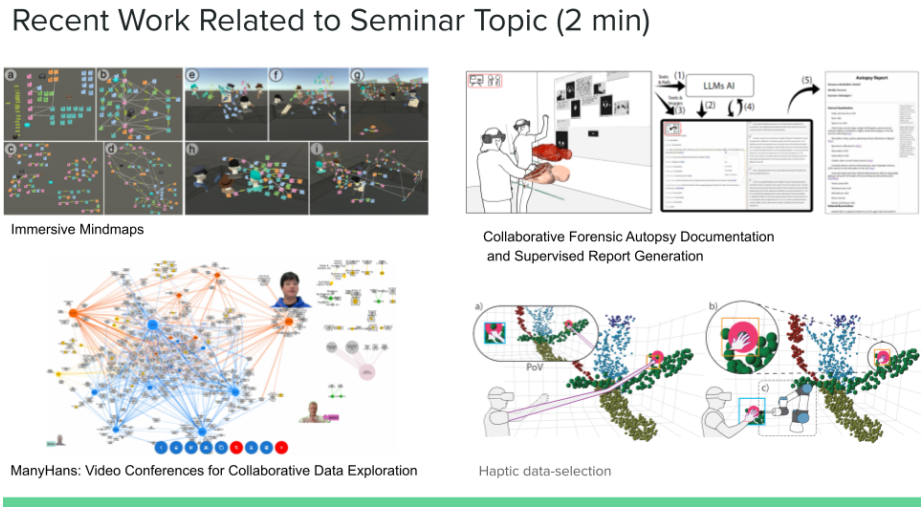

Figure 17: Tim Dwyer's recent work related to the seminar.

Dwyer introduced research interests including Information Visualisation and Immersive Analytics [28], Networks and complex data, Collaborative sensemaking and decision support, A smattering of computational geometry, and Computational multimodal interaction.

His recent work (Figure 17) related to seminar topic included Immersive Mindmaps [114], Collaborative Forensic Autopsy Documentation and Supervised Report Generation [90], ManyHans [63] (Video Conferences for Collaborative Data Exploration), and Haptic data selection.

His goals and hopes for this seminar included defining the state of the art? (What are the capabilities and limitations of current techniques, devices, modalities), forecasting the future of collaborative sensemaking (In 5 years, 20 years, 50 years...), and asking What should we be researching (to be as impactful as possible). He also hopes to organize future visits and host others in Melbourne for sabbaticals in 2025.

Mar Gonzalez Franco and **Eric Gonzalez**, Google, USA

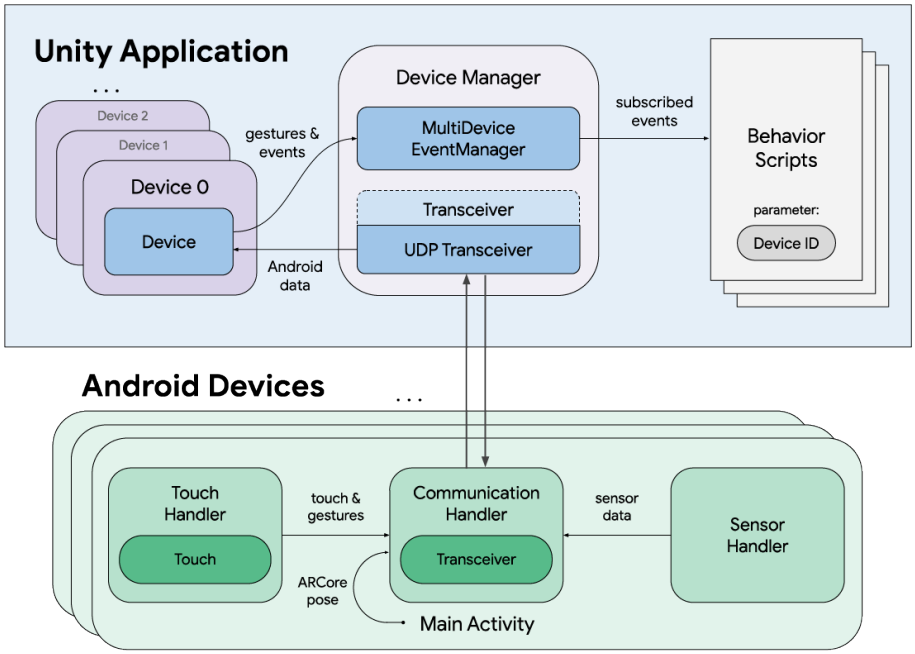

Figure 18: Mar Gonzalez Franco and Eric Gonzalez's XDTK [38], an open-source toolkit for building interactions between Android devices and a Unity application.

Mar Gonzalez Franco and Eric Gonzalez joined us remotely from Seattle for one session to introduce and demonstrate XDTK [38], an open-source toolkit for building interactions between Android devices and a Unity application (Figure 18).

Samuel Huron, Institut Polytechnique de Paris

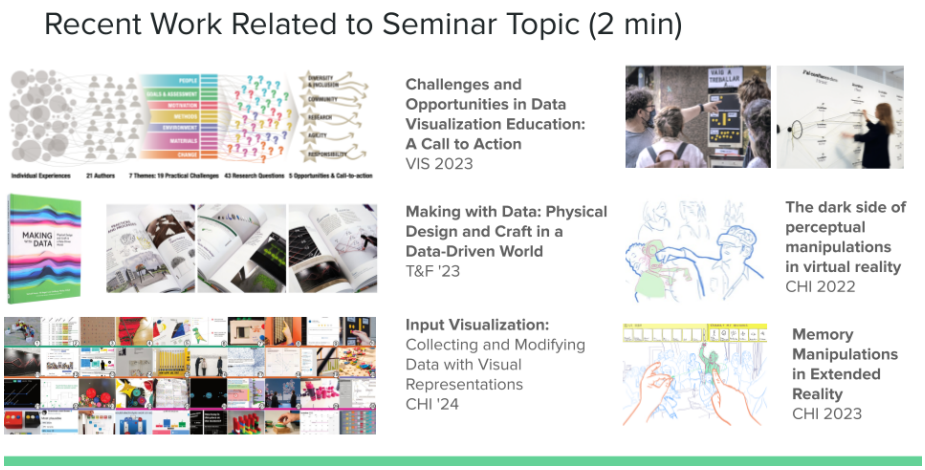

Figure 19: Samuel Huron's recent work related to the seminar.

Huron introduced his three contemporaneous research interests, including 1) democratizing data representation, how humans form data representation, and 2) how to study the negative impact of digital technology, including manipulating users in XR and electronic waste. He introduced recent work related to seminar topics including *Input Visualization* [9], *Challenges and Opportunities in Data Visualization Education: A Call to Action* [3], *Making with Data: Physical Design and Craft in a Data-Driven World* [47], The dark side of perceptual manipulations in Virtual Reality [110], and Memory Manipulations in Extended Reality [5], and *Biohybrid Devices* [83]. He finally listed goals and hopes for this seminar: to better explore synchronous communication and collaboration, particularly situations related to input data through visualization. This includes situations such as public group reflection, public activity documentation, data discussion, interactive surveys, planning, and organizing with data. On this topic he is particularly interested in the interactions between tangible and digital. He is looking to learn from other participants and start new collaborations on papers, project, and exchanges, and he is also looking for a research team to do a research sabbatical.

Petra Isenberg, INRIA / Université Paris-Saclay

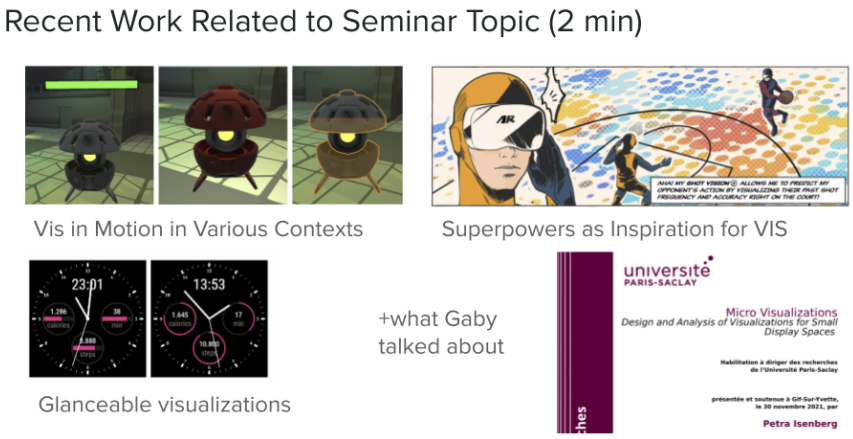

Figure 20: Petra Isenberg's recent work related to the seminar.

Isenberg introduced research interests including large + small displays for visualization, as well as embedded & situated visualizations.

Her recent work (Figure 20) related to seminar topic includes Vis in Motion in Various Contexts [115], Superpowers as Inspiration for VIS [112], glanceable visualizations [4], and speech-based interfaces for co-located collaborative analytics [71].

Her goals and hopes for this seminar included discussing the role of visualization to support remote audiences, to get more involved again in collaborative work research, and to experience the seminar remotely.

Masahiko Itoh, Hokkaido Information University

Figure 21: Masahiko Itoh's recent work related to the seminar.

Itoh introduced research interests including temporal network visualization, spatio-temporal visualization, visual analytics, sports visualization, and point clouds.

His recent work (Figure 21) related to seminar topic included snow sculpture point clouds from the Sapporo Snow Festival and 3D points cloud by Photogrammetry and NeRF. His goals and hopes for this seminar included building new connection and starting collaborations, and finding a place to stay during sabbatical.

Mahmood Jasim, Louisiana State University

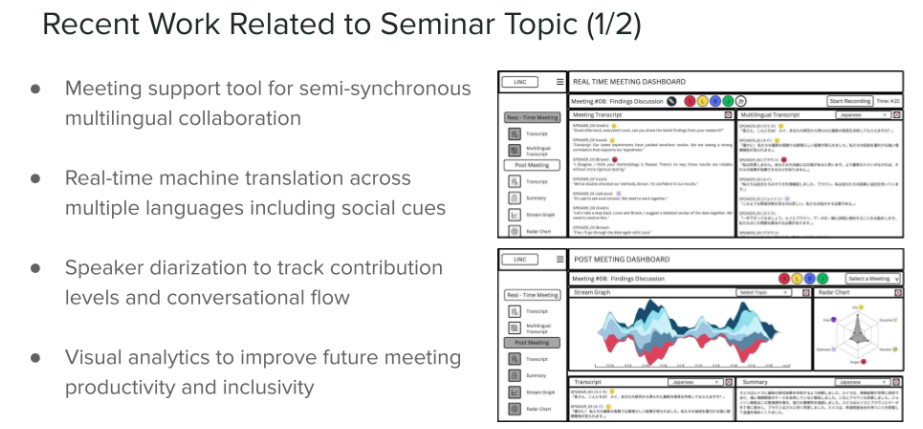

Figure 22: Mahmood Jasim's recent work related to the seminar.

Jasim introduced research interests including visualization tools for multimodal exploratory visual analytics, decision support systems for collaborative sensemaking, and interaction techniques to facilitate broader communication.

His recent work (Figure 22) relating to seminar topic includes a meeting support tool for semi-synchronous multilingual collaboration, real-time machine translation across multiple languages including social cues, speaker diarization to track contribution levels and conversational flow, and visual analytics to improve future meeting productivity and inclusivity [58].

His goals and hopes for this seminar included exploring potential for multimodal interactions to facilitate collaboration, exploring how generative AI and foundational models can be used to augment data transformations to facilitate more robust analytics, and identifying new avenues for multimodal exploratory visual analytics, and Build new connections and potential collaborations.

Alark Joshi, University of San Francisco

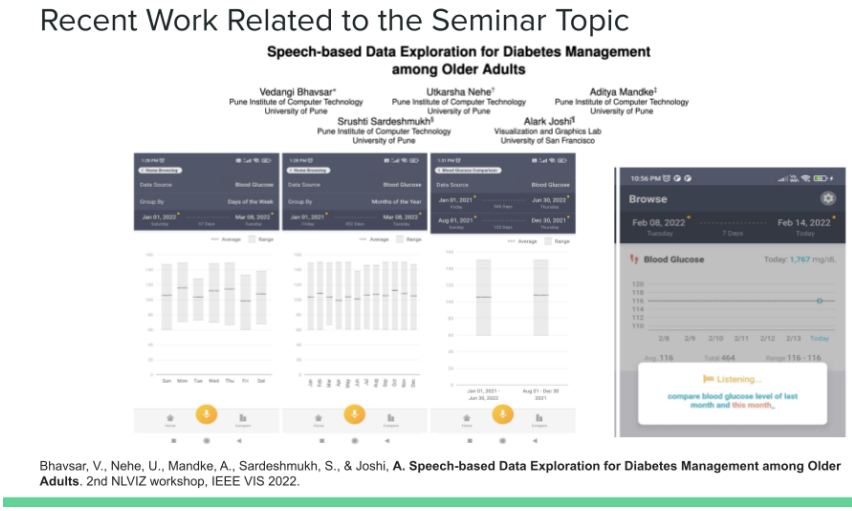

Figure 23: Alark Joshi's recent work related to the seminar.

Joshi introduced research interests including multimodal interaction for data exploration [44, 37], collaborative decision-making with visualization, and visualization literacy [33].

His recent work (Figure 23) related to seminar topic includes visualization literacy for parallel coordinate plots literacy [88] and treemap literacy [34].

His goals and hopes for the seminar included identifying potential collaborators for projects related to work with multimodal interaction, collaborative learning in VR settings, empowering older adults, developing a state-of-the-art report to identify future directions for research for the community, and building and sustaining a Community of Practice (CoP) for multimodal experiences with data.

Kiyoshi Kiyokawa, Nara Institute of Science and Technology

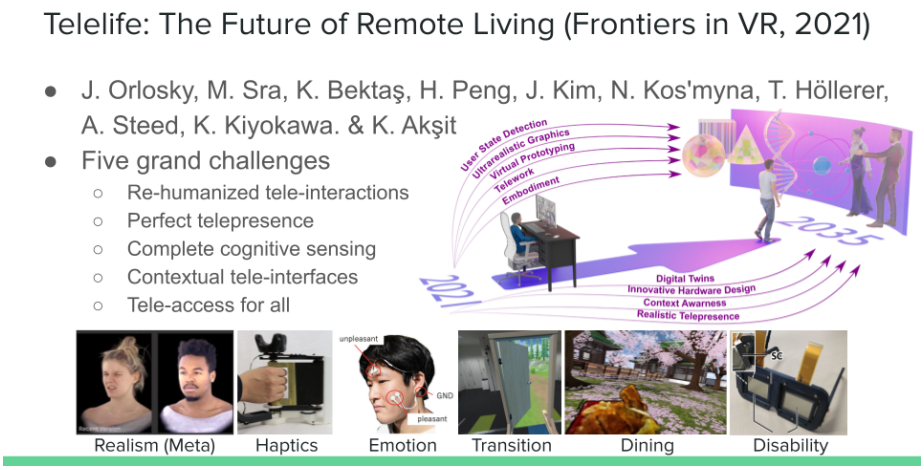

Figure 24: Kiyoshi Kiyokawa's recent work related to the seminar.

Kiyokawa introduced research interests including VR, AR, AH, collaboration, displays, multimodal, and assistive interfaces.

His recent work (Figure 24) related to seminar topic including *Telelife: The Future of Remote Living* [87].

His goals and hopes for the seminar included making new connections, identifying research challenges, starting new collaborations, and having fun.

Hideaki Kuzuoka, University of Tokyo

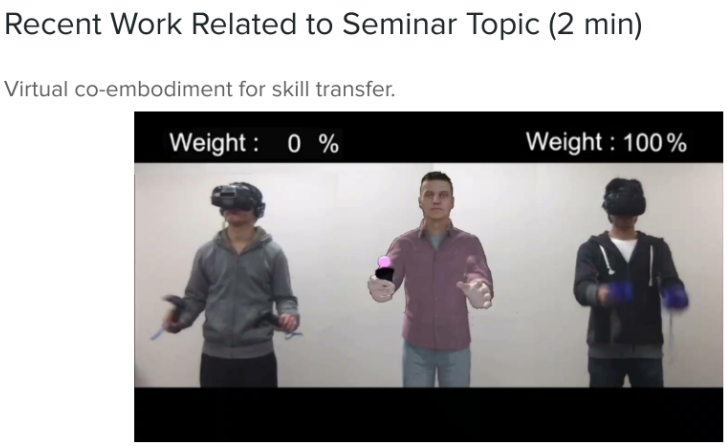

Figure 25: Hideaki Kuzuoka's recent work related to the seminar.

Kuzuoka introduced research interests including CSCW (Remote assistance on physical tasks and casual communication), Human-Robot Interaction, and Virtual Reality.

His recent work (Figure 25) related to seminar topic includes augmented reality (gesture overlay, robotic approach), and virtual co-embodiment for skill transfer.

His goals and hopes for this seminar included merging physical and digital, incorporating AI into XR, going beyond face-to-face, and designing/finding appropriate way for showing social cues.

Bongshin Lee, Yonsei University

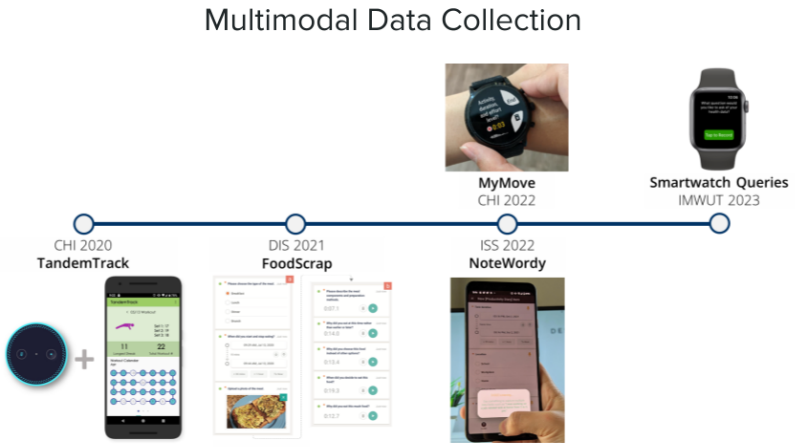

Figure 26: Bongshin Lee's recent work related to the seminar.

Lee introduced her research interests in advancing human-data interaction, including multimodal interaction for data visualization, mobile data visualization, data-driven storytelling, and inclusive data visualization.

Her recent work related to seminar topic includes *Post-WIMP Interaction for Information Visualization* [68], with projects like InChorus [103], DataHand [60], and DataBreeze [104], as well as multimodal data collection (Figure 26).

Her goals and hopes for this seminar include learning about the latest advancements and innovative approaches for data-driven remote communication, exchanging experiences and ideas, exploring future collaboration opportunities, and promoting accessibility and inclusivity in the field of data-driven remote communication research.

Gabriela Molina León, University of Bremen

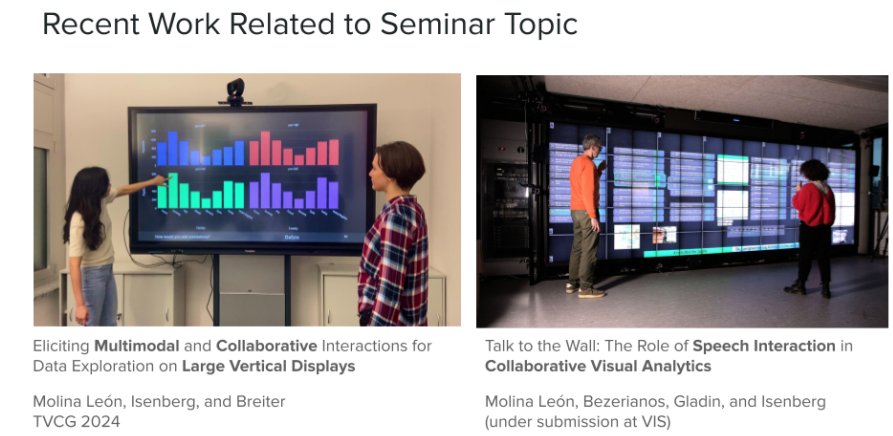

Figure 27: Gabriela Molina León's recent work related to the seminar.

Gabriela Molina León introduced her main research interests around the topics of multimodal interaction, collaborative visual analytics, and InfoVis beyond the desktop. She presented her recent work (Figure 27) on multimodal and collaborative interaction on large vertical displays, including an elicitation study [70] on how users prefer to interact with visualizations of spatio-temporal data using mainly touch and speech, and an exploratory study [69, 71] on how speech interaction can support collaborative sensemaking on wall-sized displays. Her goals for this seminar included not only getting to know the other participants and their latest work, but also brainstorming and starting collaborations around the questions: How to best combine multimodal and collaborative interactions? How to support accessibility in such complex setups? How can artificial intelligence enhance collaboration?

Arnaud Prouzeau, Inria

Figure 28: Arnaud Prouzeau's recent work related to the seminar.

Prouzeau introduced research interests including immersive analytics [28] (VR, AR, Large Display, ...), collaborative systems, and control room experiences.

His recent work (Figure 28) related to seminar topic includes remote practical activities [65], collaborative actuated tangible interactions.

His goals and hopes for this seminar included interesting discussions around collaboration and data, discussing the idea of experience, keep building a community, and establishing potential collaborations.

Harald Reiterer, University of Konstanz

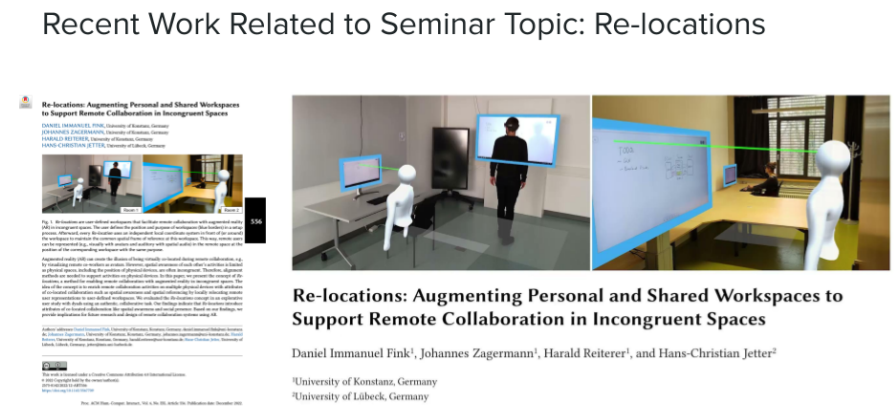

Figure 29: Harald Reiterer's recent work related to the seminar.

Reiterer introduced research interests including mixed reality and cross-device interaction.

His recent work (Figure 29) related to seminar topic include immersive analytics with cross-device interaction [46], immersive analysis of mixed-reality study data using cross-device interaction [45], and augmenting personal and shared workspaces to support remote collaboration in incongruent spaces [32].

His goals and hopes for this seminar included discussing the potential of cross-device interaction to achieve the seminar's goals, identifying possible design solutions supporting awareness management for MR collaboration, such as user representations, sound, re-locations, etc., and the applicability of *Fluid Interaction* and *Complementary Interfaces* to achieve the seminar's goals.

Bektur Ryskeldiev, Mercari R4D

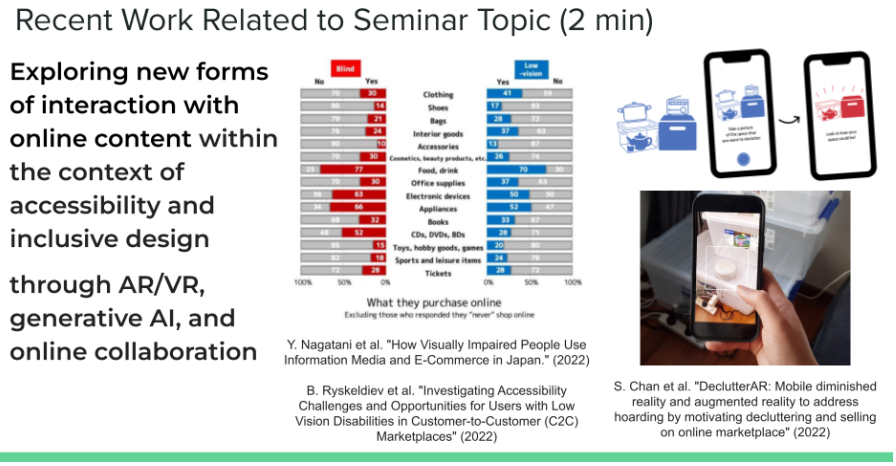

Figure 30: Bektur Ryskeldiev's recent work related to the seminar.

Ryskeldiev introduced research interests including AR / VR, Telepresence, HCI, Media Art, and Accessibility.

His recent work (Figure 30) related to seminar topic included *StreamSpace* [93]: Pervasive mixed reality telepresence for remote collaboration on mobile devices, How Visually Impaired People Use Information Media and E-Commerce in Japan, Investigating Accessibility Challenges and Opportunities for Users with Low Vision Disabilities in Customer-to-Customer (C2C) Marketplaces [94], and *DeclutterAR*: Mobile diminished reality and augmented reality to address hoarding by motivating decluttering and selling on online marketplace [14].

His goals and hopes for this seminar included understanding research directions for remote collaboration in AR/VR, outlining opportunities for accessibility and inclusive design, and finding research collaborators.

David Saffo, JPMorgan Chase & Co.

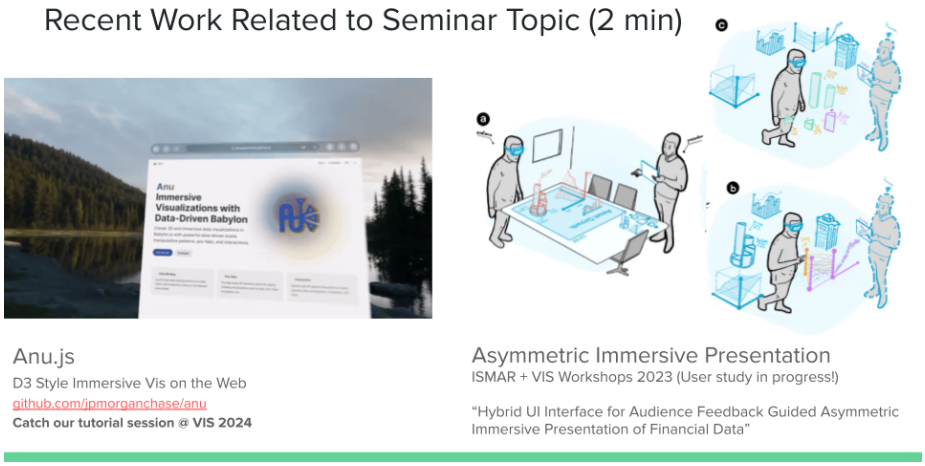

Figure 31: David Saffo's recent work related to the seminar.

Saffo introduced research interests including Immersive Analytics [28], Immersive Engagement, and Immersive Systems.

His recent work (Figure 31) related to seminar topic includes Anu.js and Asymmetric Immersive Presentation [40, 41].

His goals and hopes for this seminar included engaging and helping build the community, brainstorm application areas for immersive presentation techniques, discussing system design and implementation challenges opportunities, and to he inspired by participants.

Jonathan Schwabish, Urban Institute

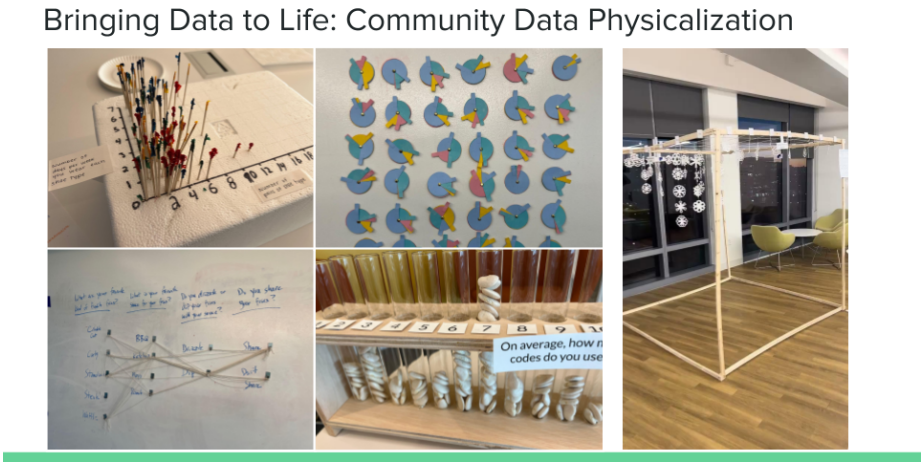

Figure 32: Jonathan Schwabish's recent work related to the seminar.

Schwabish [97] introduced research interests including economics, nutrition & disability policy, data literacy, physical dataviz, data equity, inclusion, and accessibility.

His recent work (Figure 32) related to seminar topic include *Bringing Data to Life: Community Data Physicalization.*

His goals and hopes for this seminar included community building, starting project collaborations, identify funding opportunities, exploring effective ways to teach data visualization to kids, accessibility considerations (e.g., stroke, TBI, etc.), and to have some fun.

Brian Smith, Columbia University

Figure 33: Brian Smith's recent work related to the seminar.

Smith introduced his research interests in determining how computers can facilitate meaningful experiences for people, spanning the realms of accessibility, social computing, and games. Regarding accessibility, he introduced his recent work in giving blind and low-vision people more direct and immersive means of exploring images, videos, and video games. Regarding social computing, he introduced his recent work (Figure 33) on countering the single-click interactions that we often see on social apps by enabling more authentic and effortful forms of communication. Regarding games, he introduced his recent work on understanding players' experiences from their interactions with a game including their button presses, then using that understanding to craft more meaningful experiences for players. Smith's stated goals for the seminar were to make friends and colleagues from the visualization community, find opportunities for product impact, and have fun in Japan.

Yasuyuki Sumi, Future University Hakodate

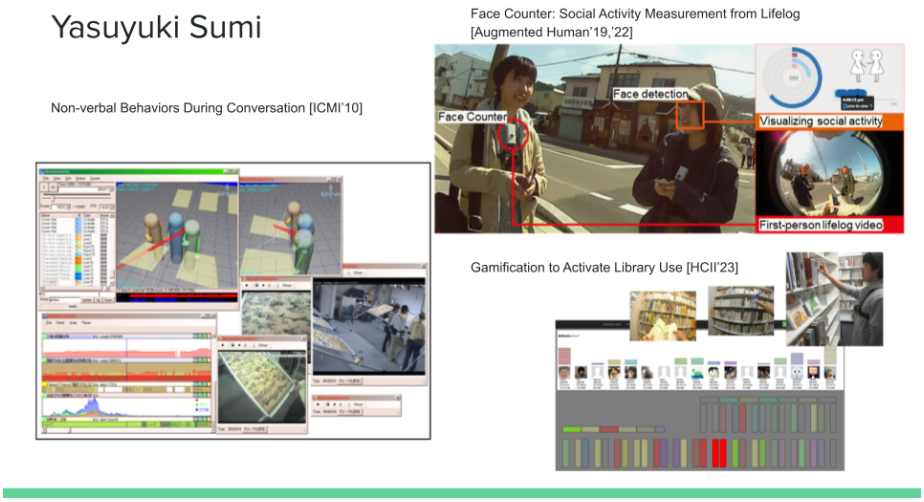

Figure 34: Yasuyuki Sumi's recent work related to the seminar.

Sumi introduced research interests including context-aware mobile assistants, collaborative capturing of experiences.

Sumi presented several of his own past studies (Figure 34) to discuss the augmented multimodal interaction between people in various spatio-temporal situations. First, he presented his past work [106, 107] that the measurement and analysis of verbal and non-verbal behaviors in in-person multiparty conversations (e.g., utterances, standing position, gestures, head movements, gaze, etc.) required the use of various special sensors 15 years ago. He then introduced a system called FaceCounter [86], which can measure the quantity and quality of daily social interactions using first-person view lifelog data, thanks to the development of computer vision in recent years. He also introduced other related research, including PhotoChat [105] that supports casual conversation among spatially dispersed users, an in-car conversation distribution system [78] that facilitates auralization of voices of the city, and a conversation embedding system [36] that facilitates sharing conversational knowledge across time in a shared space.

His goals and hopes for this seminar included embedding conversational knowledge in situation (contents vs. context?), getting to know participants and their work, and finding collaborators.

Ryo Suzuki, University of Colorado, Boulder (formerly, U. Calgary)

Figure 35: Ryo Suzuki's recent work related to the seminar.

Ryo Suzuki presented his research (Figure 35), which centers on augmenting human thought and creativity through interactive and immersive technologies. He highlighted recent projects that span augmented reality (AR), human-computer interaction (HCI), and artificial intelligence (AI), including RealityTalk [73], an AR platform for interactive communication, RealityCanvas [113], a space for collaborative creativity, RealitySketch [108], a system enabling users to sketch in real-world environments, Augmented Math [20], tools for enhancing mathematical thinking with AR, HoloBots [49], AR-driven robotic systems, and Sketched Reality [59], a hybrid platform that blends sketching with reality manipulation. In terms of his goals for the seminar, Suzuki emphasized the value of building connections with fellow researchers and fostering open discussions to explore future collaboration opportunities. He expressed a strong interest in co-authoring position or meta-level papers that could contribute to shaping the field. Additionally, Suzuki aimed to gain a deeper understanding of how AR can evolve in the areas of remote presentation, collaboration, and education, while exploring how people think about AI and AR in relation to the seminar's topics.

Anthony Tang, Singapore Management University

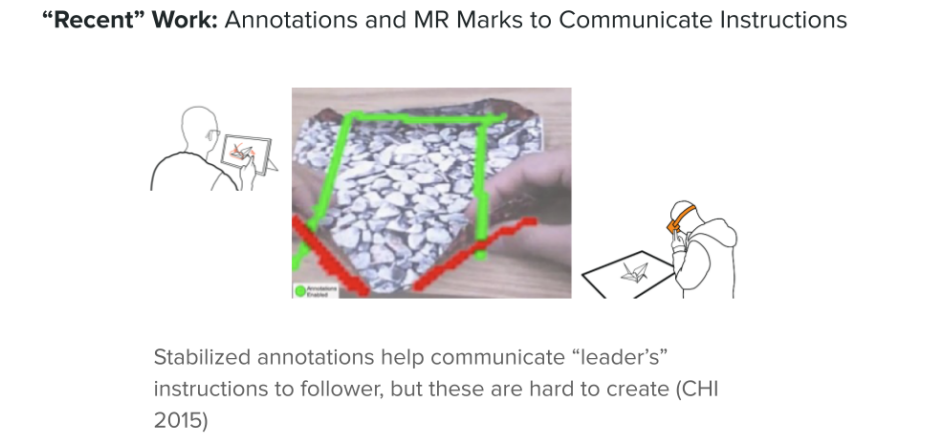

Figure 36: Anthony Tang's recent work related to the seminar.

In the seminar, Anthony shared some of the work his team has been doing around bodily representations in remote collaboration (Figure 36) . He shared different ways to visualize presence—from simple video-like representations of arms to full VR avatars, and even more abstract forms that embody expressions or emotional states [109]. These different representations have really opened up some exciting conversations about how we communicate non-verbally in digital spaces. Then, when it comes to MR, his team has been focusing a lot on visual annotations [29]. These marks can overlay onto the real world and help guide or instruct in ways that are persistent and useful in real-time collaboration.

He also talked about what we might learn from video games [26, 91], especially how they use environmental cues or first-person perspectives to communicate directly with the player. These are cues that only the player sees, but they're incredibly powerful in guiding actions and understanding. That feels like a huge opportunity for remote collaboration too. Overall, his goal for the seminar was to connect with others who are working in these areas, especially to find people interested in joining projects, and to enjoy deep discussions.

Yalong Yang, Georgia Tech

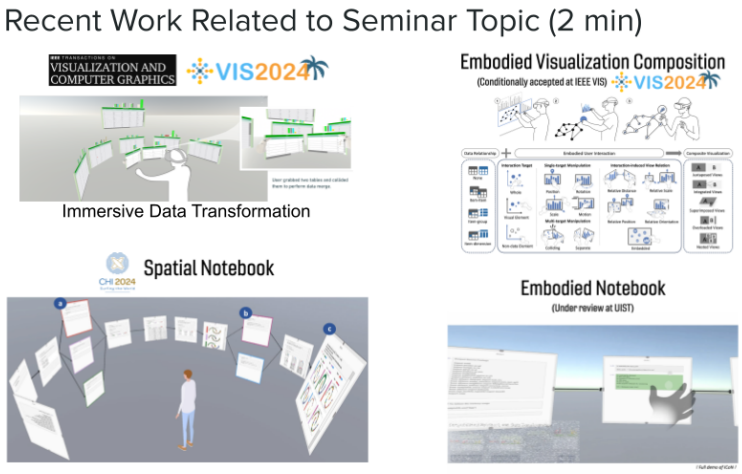

Figure 37: Yalong Yang's recent work related to the seminar.

Yang introduced research interests related to future workspaces.

His recent work related (Figure 37) to seminar topic includes immersive data transformation (to appear, TVCG), embodied visualization composition [120], spatial notebooks[50], multi-window cross-reality collaboration (to appear, ACM ISS 2024), and hybrid visualization for analytics [119].

His goals and hopes for this seminar included understanding the challenges in current practices, brainstorming innovative solutions with emerging technologies, and identify the challenges, considering evaluation metrics, discussing "real" use cases (in what scenarios VR/AR could work well, and why? Also, the converse), and forming collaborations.

Jian Zhao, University of Waterloo

Figure 38: Jian Zhao's recent work related to the seminar.

Zhao's research interests include mew visualization tools that support programmers' workflows, new AI-infused systems that augment design and UX, and new interaction techniques.

His recent work (Figure 38) related to seminar topic includes CoPrompt [31], Piet [99], EmoWear [2], and VisConductor [30].

His goals and hopes for this seminar included learning about new perspectives on presentation and collaboration techniques, deriving future agendas for relevant research topics, discussing how LLMs/ GenAI impact related approaches/ problems, building connections and collaborations, and planning the next MERCADO workshop.

Emerging Challenges in Data-Rich Multimodal Remote Collaboration

We identified 19 emerging challenges across four categories: (1) Technology, Tools, and Techniques; (2) People; (3) Artificial Intelligence (AI); and (4) Evaluation.

As of writing, we plan to discuss these challenges in a research publication affiliated either with the ACM SIGCHI conferences (CHI, UIST, ISS, CSCW) or the IEEE Visualization and Graphics Technical Community (VGTC, including VIS and PacificVis).

Classifying Technology, Tools, & Techniques (T3) for Data-Rich Multimodal Remote Collaboration

Lead contributor: Tim Dwyer

C1: What advances in technology / tools are required to make distributed and hybrid collaboration "as good as" fully face-to-face?

Writing leads: Tim Dwyer, Wolfgang Buschel, Bongshin Lee, Takayuki Itoh.

It is clear that, as of 2024, while technologies that can immerse participants in remote synchronous collaborative meetings are steadily improving, there remains a significant gap in experience between face-to-face discussion and collaboration and the best available technology-mediated remote experience. We reflected in our discussions that evidence for this disparity is played out by thousands around the world everyday who choose to travel—at significant cost and inconvenience—to work together rather than collaborate online. This consideration was front of mind as in-person participants from around the world at this particular seminar. We identified a number of questions that we believe represent fundamental challenges in termss of the Technology, Tools and Techniques (T3) for our research community. Namely: (1) *What is the minimum bar for success for distributed collaboration?*; (2) *What is the ideal - e.g. "superpower" techniques that go beyond face-to-face?*; (3) *What is the gap between the remote group and in-person group?*; (4) *What are "engineering challenges" (e.g. robustness of streaming audio / video) versus the real gaps (e.g. feature/conceptual gaps)?*

C2: What are the affordances of different T3 that support collaboration and/or presentation?

Writing leads: Tim Dwyer, Brian Smith, Harald Reiterer, Takayuki Itoh

In considering the differences in capabilities and modalities provided by various T3 it became clear that some are more suitable than others for various types of collaboration in different circumstances. In other words, different types of collaboration require different types of affordances. We this discussed *topologies* for collaboration, i.e. the relationships between people and joint activity. We considered symmetric and asymmetric topologies, i.e. all participants having equivalent (symmetric) affordances for collaboration versus having asymmetric capability, such as when one person has privileged control, for example as a presenter. But there is nuance. We asked: *When does presentation become collaboration?*; *How do we model the fluid roles of participants during collaborative*

data analysis?; *Which theories/frameworks from CSCW research are applicable to collaborative data sensemaking techniques?*.

C3: Interoperability across devices, modalities, and data

Writing leads: Brian Smith, Wolfgang Buschel, Bonsghin Lee, Masahiko Itoh

We drilled into specific combinations of hardware, modalities, and the specific needs of data visualisation for different types of data in collaborative settings. We considered devices and how they can adapt, for example: transformation of visualisation from 2D to 3D, or 3D to 2D. We wondered how different support for video and audio could still deliver inclusivity (e.g. VR, desktop or mobile). Finally, we considered tool interoperability: flexibility versus security; business concerns: market exclusivity versus compatibility; and Granularity - *what are the fundamental building blocks of collaborative data visualisation?*.

People and Data-Rich Multimodal Remote Collaboration

Lead contributor: David Saffo

C4: Scale, or, supporting all members in the current group no matter how big or small)

Contributors: David Saffo, Bektur Ryskeldiev, Samuel Huron.

The modality, form, and structure of a presentation or collaboration will change drastically depending on the number of participants. Designing tools with the target scale in mind is essential to ensure they effectively meet the needs of specific group sizes. This means creating interfaces and functionalities tailored to either intimate teams or large audiences, optimizing user experience accordingly. By focusing on the intended scale, tools can provide the right balance of features and simplicity, ensuring that both collaborative and presentation experiences are engaging and efficient for the intended number of participants. To achieve this we need to consider what technology, interfaces, and content modalities are appropriate for the target scale and context. For example, while immersive head-mounted displays can pose benefits for small to medium groups, it would be hard to scale that experience up to a larger audience since these devices are still often niche and not widely adopted by a general audience. In these scenarios, researchers must find other ways to engage large audiences in multi-modal collaboration and presentation.

C5: Dynamic roles (presenter, audiences, and future roles)

Contributors: Jon Schwabish, Bektur Ryskeldiev, Samuel Huron.

User roles during collaboration and presentations are rarely static, supporting dynamic roles is crucial for ensuring successful and efficient collaboration. A presentation of collaboration scenario often requires participants to adopt different roles targeting a specific task, such as presentor, editor, facilitator, or note-taker. Supporting flexibility in user roles allows participants to transition between roles such as presenter, audience, and collaborator. By supporting dynamic roles, tools empower users to lead discussions, contribute insights, or absorb information as needed, enhancing engagement and productivity. This

adaptability also prepares teams and audiences for future needs, as roles can evolve with the project's demands and the participants' skills.

C6: Agency / Supporting people agency and agency awareness

Contributors: Sheelagh Carpendale, Yalong Yang.

Supporting agency in both presentation and collaboration means empowering individuals to make informed decisions and actively contribute to the collective effort. Tools should enhance users' ability to express ideas, access necessary information, and understand their impact within the group or audience. Agency awareness involves recognizing each participant's contributions and ensuring that their voices are heard and valued. By fostering a sense of ownership and responsibility, platforms can motivate participants and audiences to engage more deeply and creatively.

C7: Narrative, progression, componentization (supporting varying narrative structure or format)

Contributors: Yalong Yang, David Saffo, Samuel Huron, Sheelagh Carpendale.

Successful tools for remote collaboration and presentation support diverse narrative structures and progression paths, allowing individuals to tailor content for presentations or collaborations to specific needs. Componentization involves breaking down tasks or content into manageable parts, enabling participants to focus on individual elements while maintaining a coherent overall narrative. This approach facilitates clarity and organization, helping teams and presenters track progress and adapt to changes. By accommodating various narrative formats, tools can support creativity and innovation, ensuring that the intended story unfolds in alignment with goals.

C8: Transparency, or, reducing (the level of indirection/degree of separation) between the person and the data (making the tool less intrusive), or, immediacy

Contributors: Sheelagh Carpendale, Samuel Huron.

Transparency in tools reduces the separation between users and the data they interact with, making both collaboration and presentation experiences more immediate and trustworthy. By allowing audiences to directly engage with the presented data, we can lower the level of indirection of tools, allowing participants and audiences to investigate the validity of the content in real time. Transparency involves providing clear, direct access to information and processes, enhancing understanding and efficiency. This immediacy fosters trust and confidence, encouraging more effective and engaged participation in both collaborative and presentation efforts.

C9: Accessibility, or, input vs output (how much of the overlap is there between input and output modalities?)

Contributors: Jon Schwabish, Bektur Ryskeldiev, Alark Joshi, Yasayuki Sumi.

Given the multimodal nature of remote collaboration and visualization systems, accessibility limitations can present a significant challenge for users with

visual, auditory, mobility, and cognitive impairments. The first challenge in creating accessible remote collaboration systems would be to consider the limitations and overlap in input and output modalities. For example, while users with visual and auditory impairments can experience difficulties in receiving information (an input modality for the user), they can communicate with other users through speech (an output modality). Thus, we believe that an accessible remote collaboration and visualization system should be able to balance the available communication and display methods available to users.

C10: Inclusivity, or, direct vs indirect solutions (using technology to solve a specific problem vs making environment more inclusive)

Contributors: Jon Schwabish, Bektur Ryskeldiev, Alark Joshi, Brian Smith.

Similarly to the previous challenge, a system should consider whether accessibility solutions should be direct or indirect. Direct solutions include accessibility tools designed to solve a particular problem, for a specific group of users, whereas an indirect solution is represented by guidelines and standards that would raise the inclusivity system as a whole for all the users of the system. Examples of direct solutions include sensory substitution tools, such as screen readers, while indirect solutions include guidelines and recommendations for data chart accessibility.

AI for Data-Rich Multimodal Remote Collaboration

Lead contributor: Tony Tang

C11: Interaction paradigms, or, design guidelines for AI as collaborators, mediators, and assistants

Contributors: Gabriela Molina León, Mahmood Jasim, Tony Tang, Christophe Hurter, Jian Zhao, Kiyoshi Kiyokawa, Zhu-Tian Chen, Ryo Suzuki

AI has the potential to enhance collaboration, but we need to understand and trust AI tools to make good use of them. We need to comprehend how models work and be able to predict their behaviour (e.g., make assumptions). Simultaneously, we want models to understand us better. That includes not only natural language but also the embodied context of human collaborators. Otherwise, we risk running into human-AI miscommunication, adding additional obstacles to the interaction, and getting stuck with unpredictable responses. Therefore, we need to develop interaction paradigms and design guidelines on how AI models can best support us as collaborators, mediators, and assistants. Such principles will help us create effective communication models that correctly reflect AI functionality and behaviour, and that actively learn from the human context to improve their performance.

C12: Interaction provenance, or, representing, understanding, and using shared interaction history

Contributors: Tony Tang, Mahmood Jasim, Jian Zhao, Andrew Cunningham, Zhu-Tian Chen

AI systems need to keep track of past interactions to support continuity and context in collaboration. A key aspect is how this history is presented and made visible to participants and the AI system itself, ensuring transparency in what the AI "remembers" or considers relevant. This is pressing because without proper representation of interaction provenance, AI systems may fail to integrate previous actions and insights into ongoing collaboration. This lack of continuity risks disjointed communication, inefficient task progression, and diminished trust in the AI's role. To address this, it is essential to identify ways to represent the system's understanding of shared interaction clearly, enabling both human collaborators and AI systems to make informed decisions based on a mutual awareness of past interactions.

C13: Reliability, or reliability and alignment contextual expectations in collaborative interactions

Contributors: Christophe Hurter, Mahmood Jasim, Bektur Ryskeldiev, Gabriela Molina León, Tony Tang.

AI always processes data in the context of the information that the system has and has learned to work with. Even though generative AI today tends to demonstrate that machines can be creative, context and its understanding remain significant limitations in achieving efficient alignment between user expectations and machine capabilities. While such limitations exist, they can be mitigated through tight and continuous collaboration between humans and machines. In this way, a dialogue can be initiated to enrich the shared understanding of the context and the tasks to be performed. Overall, this human-machine interaction plays a major role in the notion of trust, which is currently asymmetrical across two aspects. The only trust that exists is the trust that humans place in machine responses. The reverse trust—how much machines can rely on user requests—is still barely investigated. Additionally, the trust users have in machines is often fragile and asymmetrical; a machine can easily lose user trust after a single failure. Regaining trust is challenging and may require time or extensive dialogue to explain the machine's rationale. These aspects of human-machine team dynamics and trust are essential for effective collaboration with the use of AI tools.

C14: Sustainability, or matching energy consumption to tempo of interaction

Contributors: Andrew Cunningham, Christophe Hurter, Mahmood Jasim, Ryo Suzuki, Zhu-Tian Chen, Gabriela Molina León

The growing use of AI in multimodal remote collaboration brings with it significant computational demands, which in turn impact energy consumption. This challenge revolves around balancing the need for real-time responsiveness with sustainable AI processing. In collaboration environments, the AI must adapt its tempo to match human interaction, ensuring that it is neither too fast (over-processing) nor too slow (hindering the natural flow of communication). At the same time, the environmental footprint of compute-intensive models must be taken into account, particularly when these systems are used frequently over long periods. Thoughtful choices about AI models—taking into account task-specific needs like speed, performance, and sustainability—are essential. Models should be optimized to match the context of use, ensuring that energy

consumption is appropriate and that AI processing aligns with the demands of human interaction, without compromising either efficiency or environmental responsibility.

C15: Privacy, or balancing opportunities of personalization with the challenges of privacy and data collection + VI

Contributors: Ryo Suzuki, Andrew Cunningham, Kiyoshi Kiyokawa, Zhu-Tian Chen, Mahmood Jasim, Christophe Hurter.

Personalization in AI has the potential to significantly enhance collaboration by tailoring experiences to individual preferences, behaviors, and needs. However, this brings challenges related to privacy, data collection, and social acceptance. The more personalized an AI system becomes, the more data it requires, raising questions about how much information individuals are willing to share, especially in sensitive or high-stakes collaborative environments. Furthermore, there is a fine balance between providing personalized assistance and maintaining fidelity to the original intentions or actions of the human collaborators. Systems must also be designed to be inclusive, acknowledging and respecting diversity while avoiding intrusive or excessive data collection. AI must navigate the tension between offering rich, personalized experiences and ensuring that participants feel their data and privacy are protected. Social acceptability becomes key, particularly in how the AI's behavior and presence are perceived by everyone in the collaboration space, ensuring it remains respectful and non-invasive to all participants involved.

Evaluating Data-Rich Multimodal Remote Collaboration Experiences

Lead contributor: Arnaud Prouzeau

C16: Expanding the scope of evaluating users to group specificity

Contributors: Samuel Huron, Lyn Bartram, Maxime Cordeil, Arnaud Prouzeau

When evaluating multimodal remote collaborative systems, expanding the scope from individual users to account for group-specific factors presents a significant challenge. Group dynamics play a pivotal role, with hierarchy, personality differences, and the representativity of the group's social or professional background all affecting how participants interact with one another, which in turn impacts performance during task completion in the evaluation. Additionally, the context of use must be carefully considered; evaluations should take place in settings that reflect real-world conditions to ensure accurate and consistent results. For example, conducting a study in a co-located setting (to reduce technical complexity) for a system primarily used in remote settings will fail to generate the appropriate insights. Finally, in laboratory settings, most systems are evaluated in isolation. However, in the real world, systems are part of a broader ecosystem of applications, where users may communicate via other tools (backchannel communication) or share their screens through separate applications, deploying communication mechanisms that were not assessed during the evaluation.

C17: Collaborative study design + getting the right research questions

Contributors: Lyn Bartram, Samuel Huron, Maxime Cordeil, Arnaud Prouzeau

Designing a collaborative study presents a complex challenge that begins with identifying the primary motivation: is the goal to evaluate the system itself or to understand the nature of collaboration? This distinction not only impacts what is measured—whether it's task performance or collaboration effectiveness—but also influences the choice of methods, such as empirical, observational, participatory, or co-design approaches. Establishing an appropriate baseline is equally crucial. Should the study compare against face-to-face interactions, where participants are highly experienced, or against a common digital platform like Zoom? Another important consideration is group size: is studying pairs sufficient? While involving more participants can be complicated, as it requires recruiting more people overall, most ecological settings typically involve more users collaborating simultaneously. Additionally, selecting participants with the appropriate expertise can be a significant challenge. Finally, metrics to measure key aspects like workspace awareness or the sense of presence often rely on questionnaires, which, while sometimes standardized, are not always reliable. We believe balancing these quantitative measures with qualitative observations is essential for a more comprehensive evaluation.

C18: Collaborative study logistics

Contributors: Maxime Cordeil, Lyn Bartram, Samuel Huron, Arnaud Prouzeau

Conducting such studies remotely presents unique logistical difficulties. As multiple devices are introduced into the study, the complexity increases, leading to potential technical issues that can affect both data capture and user experience. Furthermore, participants may be spread across different time zones, as is increasingly the case with global teams of analysts, making coordination challenging, with users available at different times ('everywhere and everywhen'). Capturing meaningful data also poses limitations: some systems cannot be instrumented for detailed custom monitoring (e.g., commercial tools like Zoom or Slack do not provide access to log data). Additionally, methods like 'think aloud,' which are valuable in traditional usability studies, may not be feasible in a distributed, remote setting, as they require an experimenter to be present with each participant and can interfere with the oral communication between participants. The remote setting and the difficulty of having an experimenter on-site with each participant also mean that certain elements of collaboration, such as facial expressions or deictic gestures, may be missed. These challenges make it difficult to gather the rich, real-time data often needed for evaluating collaborative systems.

C19: Collaborative study data analysis

Contributors: Arnaud Prouzeau, Maxime Cordeil, Lyn Bartram, Samuel Huron

Analyzing data from studies on collaborative remote systems presents significant challenges. Applying different frameworks for analysis, such as qualitative, quantitative, or mixed-method approaches, can provide distinct insights but also add complexity to the process of interpreting findings cohesively. One critical issue is constructing meaning from the data and reporting results. Studies

often yield both individual-level data (from participants within the group) and group-level dynamics, but merging these two perspectives doesn't guarantee a comprehensive understanding of the collaborative processes at play. This 'completeness problem' arises because insights from separate groups don't necessarily reflect the full scope of collaboration across different settings. Additionally, comparing results across different contexts is another challenge. Variability in remote setups, participant expertise, technological tools, and task complexity can make it difficult to generalize findings, as results from one context may not directly translate to another. Finally, the volume and diversity of data generated in these studies—from interaction logs to qualitative feedback—can overwhelm traditional analysis techniques. This makes it difficult to identify overarching patterns while, at the same time, complicating the detection of often subtle, weak signals of collaboration that may be spread across different modalities or interactions.

References

[1] M. Aldugom, K. Fenn, and S. W. Cook. Gesture during math instruction specifically benefits learners with high visuospatial working memory capacity. *Cognitive Research: Principles and Implications*, 5(1), 2020. doi: 10.1186/s41235-020-00215-8

[2] P. An, J. S. Zhu, Z. Zhang, Y. Yin, Q. Ma, C. Yan, L. Du, and J. Zhao. EmoWear: Exploring emotional teasers for voice message interaction on smartwatches. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2024. doi: 10.1145/3613904.3642101

[3] B. Bach, M. Keck, F. Rajabiyazdi, T. Losev, I. Meirelles, J. Dykes, R. S. Laramee, M. AlKadi, C. Stoiber, S. Huron, C. Perin, L. Morais, W. Aigner, D. Kosminsky, M. Boucher, S. Knudsen, A. Manataki, J. Aerts, U. Hinrichs, J. C. Roberts, and S. Carpendale. Challenges and opportunities in data visualization education: A call to action. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 30(1), 2024. doi: 10.1109/TVCG.2023.3327378

[4] T. Blascheck, F. Bentley, E. K. Choe, T. Horak, and P. Isenberg. Characterizing glanceable visualizations: From perception to behavior change. In *Mobile Data Visualization*. CRC Press, 2021.

[5] E. Bonnail, W.-J. Tseng, M. Mcgill, E. Lecolinet, S. Huron, and J. Gugenheimer. Memory manipulations in extended reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2023. doi: 10.1145/3544548.3580988

[6] M. Brehmer. The information in our hands, 2021. Information+ 2021 conference presentation. https://vimeo.com/592591860.

[7] M. Brehmer, M. Cordeil, C. Hurter, and T. Itoh. The MERCADO Workshop at IEEE VIS 2023: Multimodal experiences for remote communication around data online, 2023. Workshop at IEEE VIS 2023. arxiv.org/abs/2303.11825.

[8] M. Brehmer and R. Kosara. From jam session to recital: Synchronous communication and collaboration around data in organizations. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VIS)*, 28(1), 2022. doi: 10.1109/TVCG.2021.3114760

[9] N. Bressa, J. Louis, W. Willett, and S. Huron. Input visualization: collecting and modifying data with visual representations. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2024. doi: 10.1145/3613904.3642808

[10] W. Büschel, A. Lehmann, and R. Dachselt. Miria: A mixed reality toolkit for the in-situ visualization and analysis of spatio-temporal interaction data. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi: 10.1145/3411764.3445651

[11] S. Butscher, S. Hubenschmid, J. Müller, J. Fuchs, and H. Reiterer. Clusters, trends, and outliers: How immersive technologies can facilitate the collaborative analysis of multidimensional data. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2018. doi: 10.1145/3173574.3173664

[12] W. Buxton. Ubiety, mobility & transitions: The quest for ecological intelligence. In *Keynote presentation of the ACM International Conference on Interactive Surfaces and Spaces (ISS)*, 2019. doi: 10.1145/3343055.3359724

[13] P. Cash and A. Maier. Prototyping with your hands: The many roles of gesture in the communication of design concepts. *Journal of Engineering Design*, 27(1-3), 2016. doi: 10.1080/09544828.2015.1126702

[14] S. W. Chan, B. Ryskeldiev, and S. Nanayakkara. DeclutterAR: Mobile diminished reality and augmented reality to address hoarding by motivating decluttering and selling on online marketplace. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2022. doi: 10.1109/ISMAR-Adjunct57072.2022.00187

[15] Z.-T. Chen, D. Chiappalupi, T. Lin, Y. Yang, J. Beyer, and H. Pfister. RL-LABEL: A deep reinforcement learning approach intended for AR label placement in dynamic scenarios. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 2023. doi: 10.1109/TVCG.2023.3326568

[16] Z.-T. Chen, Y. Su, Y. Wang, Q. Wang, H. Qu, and Y. Wu. MARVisT: Authoring glyph-based visualization in mobile augmented reality. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 26(8), 2020. doi: 10.1109/TVCG.2019.2892415

[17] Z.-T. Chen, Q. Yang, J. Shan, T. Lin, J. Beyer, H. Xia, and H. Pfister. iball: Augmenting basketball videos with gaze-moderated embedded visualizations. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–18, 2023.

[18] Z.-T. Chen, Q. Yang, X. Xie, J. Beyer, H. Xia, Y. Wu, and H. Pfister. Sporthesia: Augmenting sports videos using natural language. *IEEE transactions on visualization and computer graphics*, 29(1):918–928, 2022.

[19] Z.-T. Chen, S. Ye, X. Chu, H. Xia, H. Zhang, H. Qu, and Y. Wu. Augmenting sports videos with viscommentator. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VIS)*, 28(1), 2021. doi: 10.1109/TVCG.2021.3114806

[20] N. Chulpongsatorn, M. S. Lunding, N. Soni, and R. Suzuki. Augmented Math: Authoring ar-based explorable explanations by augmenting static math textbooks. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2023. doi: 10.1145/3586183.3606827

[21] J. J. Y. Chung, H. V. Shin, H. Xia, L.-y. Wei, and R. H. Kazi. Beyond show of hands: Engage viewers via expressive and scalable visual communication in live stream. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi: 10.1145/3411764.3445419

[22] Cisco Webex Meetings, 2022. https://www.webex.com/meetings.html.

[23] J. S. Clarke, J. P. Cornelissen, and M. P. Healey. Actions speak louder than words: How figurative language and gesturing in entrepreneurial pitches influences investment judgments. *Academy of Management Journal*, 62, 2019. doi: 10.5465/amj.2016.1008

[24] M. Cordeil, T. Billy, N. Mellado, L. Barthe, N. Couture, and P. Reuter. ImmersiveIML–immersive interactive machine learning for 3D point cloud classification: the neural network at your fingertips. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR)*, 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00025

[25] F. Daneshzand, C. Perin, and S. Carpendale. Kiriphys: Exploring new data physicalization opportunities. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 29(1), 2022. doi: 10.1109/TVCG.2022.3209365

[26] K. R. Dillman, T. T. H. Mok, A. Tang, L. Oehlberg, and A. Mithcell. A visual interaction cue framework from video game environments for augmented reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2018. doi: 10.1145/3173574.3173714

[27] S. Drucker, S. Huron, R. Kosara, J. Schwabish, and N. Diakopoulos. Communicating data to an audience. In N. Henry Riche, C. Hurter, N. Diakopoulos, and S. Carpendale, eds., *Data-Driven Storytelling*. CRC Press, 2018. https://tinyurl.com/drucker2018.

[28] B. Ens, B. Bach, M. Cordeil, U. Engelke, M. Serrano, W. Willett, A. Prouzeau, C. Anthes, W. Büschel, C. Dunne, T. Dwyer, J. Grubert, J. H. Haga, N. Kirshenbaum, D. Kobayashi, T. Lin, M. Olaosebikan, F. Pointecker, D. Saffo, N. Saquib, D. Schmalstieg, D. A. Szafir, M. Whitlock, and Y. Yang. Grand challenges in immersive analytics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi: 10.1145/3411764.3446866

[29] O. Fakourfar, K. Ta, R. Tang, S. Bateman, and A. Tang. Stabilized annotations for mobile remote assistance. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2016. doi: 10.1145/2858036.2858171

[30] T. Femi-Gege, M. Brehmer, and J. Zhao. VisConductor: Affect-varying widgets for animated data storytelling in gesture-aware augmented video presentation. In *Proceedings of the ACM International Conference on Interactive Surfaces and Spaces (ISS)*, 2024. arxiv.org/abs/2406.17986.

[31] L. Feng, R. Yen, Y. You, M. Fan, J. Zhao, and Z. Lu. CoPrompt: Supporting prompt sharing and referring in collaborative natural language programming. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2024. doi: 10.1145/3613904.3642212

[32] D. I. Fink, J. Zagermann, H. Reiterer, and H.-C. Jetter. Re-locations: Augmenting personal and shared workspaces to support remote collaboration in incongruent spaces. *Proceedings of the ACM on Human-Computer Interaction (ISS)*, 6, 2022. doi: 10.1145/3567709

[33] E. E. Firat, A. Joshi, and R. S. Laramee. Interactive visualization literacy: The state-of-the-art. *Information Visualization*, 21(3), 2022. doi: 10.1177/14738716221081831

[34] E. E. Firat, C. Lang, B. Srinivas, I. Peng, R. S. Laramee, and A. Joshi. A constructivism-based approach to treemap literacy in the classroom. *Computer Graphics Forum (Proceedings of EuroVis)*, 42(2), 2023. doi: 10.2312/eged.20231016

[35] Flow Immersive, 2022. https://flowimmersive.com.

[36] R. Fujikura and Y. Sumi. Facilitating experiential knowledge sharing through situated conversations. In *Proceedings of the ACM Augmented Humans International Conference*, 2020. doi: 10.1145/3384657.3384798

[37] P. S. Games and A. Joshi. An evaluation-guided approach for effective data visualization on tablets. In *Visualization and Data Analysis*, vol. 9397, 2015. doi: 10.1117/12.2076523

[38] E. J. Gonzalez, K. Patel, K. Ahuja, and M. Gonzalez-Franco. XDTK: A cross-device toolkit for input & interaction in XR. In *Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 2024. doi: 10.1109/VRW62533.2024.00092

[39] Google Meet, 2022. https://meet.google.com/.

[40] M. Gottsacker, M. Chen, D. Saffo, F. Lu, and B. MacIntyre. Asymmetric immersive presentation system for financial data visualization. In *Proceedings of MERCADO Workshop at IEEE VIS 2023: Multimodal Experiences for Remote Communication Around Data Online*, 2023. osf.io/cvqg9.

[41] M. Gottsacker, M. Chen, D. Saffo, F. Lu, and B. MacIntyre. Hybrid user interface for audience feedback guided asymmetric immersive presentation of financial data. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00046

[42] J. E. Grønbæk, B. Saatçi, C. F. Griggio, and C. N. Klokmose. MirrorBlender: Supporting hybrid meetings with a malleable video-conferencing system. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi: 10.1145/3411764.3445698

[43] B. D. Hall, L. Bartram, and M. Brehmer. Augmented chironomia for presenting data to remote audiences. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2022. doi: 10.1145/3526113.3545614

[44] T. Horak, W. Aigner, M. Brehmer, A. Joshi, and C. Tominski. Responsive visualization design for mobile devices. In *Mobile Data Visualization*, pp. 33–66. CRC Press, 2021.

[45] S. Hubenschmid, J. Wieland, D. I. Fink, A. Batch, J. Zagermann, N. Elmqvist, and H. Reiterer. Relive: Bridging in-situ and ex-situ visual analytics for analyzing mixed reality user studies. In *Proceedings of*

the ACM Conference on Human Factors in Computing Systems (CHI), 2022. doi: 10.1145/3491102.3517550

[46] S. Hubenschmid, J. Zagermann, S. Butscher, and H. Reiterer. Stream: Exploring the combination of spatially-aware tablets with augmented reality head-mounted displays for immersive analytics. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi: 10.1145/3411764.3445298

[47] S. Huron, T. Nagel, L. Oehlberg, and W. Willett. *Making with Data: Physical Design and Craft in a Data-Driven World*. CRC Press, 2022.

[48] C. Hurter, B. Rogowitz, G. Truong, T. Andry, H. Romat, L. Gardy, F. Amini, and N. H. Riche. Memory recall for data visualizations in mixed reality, virtual reality, 3D and 2D. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 2024. doi: 10.1109/TVCG.2023.3336588

[49] K. Ihara, M. Faridan, A. Ichikawa, I. Kawaguchi, and R. Suzuki. HoloBots: Augmenting holographic telepresence with mobile robots for tangible remote collaboration in mixed reality. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2023. doi: 10.1145/3586183.3606727

[50] S. In, E. Krokos, K. Whitley, C. North, and Y. Yang. Evaluating navigation and comparison performance of computational notebooks on desktop and in virtual reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2024. doi: 10.1145/3613904.3642932

[51] T. Itoh and K. Klein. Key-node-separated graph clustering and layout for human relationship graph visualization. *IEEE Computer Graphics and Applications*, 35(6), 2015.

[52] T. Itoh and K. Koyamada. Automatic isosurface propagation by using an extrema graph and sorted boundary cell lists. *IEEE Transactions on Visualization and Computer Graphics*, 1(4), 1995.

[53] T. Itoh, A. Kumar, K. Klein, and J. Kim. High-dimensional data visualization by interactive construction of low-dimensional parallel coordinate plots. *IEEE Transactions on Visualization and Computer Graphics*, 26(6), 2020.

[54] T. Itoh, C. Muelder, K.-L. Ma, and J. Sese. A hybrid space-filling and force-directed layout method for visualizing multiple-category graphs. In *Proceedings of IEEE Pacific Visualization Symposium*, 2009.

[55] T. Itoh, H. Takakura, A. Sawada, and K. Koyamada. Hierarchical visualization of network intrusion detection data in the ip address space. *IEEE Computer Graphics and Applications*, 26(2), 2009.

[56] T. Itoh, Y. Yamaguchi, Y. Ikehata, and Y. Kajinaga. Hierarchical data visualization using a fast rectangle-packing algorithm. *IEEE Transactions on Visualization and Computer Graphics*, 10(3), 2004.

[57] T. Itoh, Y. Yamaguchi, and K. Koyamada. Fast isusurface generation using the volume thinning algorithm. *IEEE Transactions on Visualization and Computer Graphics*, 7(1), 2001.

[58] M. Jasim, A. Sarvghad, and N. Mahyar. CommunityClick-Virtual: Multi-modal interactions for enhancing participation in virtual meetings. In *Proceedings of MERCADO Workshop at IEEE VIS 2023: Multimodal Experiences for Remote Communication Around Data Online*, 2023. osf.io/preprints/osf/szep3.

[59] H. Kaimoto, K. Monteiro, M. Faridan, J. Li, S. Farajian, Y. Kakehi, K. Nakagaki, and R. Suzuki. Sketched Reality: Sketching bi-directional interactions between virtual and physical worlds with AR and actuated tangible UI. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2022. doi: 10.1145/3526113.3545626

[60] Y.-H. Kim, B. Lee, A. Srinivasan, and E. K. Choe. Data@ hand: Fostering visual exploration of personal data on smartphones leveraging speech and touch interaction. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2021. doi: 10.1145/3411764.3445421

[61] A. King. From sage on the stage to guide on the side. *College Teaching*, 41(1), 1993.

[62] R. Kitagawa and T. Itoh. Visualization of swiping motion of competitive karuta using 3d bone display. In *Proceedings of the IEEE International Conference Information Visualisation (IV)*, 2023. doi: 10.1109/IV60283.2023.00065

[63] A. Kristanto, M. Cordeil, B. Tag, N. H. Riche, and T. Dwyer. Hanstreamer: An open-source webcam-based live data presentation system. In *Proceedings of MERCADO Workshop at IEEE VIS 2023: Multimodal Experiences for Remote Communication Around Data Online*, 2023. arxiv.org/abs/2309.12538.

[64] K. Krug, W. Büschel, K. Klamka, and R. Dachselt. Clear sight: Exploring the potential of interacting with transparent tablets in augmented reality. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2022. doi: 10.1109/ISMAR55827.2022.00034

[65] J. Le Meudec, A. Bezerianos, C. Mercier, and A. Prouzeau. Learning with immersive technologies: a VR task to study collaborative learning. In *IHM 2024-35e Conférence Internationale Francophone sur l'Interaction Humain-Machine*, 2024. hal.science/hal-04487309v1.

[66] B. Lee, M. Brehmer, P. Isenberg, E. K. Choe, R. Langner, and R. Dachselt. Data visualization on mobile devices. In *Extended Abstract Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2018. doi: 10.1145/3170427.3170631

[67] B. Lee, D. Brown, B. Lee, C. Hurter, S. Drucker, and T. Dwyer. Data visceralization: Enabling deeper understanding of data using virtual reality. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)*, 27(2), 2021. doi: 10.1109/TVCG.2020.3030435

[68] B. Lee, A. Srinivasan, P. Isenberg, J. Stasko, et al. Post-wimp interaction for information visualization. *Foundations and Trends® in Human–Computer Interaction*, 14(1):1–95, 2021.

[69] G. M. León, A. Bezerianos, O. Gladin, and P. Isenberg. Talk to the wall: The role of speech interaction in collaborative visual analytics. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VIS)*, 2024. doi: 10.1109/TVCG.2024.3456335

[70] G. M. León, P. Isenberg, and A. Breiter. Eliciting multimodal and collaborative interactions for data exploration on large vertical displays. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 30(2), 2023. doi: 10.1109/TVCG.2023.3323150

[71] G. M. Léon, P. Isenberg, and A. Breiter. Talking to data visualizations: Opportunities and challenges. In *Proceedings of MERCADO Workshop at IEEE VIS 2023: Multimodal Experiences for Remote Communication Around Data Online*, 2023. arxiv.org/abs/2309.09781.

[72] M. Levy, S. Gruber, Z. Wang, and S. Carpendale. From science to story: Communicating permafrost concepts with data comics. In *Proceedings of the International Conference on Permafrost*, 2024. doi: 10.52381/ICOP2024.157.1

[73] J. Liao, A. Karim, S. S. Jadon, R. H. Kazi, and R. Suzuki. RealityTalk: Real-time speech-driven augmented presentation for ar live storytelling. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2022. doi: 10.1145/3526113.3545702

[74] Z. Liu, T. Itoh, J. Q. Dawson, and T. Munzner. The sprawlter graph readability metric:combining sprawl and area-aware clutter. *IEEE Transactions on Visualization and Computer Graphics*, 26(6), 2020.

[75] W. Luo, Z. Yu, R. Rzayev, M. Satkowski, S. Gumhold, M. McGinity, and R. Dachselt. PEARL: Physical environment based augmented reality lenses for in-situ human movement analysis. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2023. doi: 10.1145/3544548.3580715

[76] N. Mahyar and M. Tory. Supporting communication and coordination in collaborative sensemaking. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VAST)*, 20(12), 2014. doi: 10.1109/TVCG.2014.2346573

[77] D. Matsumoto, M. G. Frank, and H. S. Hwang. *Nonverbal communication: Science and applications: Science and Applications*. SAGE, 2013.

[78] K. Matsumura and Y. Sumi. What are you talking about while driving? an analysis of in-car conversations aimed at conversation sharing. In *Proceedings of the International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 2014. doi: 10.1145/2667317.2667417

[79] J. McDade, A. Drogemuller, A. Jing, N. Ireland, J. Walsh, B. Thomas, W. Mayer, and A. Cunningham. Cadet: A collaborative agile data exploration tool for mixed reality. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2022. doi: 10.1109/ISMAR-Adjunct57072.2022.00195

[80] Microsoft Teams, 2022. https://www.microsoft.com/en-us/microsoft-teams/group-chat-software.

[81] mmhmm, 2022. https://mmhmm.app.

[82] R. Neogy, J. Zong, and A. Satyanarayan. Representing real-time multi-user collaboration in visualizations. In *Short Paper Proceedings of the IEEE Visualization Conference (VIS)*, 2020. doi: 10.1109/VIS47514.2020.00036

[83] M. Nicolae, V. Roussel, M. Koelle, S. Huron, J. Steimle, and M. Teyssier. Biohybrid devices: Prototyping interactive devices with growable materials. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2023. doi: 10.1145/3586183.3606774

[84] OBS Studio, 2022. https://obsproject.com.

[85] K. Okada, M. Yoshida, T. Itoh, T. Czauderna, and K. Stephens. VR system for spatio-temporal visualization of tweet data. In *Proceedings of the IEEE International Conference Information Visualisation (IV)*, 2018. doi: 10.1109/iV.2018.00026

[86] A. Okuno and Y. Sumi. Social activity measurement by counting faces captured in first-person view lifelogging video. In *Proceedings of the ACM Augmented Human International Conference*, 2019. doi: 10.1145/3311823.3311846

[87] J. Orlosky, M. Sra, K. Bektaş, H. Peng, J. Kim, N. Kos' myna, T. Höllerer, A. Steed, K. Kiyokawa, and K. Akşit. Telelife: The future of remote living. *Frontiers in Virtual Reality*, 2:763340, 2021. doi: 10.3389/frvir.2021.763340

[88] I. Peng, E. E. Firat, R. S. Laramee, and A. Joshi. Evaluating Bloom's taxonomy-based learning modules for parallel coordinates literacy. In *Proceedings of Eurographics Education Papers*, 2022. doi: 10.2312/eged.20221042

[89] K. Perlin, Z. He, and K. Rosenberg. Chalktalk: A visualization and communication language – as a tool in the domain of computer science education, 2018. arXiv preprint https://arxiv.org/abs/1809.07166.

[90] V. Pooryousef, M. Cordeil, L. Besançon, R. Bassed, and T. Dwyer. Collaborative forensic autopsy documentation and supervised report generation using a hybrid mixed-reality environment and generative AI. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 2024. doi: 10.1109/TVCG.2024.3456212

[91] L. Poretski and A. Tang. Press A to jump: Design strategies for video game learnability. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2022. doi: 10.1145/3491102.3517685

[92] D. Rodighiero, L. Derry, D. Duhaime, J. Kruger, M. C. Mueller, C. Pietsch, J. Steward, and J. T. Schnapp. Surprise machines: Revealing harvard art museums' image collection. *Information Design Journal*, 2022. doi: 10.1075/idj.22013.rod

[93] B. Ryskeldiev, M. Cohen, and J. Herder. Streamspace: Pervasive mixed reality telepresence for remote collaboration on mobile devices. *Journal of Information Processing*, 26, 2018. doi: 10.2197/ipsjjip.26.177

[94] B. Ryskeldiev, K. Hara, M. Kobayashi, and K. Kusano. Investigating accessibility challenges and opportunities for users with low vision disabilities in customer-to-customer (C2C) marketplaces. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS)*, 2022. doi: 10.1145/3517428.3550390

[95] N. Saquib, R. H. Kazi, L.-Y. Wei, and W. Li. Interactive body-driven graphics for augmented video performance. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2019. doi: 10.1145/3290605.3300852

[96] M. Schwab, D. Saffo, Y. Zhang, S. Sinha, C. Nita-Rotaru, J. Tompkin, C. Dunne, and M. A. Borkin. VisConnect: Distributed event synchronization for collaborative visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of InfoVis)*, 2020. doi: 10.1109/TVCG.2020.3030366

[97] J. Schwabish. *Better Presentations: A Guide for Scholars, Researchers, and Wonks*. Columbia University Press, 2016.

[98] V. Setlur. Exploring spatial computing and immersive analytics with Vision Pro, 2024. Tableau Blog. tableau.com/blog/exploring-spatial-computing-and-immersive-analytics-vision-pro.

[99] X. Shi, Y. Wang, Y. Wang, and J. Zhao. Piet: Facilitating color authoring for motion graphics video. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2024. doi: 10.1145/3613904.3642711

[100] Slack, 2022. https://slack.com/.

[101] Slack help center: Use huddles in Slack, 2022. https://slack.com/help/articles/4402059015315-Use-huddles-in-Slack.

[102] A. Srinivasan and M. Brehmer. Combining voice and gesture for presenting data to remote audiences. In *Proceedings of MERCADO Workshop at IEEE VIS 2023: Multimodal Experiences for Remote Communication Around Data Online*, 2023. arjun010.github.io/static/papers/mm-presentation-mercado23.pdf.

[103] A. Srinivasan, B. Lee, N. H. Riche, S. M. Drucker, and K. Hinckley. InChorus: Designing consistent multimodal interactions for data visualization on tablet devices. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2020. doi: 10.1145/3313831.3376782

[104] A. Srinivasan, B. Lee, and J. Stasko. Interweaving multimodal interaction with flexible unit visualizations for data exploration. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 27(8), 2020. doi: 10.1109/TVCG.2020.2978050

[105] Y. Sumi, J. Ito, and T. Nishida. Photochat: communication support system based on sharing photos and notes. In *Extended Abstract Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2008. doi: 10.1145/1358628.1358837

[106] Y. Sumi, S. Ito, T. Matsuguchi, S. Fels, S. Iwasawa, K. Mase, K. Kogure, and N. Hagita. Collaborative capturing, interpreting, and sharing of experiences. *Personal and Ubiquitous Computing*, 11, 2007. doi: 10.1007/s00779-006-0088-1

[107] Y. Sumi, M. Yano, and T. Nishida. Analysis environment of conversational structure with nonverbal multimodal data. In *Proceedings of the International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, 2010. doi: 10.1145/1891903.1891958

[108] R. Suzuki, R. H. Kazi, L.-Y. Wei, S. DiVerdi, W. Li, and D. Leithinger. RealitySketch: Embedding responsive graphics and visualizations in AR through dynamic sketching. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2020. doi: 10.1145/3379337.3415892

[109] A. Tang, M. Pahud, K. Inkpen, H. Benko, J. C. Tang, and B. Buxton. Three's company: understanding communication channels in three-way distributed collaboration. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW)*, 2010. doi: 10.1145/1718918.1718969

[110] W.-J. Tseng, E. Bonnail, M. McGill, M. Khamis, E. Lecolinet, S. Huron, and J. Gugenheimer. The dark side of perceptual manipulations in virtual reality. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI)*, 2022. doi: 10.1145/3491102.3517728

[111] VisInPractice @ IEEE VIS, 2021. ieeevis.org/year/2021/info/visinpractice.

[112] W. Willett, B. A. Aseniero, S. Carpendale, P. Dragicevic, Y. Jansen, L. Oehlberg, and P. Isenberg. Perception! Immersion! Empowerment! Superpowers as inspiration for visualization. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VIS)*, 28(1), 2021. doi: 10.1109/TVCG.2021.3114844

[113] Z. Xia, K. Monteiro, K. Van, and R. Suzuki. RealityCanvas: Augmented reality sketching for embedded and responsive scribble animation effects. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2023. doi: 10.1145/3586183.3606716

[114] Y. Yang, T. Dwyer, Z. Swiecki, B. Lee, M. Wybrow, M. Cordeil, T. Wulandari, B. H. Thomas, and M. Billinghurst. Putting our minds together: Iterative exploration for collaborative mind mapping. In *Proceedings of*

the Augmented Humans International Conference (AHs), 2024. doi: 10.1145/3652920.3653043

[115] L. Yao, F. Bucchieri, V. McArthur, A. Bezerianos, and P. Isenberg. User experience of visualizations in motion: A case study and design considerations. *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 2024. doi: 10.1109/TVCG.2024.3456319

[116] L. Yuan, W. Tong, K. Takahira, Z. Wen, Y. Yang, and H. Qu. Echoes in the gallery: A collaborative immersive analytics system for analyzing audience reactions in virtual reality exhibitions. In *Proceedings of MERCADO Workshop at IEEE VIS 2023: Multimodal Experiences for Remote Communication Around Data Online*, 2023. doi.org/10.31219/osf.io/zmyx9.

[117] S. Zeller, F. Samsel, and L. Bartram. Affective, hand-sculpted glyph forms for engaging and expressive scientific visualization. In *Proceedings of the IEEE VIS Arts Program (VISAP)*, 2022. doi: 10.1109/VISAP57411.2022.00025

[118] Z. Zhao and N. Elmqvist. The stories we tell about data: Media types for data-driven storytelling, 2022. arXiv preprint. https://arxiv.org/abs/2202.00047.

[119] X. Zhou, Y. Yang, F. Ortega, A. U. Batmaz, and B. Lee. Data-driven storytelling in hybrid immersive display environments. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00056

[120] Q. Zhu, T. Lu, S. Guo, X. Ma, and Y. Yang. CompositingVis: Exploring interactions for creating composite visualizations in immersive environments. *IEEE Transactions on Visualization and Computer Graphics (Proceedings of VIS)*, 2024. doi: 10.1109/TVCG.2024.3456210

[121] Zoom, 2022. https://zoom.us/.