Photo credit: Dave DiCello

17-803 Empirical Methods

Bogdan Vasilescu, Institute for Software Research

Exemplar Interview Studies

Thursday, February 18, 2021

“Great artists steal”

(Unclear, often misquoted: https://quoteinvestigator.com/2013/03/06/artists-steal/)

Outline for Today

- Discuss your interview guides
- Dissect exemplars of semi-structured interviews in (CS) research practice

- Administrivia:
  - Blog posts due today
  - Break day next Tuesday
  - Mandatory reading for next Thursday

Part I. Discussion of interview guide

Activity: Workshop an Interview Guide

- Goal: better understanding of how and why academic researchers collaborate on writing papers.
  - Perhaps you are studying collaboration technology in the workplace,
  - or you are a tool developer in the very early stages of trying to develop a winning collaboration technology for co-authorship,
  - or you are someone interested in improving the quality of collaboration in your field.
- Develop a short interview protocol.
  - Anticipate 15-20 minute interviews.
  - Use good technique, keeping the interviewee focused on concrete tasks.
- Be able to provide convincing rationale for all of your choices in producing the protocol – why these questions, what probes, etc.

Part II. Paper dissection mini workshop

Exemplars

- Grinter, R. E., & Palen, L. (2002). Instant messaging in teen life. In Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW) (pp. 21-30).
- Cherubini, M., Venolia, G., DeLine, R., & Ko, A. J. (2007). Let's go to the whiteboard: How and why software developers use drawings. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI) (pp. 557-566).
- Wash, R. (2010). Folk models of home computer security. In Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS) (pp. 1-16).
- Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub: Transparency and collaboration in an open software repository. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work & Social Computing (CSCW) (pp. 1277-1286).
- Tausczik, Y. R., Kittur, A., & Kraut, R. E. (2014). Collaborative problem solving: A study of Math Overflow. In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 355-367).
- Manotas, I., Bird, C., Zhang, R., Shepherd, D., Jaspan, C., Sadowski, C., Pollock, L., & Clause, J. (2016). An empirical study of practitioners' perspectives on green software engineering. In 2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE) (pp. 237-248). IEEE.
- Barwulor, C., McDonald, A., Hargittai, E., & Redmiles, E. M. (2021). "Disadvantaged in the American-dominated internet": Sex, Work, and Technology. In Proceedings of the 2021 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) (pp. 931-936).

What kind of questions do they ask?

Teen Instant Messaging Paper (Grinter & Palen, 2002)

- "Teenagers' use of Instant Messaging (IM) is on rapid rise, and has been a recent object of media attention.
- "However, little is empirically known about how and why teens use IM […] – an area rife with interesting and open research questions.
- "We believe that teenage IM adoption offers three potential insights.
  - Firstly, teenage IM adoption marks a significant entry of collaborative information technologies into the home. Studying teenagers' use of collaborative technology in the home offers new insight about its role in the domestic ecology.
  - Secondly, since most teenagers have little previous experience with technologies that convey presence between remote peers, they must learn what it means to be simultaneously private and public people.
  - Finally, teenagers are the workforce of the future, and communication habits they develop now may indicate what we can expect from them as adults.

Grinter, R. E., & Palen, L. (2002). Instant messaging in teen life. In Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW) (pp. 21-30).

Whiteboard Diagrams Paper (Cherubini et al, 2007)

- "Diagrams are important tools in every design and engineering discipline.
- "Few studies, however, have investigated diagram use in software development activities.
- "There may be fundamental differences between software engineering and other types of engineering.
  - "research suggests that developers are bound to the written form of their code, and so source code editors are the most-used tools for design.

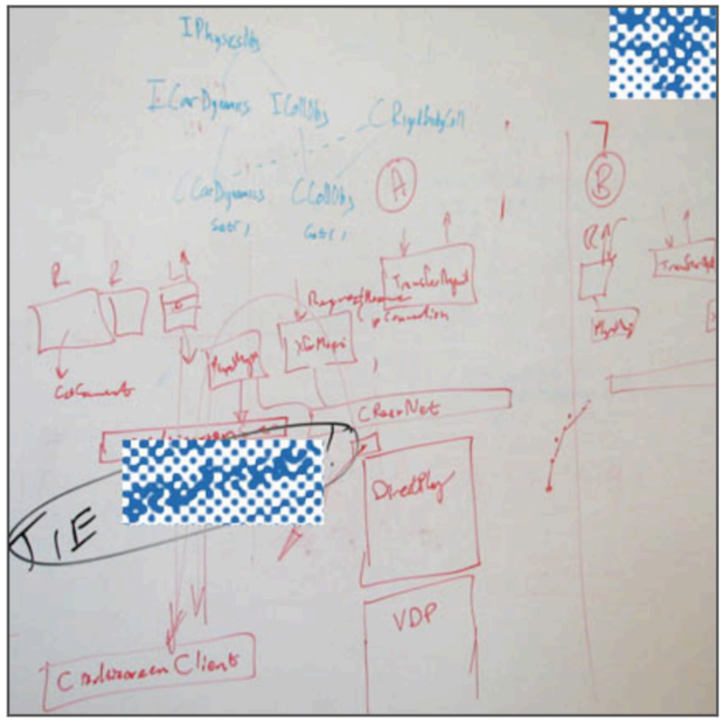


Figure 2: A developer's office whiteboard, with drawings produced during multiple ad-hoc meetings [Tom].

Cherubini, M., Venolia, G., DeLine, R., & Ko, A. J. (2007). Let's go to the whiteboard: How and why software developers use drawings. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI) (pp. 557-566).

Whiteboard Diagrams Paper (Cherubini et al, 2007)

- "Diagrams and drawings seem to have an important role in software development, but we have little understanding about the extent to which they are used, and how their use compares to the use of diagrams in other disciplines.
- "This knowledge is important in designing any kind of support tool for software development, and it may also help reveal fundamental aspects of diagram use across different engineering disciplines.

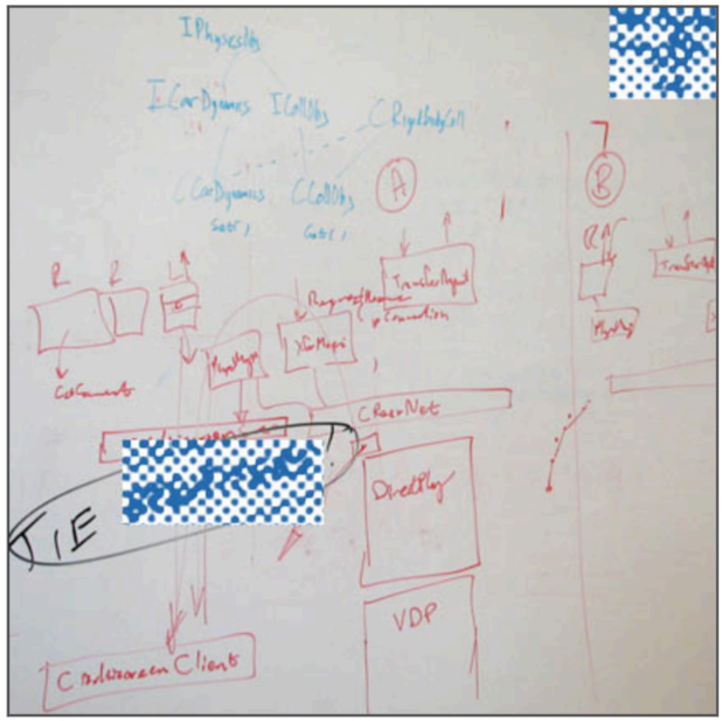


Figure 2: A developer's office whiteboard, with drawings produced during multiple ad-hoc meetings [Tom].

Cherubini, M., Venolia, G., DeLine, R., & Ko, A. J. (2007). Let's go to the whiteboard: How and why software developers use drawings. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI) (pp. 557-566).

Security Folk Models Paper (Wash, 2010)

- "Home users are installing paid and free home security software [(anti-virus, anti-spyware, firewall, …)] at a rapidly increasing rate.
- "Nonetheless, security intrusions and the costs they impose on other network users are also increasing.
- "To design better security technologies, it helps to understand how users make security decisions, and to characterize the security problems that result from these decisions.
- "I investigate the existence of folk models for home computer users. Folk models are mental models that are not necessarily accurate in the real world, thus leading to erroneous decision making, but are shared among similar members of a culture.

Wash, R. (2010). Folk models of home computer security. In Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS) (pp. 1-16).

Security Folk Models Paper (Wash, 2010)

- "Research questions:
  - How do home computer users conceptualize the information security threats that they face?
  - How do home computer users apply their mental models of security threats to make security-relevant decisions?
- "Many of these problems extend beyond the home; […] likely to generalize to a whole class of users who are unsophisticated in their security decisions. This includes many university computers, computers in small business that lack IT support, and personal computers used for business purposes."

Wash, R. (2010). Folk models of home computer security. In Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS) (pp. 1-16).

Social Coding Paper (Dabbish et al, 2012)

- "Visible cues of others' behavior on a social website are likely to support a variety of interpretations about their motivations and the community more generally. People are social creatures and make inferences about others from what they observe (e.g., [27]). Surfacing information about people's actions on artifacts is no longer a technological challenge. What is more interesting, and less understood, is what people are able to infer from such a collection of information, and how these inferences help them carry out their collective work.

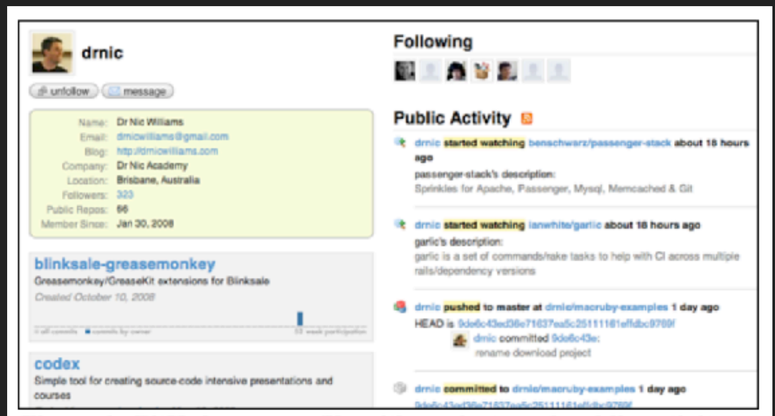


Figure 1. GitHub user profile with projects and public activity

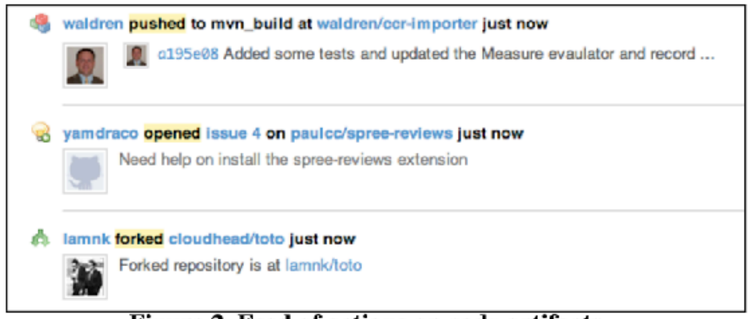


Figure 2. Feed of actions on code artifacts

Figure 3. Network view: sequence of actions on code artifacts

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub: transparency and collaboration in an open software repository. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW) (pp. 1277-1286).

Social Coding Paper (Dabbish et al, 2012)

- "We address the following two research questions to advance our understanding of transparency in online social sites:
  - What inferences do people make when transparency is integrated into a web-based workspace?
  - What is the value of transparency for collaboration in knowledge-based work?
- "Our results inform the design of social media for large-scale collaboration, and imply a variety of ways that transparency can support innovation, knowledge sharing, and community building.

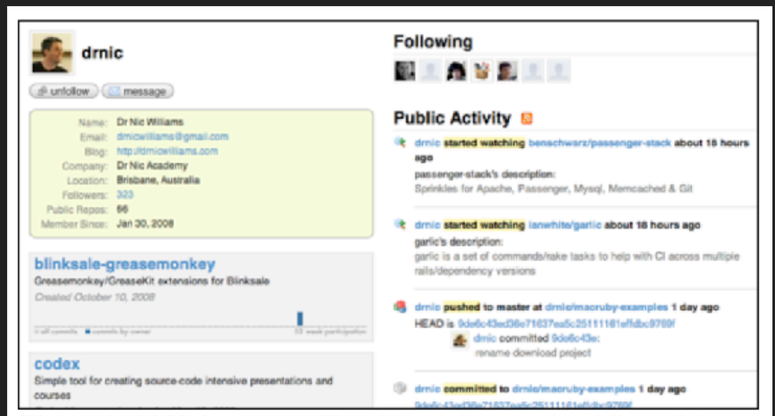


Figure 1. GitHub user profile with projects and public activity

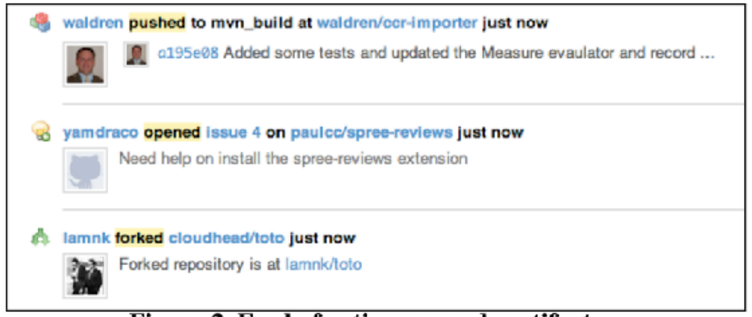


Figure 2. Feed of actions on code artifacts

Figure 3. Network view: sequence of actions on code artifacts

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub: transparency and collaboration in an open software repository. In Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW) (pp. 1277-1286).

Math Overflow Paper (Tausczik et al, 2014)

- How do people collaborate online to solve complex, difficult math problems?

Tausczik, Y. R., Kittur, A., & Kraut, R. E. (2014). Collaborative problem solving: A study of Math Overflow. In Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (pp. 355-367).

Green SE Paper (Manotas et al, 2016)

- "As the use of [mobile] applications and services has expanded, so too have concerns about the amount of energy that they consume.
- "Despite its increasing popularity as a research topic, little is known about practitioners' perspectives on green software engineering.
- "Even basic questions such as
  - 'What types of software commonly have requirements about energy usage?',
  - 'How does the importance of reducing energy usage compare to other requirements?', and
  - 'How do developers find and correct energy usage issues?'

do not have clear answers."

Manotas, I., Bird, C., Zhang, R., Shepherd, D., Jaspan, C., Sadowski, C., Pollock, L., & Clause, J. (2016). An empirical study of practitioners' perspectives on green software engineering. In 2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE) (pp. 237-248). IEEE.

Green SE Paper (Manotas et al, 2016)

- “Without understanding practitioner’s needs, researchers may find themselves in a situation where, despite the investment of significant amounts of time and effort, tools and techniques designed to make practitioners’ lives easier are underused in practice.”

Manotas, I., Bird, C., Zhang, R., Shepherd, D., Jaspan, C., Sadowski, C., Pollock, L., & Clause, J. (2016). An empirical study of practitioners' perspectives on green software engineering. In 2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE) (pp. 237-248). IEEE.

Sex Workers Paper (Barwulor et al, 2021)

- "There is a 'paucity of empirical data' about the role of technology in the business of sex work.
- "This sizeable, digitally-facilitated workforce faces significant challenges with a unique set of social, political, legal, and safety constraints.
  - "may be exacerbated by the fact that many sex workers sit at the intersection of multiple marginalized identities.
- "No existing work in HCI examines technology-enabled sex work as a business directly and empirically through interviews with sex workers themselves."

Barwulor, C., McDonald, A., Hargittai, E., & Redmiles, E. M. (2021). "Disadvantaged in the American-dominated internet": Sex, Work, and Technology. In Proceedings of the 2021 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) (pp. 931-936).

Sex Workers Paper (Barwulor et al, 2021)

- "Our findings offer insight into
  - how the internet is used for sex work by sex workers,
  - how mainstream technological platforms discriminate against this group, and
  - how technologists can better support inclusive and non-discriminatory online spaces for this sizeable, marginalized segment of the global workforce."

Barwulor, C., McDonald, A., Hargittai, E., & Redmiles, E. M. (2021). "Disadvantaged in the American-dominated internet": Sex, Work, and Technology. In Proceedings of the 2021 ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) (pp. 931-936).

What study design do they choose?

Teen Instant Messaging Paper (Grinter & Palen, 2002)

- Semi-structured interviews with 16 teenage IM users
  - “The intention was to take a grounded, bottom-up approach to the investigation, allowing the most common and significant issues to emerge from the inquiry, with few initial expectations.”
- Observation of online activity

Whiteboard Diagrams Paper (Cherubini et al, 2007)

- Initial survey (350 random Microsoft engineers)
  - Identify who used code diagrams in their work
- Semi-structured interviews with 9:
  - What, why, when, how diagram was created / used
- Large-scale survey
  - Learn more about scenarios where diagrams were created

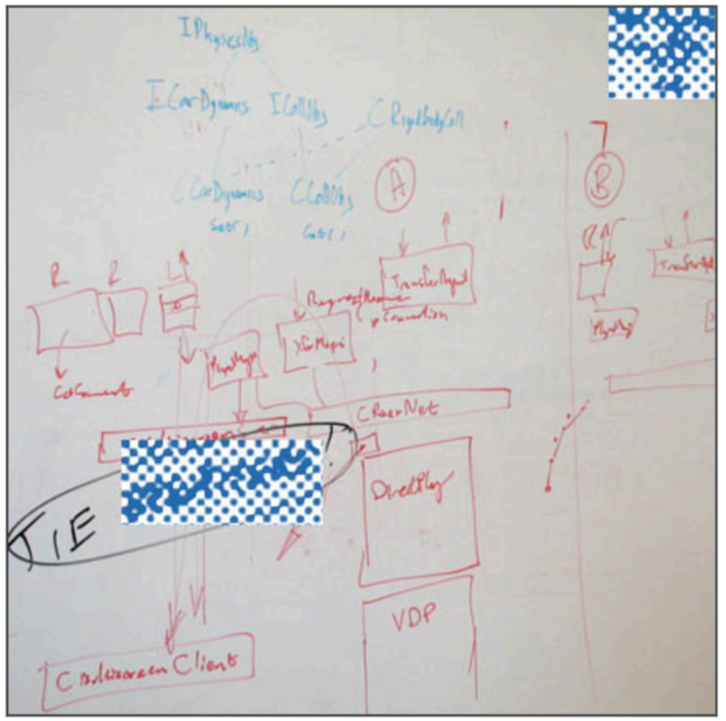


Figure 2: A developer's office whiteboard, with drawings produced during multiple ad-hoc meetings [Tom].

Security Folk Models Paper (Wash, 2010)

- Iterative methodology “as is common in qualitative research [24]:”
  - “multiple rounds of interviews punctuated with periods of analysis and tentative conclusions”
- First round: 23 semi-structured interviews in Summer 2007.
- Second round: 10 interviews in Summer 2008.
  - “This second round was more focused, and specifically searched for negative cases of earlier results [24].”

Social Coding Paper (Dabbish et al, 2012)

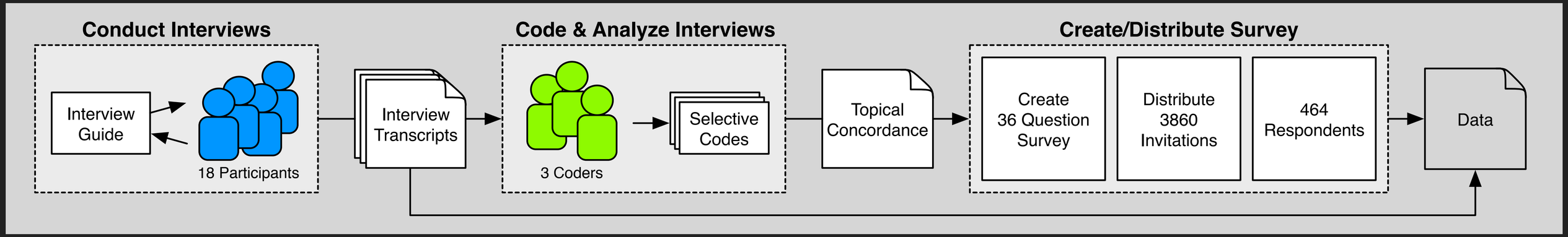
- Semi-structured interviews with 24 GitHub users
- Open coding / grounded theory approach to analyze the transparency related inferences in interview responses

Math Overflow Paper (Tausczik et al, 2014)

- Study 1: How does collaborative problem solving occur on Math Overflow?
  - Grounded theory on a sample of MO questions –> taxonomy of “collaborative acts”
  - Semi-structured interviews with active participants in the collaborations –> context
- Study 2: What value, if any, each type of collaborative act identified in Study 1 added to the solution quality?
  - Linear regression modeling

Green SE Paper (Manotas et al, 2016)

- Semi-structured interviews (exploratory, open ended)
- Survey (large scale, quantitative, confirmatory)



Sex Workers Paper (Barwulor et al, 2021)

- 29 semi-structured interviews in late 2018 with sex workers in Germany and Switzerland, where sex work is legal.
- Recorded all audio interviews and copied all chat transcripts.
- Grounded-theory open-coding process to analyze the interviews.

Specifically how do they conduct their interviews?

Teen Instant Messaging Paper (Grinter & Palen, 2002)

- "With a set of 16 teenage IM users the first author conducted in-depth interviews lasting from 1/2 to 3 hours each.
- "Interviews with P1-4 took place in the United Kingdom and P5-16 in the United States.
- "All participants lived in regions where the local economies centered on computing and telecommunications. Our assumption was that this population of people leads others in technology adoption, and that examination of such a group forecasts future practice of wider and more diverse populations.
- But:
  - Subject recruitment isn't described.
  - How was the number of participants determined?
  - Did the subjects know each other? Many were from the same school.
  - The paper doesn't include an interview script.
  - When were the interviews conducted? 2000? 2001? 2002?

Whiteboard Diagrams Paper (Cherubini et al, 2007)

- "The first two authors conducted the interviews, which typically lasted 45 mins.
- "After introductions, we explained the goals of the study, that their answers would be anonymous, that they could decline to answer any question, and that they could terminate the interview at any time. Additionally, we asked permission to audio record the conversation and photograph drawings that they showed us.
- Guide – four functional areas:
  - WHAT (e.g., "Please, tell me something about this visualization."),
  - WHY (e.g., "Why did you produce this visualization?"),
  - WHEN (e.g., "When did you use it last? For what purpose?"), and
  - HOW (e.g., "How do you use it?").
- "We did not ask the questions sequentially but we tried to respect the flow of the conversation, always trying to touch a couple of points in each of the four areas.

Security Folk Models Paper (Wash, 2010)

- "Respondents were chosen from a snowball sample of home computer users evenly divided between three midwestern U.S. cities.
  - "snowballing through recommendations ensured that the contacted respondents would be information-rich and cooperative
- "Purposefully selected respondents for maximum variation.
- Interviews averaged 45 minutes each; audio recorded and transcribed.

Social Coding Paper (Dabbish et al, 2012)

- Solicited participants via email; conducted interviews in person or via phone.
- Participants were chosen for equal representation across peripheral and heavy users (with greater than 80 watchers on at least one of their OSS projects). This was done because we thought that serious and hobby users might have different purposes and strategies, and very different information loads.
- Participants were asked to walk us through their last session on GitHub, describing how they interpreted information displayed on the site as they managed their projects, and interacted with other users' projects.
- Remote participants shared their screen during the interview using Adobe Connect so we could ask specific questions about data on the site and users could demonstrate their activities on the site.
- Interviews lasted approximately 45 minutes to one hour.
- Interviews were then transcribed verbatim to support further analysis.
- The interviews, videos and field notes supported analysis process.

Math Overflow Paper (Tausczik et al, 2014)

- Semi-structured interviews with 35 of the most active users from MO
  - “because they have the most familiarity with the site and have been exposed to a variety of different collaborations.”
- All male; current Ph.D. students or graduates (n = 6 interviewees), postdoctoral fellows (n = 1), or professors (n = 9) in mathematics.
- Conducted by phone or skype call with at minimum audio (n = 9); instant messenger (n = 2); or email (n = 5). The interviews by phone and skype were recorded and transcribed (M = 42 min); the interviews by instant messenger (M = 1,803 words) and email (M = 980 words) were saved.

Green SE Paper (Manotas et al, 2016)

- Semi-structured interviews, 30 to 60 minutes each, audio-recorded
- Interview guide had four main parts:
  - general demographics,
  - views on energy usage,
  - open-ended and clarification questions based on the second part of the interview:
    - "The interactive nature of the conversations allowed the interviewers to gather detailed information about the participant's experiences with techniques, policies, specifications, patterns, contexts, failed and successful attempts, etc."
  - the interviewers thanked the participant, explained how their responses would be used, and asked whether there was anything else they wanted to mention that was not previously covered.

Green SE Paper (Manotas et al, 2016)

- Recruitment through:
  - mailing lists related to energy use,
  - querying the employee database with energy-related keywords, and
  - communicating with product group managers
- Maximum Variation Sampling:
  - "Because our goal was to learn about as many perspectives as possible, we ensured that the participants came from a range of projects and platforms and had various roles and levels of seniority."
- Expansion using the snowball process—participants added based on recommendations from current participants—until data saturation:
  - "Using the snowball process allowed us to access the hidden population of experienced green software practitioners—practitioners who we would otherwise be unable to identify—without incurring prohibitive costs."

Sex Workers Paper (Barwulor et al, 2021)

- 29 semi-structured interviews in late 2018 with sex workers in Germany and Switzerland, where sex work is legal
- Interview guide:
  - context in the industry (e.g., how long, what type of work).
  - context on their non-work-related technology use (e.g., how long using the internet).
  - how they used technology for sex work.
- "The interview questions used in this analysis are included in the Appendix."
- "To ensure that we were up-to-date on appropriate and region-specific terms workers preferred to use when talking about their work, [we] conducted an informal analysis of four publicly-accessible online sex-worker forums. This forum analysis was not intended as a research artifact but rather used to help us develop the most effective interview protocol."

Sex Workers Paper (Barwulor et al, 2021)

- "After we drafted the interview protocol, we hired a sex worker as a consultant to review our protocol for appropriateness and to ensure a member of the community under study was involved in the research to the extent that they desired to be involved [68]. The consultant was paid market rate for their work."
- Three different approaches to recruit:
  - direct contact;
  - contact through sex-work organizations; and
  - participant-driven (snowball) sampling.

Sex Workers Paper (Barwulor et al, 2021)

- "The interviews lasted approximately one hour, with the shortest running 30 minutes and the longest running two hours.
- "Interviews were conducted by one of three researchers in either English or German, depending on the participant's preference. The English and German interviewers met after approximately every five interviews to ensure they remained consistent in interview length and mitigated any issues or variances in the protocol that had arisen.
- "Participants choose from one of three interview modes: chat, voice, or video.
- "For participant safety, all interviews were conducted using private paid 'rooms' on Appear.in, an end-to-end encrypted communication service.
- "We paid interviewees the equivalent of $75USD (75CHF or 60 Euros) for their participation in the form of an Amazon gift card or money transfer.

Sex Workers Paper (Barwulor et al, 2021)

- "We recognize the importance of our position as scholars in relation to this research [6, 9, 26, 73] and thus describe our identities, their alignment with those of our participants, and how our identities may create limitations in this work.
- "All of the researchers involved in this work identify as women. This is a limitation of the work, as we have participants who identified as other genders and whose experiences may have been better interpreted by researchers with those identities.
- "We have differing sexual orientations, as did our participants. We also have differing nationalities (German, Hungarian, American, and Liberian), some of which overlapped with the regions in which our participants mentioned being born. We have differing races (Black and white) and ages (early 20s through mid-40s), as did our participants.
- "Our lack of researchers of races other than Black and white, and our lack of researchers 50 and older is a limitation of this work, as it does not mirror all of the demographic axes of our participants.

Summary

We Have Seen:

- Quite similar types of research questions
  - Open ended
  - Goal is to build understanding / theory
    - Grounded, bottom up
- Quite diverse study designs
  - From only interviews
  - To various mixed methods with interviews as one step
- Generally very careful execution
- Generally very detailed and transparent reporting
  - Clear motivation for why method was chosen
  - … why a specific design decision (e.g., sampling strategy)
  - Honest and forward acknowledgement of limitations
- Generally simple, straightforward writing

Credits

- Graphics:
  - Dave DiCello photography (cover)