Active and Passive Research through Social Media: The Case for Repurposing Reddit, Instagram and WhatsApp Features in HCI Research Practices

JUAN CARLOS ALVAREZ DE LA VEGA, Northumbria University, United Kingdom

MARTA E. CECCHINATO, Northumbria University, United Kingdom

DANIEL LAMBTON-HOWARD, Open Lab, Newcastle University, United Kingdom

DANIEL HARRISON, Northumbria Universityy, United Kingdom

1 INTRODUCTION

Social media and networking platforms are unique digital spaces that offer many opportunities for research from data collection and participant feedback, to better understanding of people's lives and the ways in which they navigate technologies. In this position paper, we discuss how social media can be used in research in both *passive* and *active* ways, calling for more work to consider in particular active ways of using social media for research. We argue that there is significant scope for the leveraging of the unique affordances of social media and networking platforms in support of these. Building on the concept of *'unplatformed design'*, we conceptualise social media platforms' as a design material, one that allows for reconfiguration and appropriation beyond their initial intended usage [11]. Drawing on our own experiences of using platforms such as Reddit, Instagram and WhatsApp, we discuss the material elements of each of these platforms, the ways in which they can be appropriated within research contexts, and point to the potential for appropriating and repurposing their intrinsic affordances and features.

2 RESEARCHING WITH REDDIT

Traditionally, HCI and social computing scholars have been interested in understanding discourse and dynamics of online communities. One such community is the website Reddit, which describes itself as "home to thousand of communities, endless conversations and authentic human connection" [13]. Reddit's application programme interface (API) allows for the flexible retrieval of posts, comments, and associated metadata from its communities, thus presenting an enormous corpus of

data for research. The readily accessible data stream from Reddit provides opportunities to research online communities that may otherwise be difficult to reach, such as geographically dispersed gig workers [1]. Moreover, Reddit provides a space for the sharing, discussing and community rating of multimedia content. This has proven useful for enabling people without specialist knowledge to engage in meaningful conversations, such as ethical design [7]. Indeed, ready access to Reddit data for research has been instrumental in wider explorations and understanding of people's interactions and use of technology.

Research that simply treats social media platforms as a source of data can be conceived as being a relatively passive form of research. By passive, we mean treating social media data as an emergent property of discourse within online communities, rather than actively engaging with the online community themselves within the research process. Of course, we do not seek to denigrate this type of research, which requires researchers to sort, navigate and analyse often vast and complex datasets. There is indeed value in more passive ways of data collection such as this, especially when researchers want to avoid any risk of introducing their own assumptions or biases. Using social media for passive data collection has been a popular and easy way to gain access to topics and communities, within and outside HCI. However, we argue here that there is scope for exploring what a more active approach to researching through social media platforms would look like.

3 RESEARCHING WITH INSTAGRAM

Instagram's visually rich content and user-generated nature has equal potential for exploratory research. For example, Ayobi et al. [4] analysed Instagram users' paper bullet journal practices and preferences to draw design directions for self-tracking technologies. In their study, they selected a number of relevant hashtags to narrow their search and collect data. In this way, Ayobi et al. [4] utilised the features of Instagram (such as its hashtags), as a research resource to tackle the design of their research study. This early research was critical to later inform the development of an application that supports mindful and meaningful interactions with health and wellbeing data for people living with multiple sclerosis [3].

Another research affordance of Instagram is its use to access niche communities, such as around craftsmanship [6]. For example, Harrison and Cecchinato[8] began to explore this affordance in the bicycle craft community, harnessing Instagram's in-built features. Differently to Ayobi, we more actively leveraged various features of Instagram to conduct our research, including both data collection and analysis. Firstly, we created a specific *Instagram account* linked to the research project – this could be accessed by anyone on the research team with the login credentials. It also allowed us to source members of this craft niche through hashtag searching, following other relevant accounts, as well as allowing potential participants find the account and contact us directly. By creating an account, we are also able to engage with participants, with minimal burden on their part. For example, consent can be obtained from individual participants, interviews can be conducted directly through the platform via *direct messages*, and survey-like questions can be asked through the *Stories* feature (via polls and question boxes). Moreover, the *Save* and *Collections* features of the platform have been appropriated to aid the analysis and organise posts into themes. However, not all data can be saved in collections; for example users' own Stories and conversations via direct messages cannot be easily exported. While similar approaches of using social media to collect visual data have been used before [9], the affordances of Instagram allow the researchers to go beyond just collecting a repository of data. As such, this approach echoes the *'pragmatic utilities'* of *"drawing attention to the qualities of technologies that can be employed in [research] design"* [10] p.9. Much can be learned from this example of appropriating Instagram's features to structure and enable multiple stages of a research project, both in terms of how research can be conducted, but also how the platform could be redesigned to support more research-like features.

However, unlike Reddit data, research through Instagram requires that researchers are in some way subject to the platform's discovery and visibility algorithms, meaning that replicating findings might not be possible. This is a consequence of Instagram's materiality in terms of its *'representation of activity'* [11]. Moreover, using a singular commercial platform to both collect data and analyse it can be both an advantage (e.g., storing data all in one place) and a limitation (e.g., being locked into the platform and having no control over it), so data management practices need to be carefully considered and back-up plans are necessary.
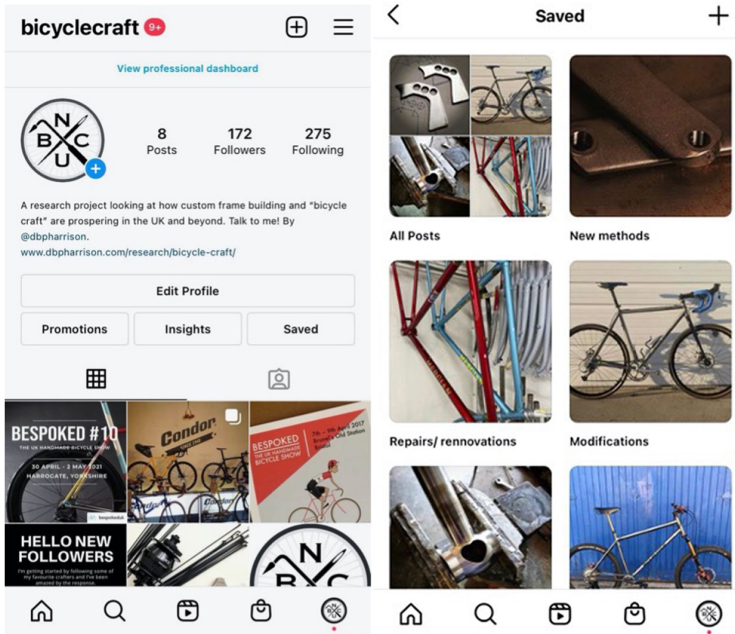

Fig. 1. Bicycle craft Instagram account (left) and Saved posts (right)

4 RESEARCHING WITH WHATSAPP

WhatsApp – although more of social networking platform than a social media – is the most popular messaging app worldwide. Its popularity lends itself as an ideal space for communicating with participants, and conducting remote data generation.

For example, WhatsApp has been repurposed for conducting HCI research at scale. Lambton-Howard et al. [10] harnessed WhatsApp's multimedia and group formation affordances to develop and deploy a large-scale 'strategic foresight' game event, with 487 players across five countries as part of a collaboration with the International Federation of Red Cross and Red Crescent Societies. This novel engagement method highlighted the value of using social media technologies for a geographically distributed coordinated engagement. It also showed the critical importance of meeting participants in digital spaces that they are already familiar with within research (as opposed to imposing new technologies), although the onboarding and consenting process in this case happens outside of the platform. Most importantly, this project cast light on how paying attention to the configuration of the material qualities of social media platforms can be used to design new types of interaction. This is referred to by the authors as *'unplatformed design'* [11]. Yet, many of these interactions have been underexplored, raising essential areas for future work.

WhatsApp's affordances have proven equally useful for smaller-scale research practices. For instance, Rainey et al. [12] provided coordinated support and training through WhatsApp on an engagement aimed at distributed participatory qualitative processes. Here, participants were sent training materials (e.g., short videos), gave feedback on each other's work, and asked questions to the research team. Another example is repurposing WhatsApp's multimedia affordances to support qualitative diary methods, whereby participants can act upon researcher-directed prompts and capture moments of their everyday work-life experiences. In our study on freelancers' ways of managing work-life balance [2], we have set up daily prompts for participants whereby they need to respond to daily prompts (e.g., Fig. 2). The ability to share audio, photos, videos, links as well as text and voice notes is important when wanting to offer flexibility to participants. Moreover, the encryption feature allows participants to know that their data is protected, unlike other social media platforms. This makes it a preferable tool when researching more sensitive topics. The direct conversation with the researcher means that participants can more easily ask questions at any point of the study and build a more direct rapport between the researcher and the participants. Also, the researcher can further probe any responses as and when they are logged, mitigating any recall issues that might occur in a post-diary interview. Here WhatsApp offers more flexibility in terms of data collection compared to more traditional experience sampling methods.
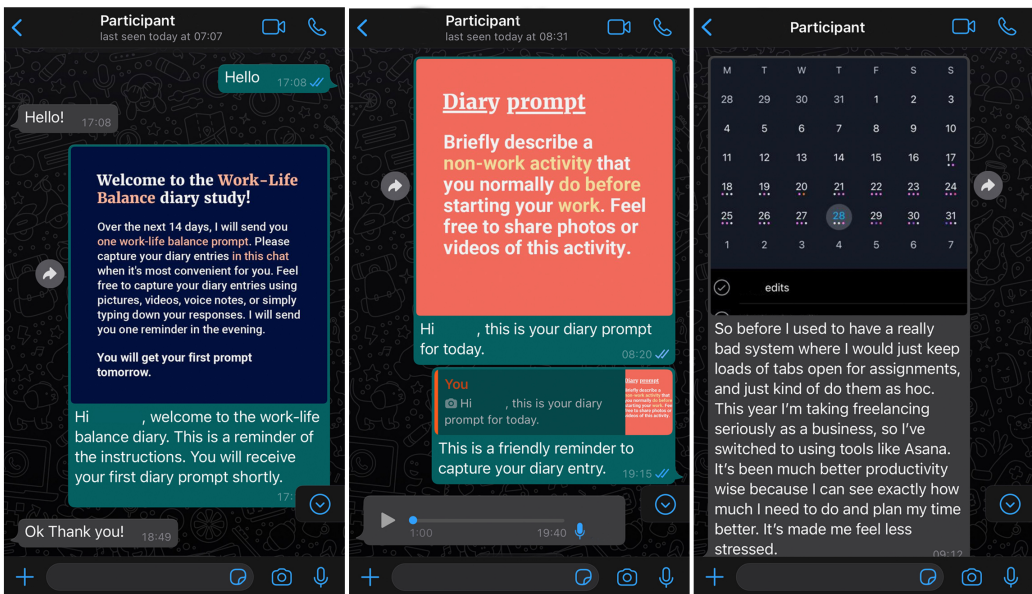

Fig. 2. Three anonymised screenshots illustrating WhatsApp's features supporting the qualitative diary method mentioned in [2].

5 DISCUSSION AND CONCLUSION

Relying on social media and social networking platforms for data collection offer the benefit of overcoming *"the hurdles associated with attempts to migrate users onto new unfamiliar platforms"* ([10] p. 10) for engaging with research. We have shared ways in which Reddit, Instagram and WhastApp have been repurposed to exemplify some of the ways in which research could be conducted within HCI. Our position paper highlights how social media platforms can be appropriated for all stages

of research. However, the features are generally not designed with research in mind, meaning that compromises around privacy, control and use of data need to be carefully considered.

More passive forms of data collection on social media platforms have been the norm in HCI research, overlooking the opportunities of repurposing features of the system for conducting research. We have highlighted examples of our previous work on how social media and social networking platforms can be actively embedded in the research design process to expand already well-established methods, such as experience sampling. Moreover, these platforms bring about an opportunity for including people in research processes and coordinated action at global scales that would be traditionally hard. Yet, much of these practices still lack a common vocabulary amongst the research community. For example, the ethics of passive collection of data on social media has been discussed at length in existing literature on platforms such as Twitter [5], but active collection requires further considerations and a different approach. We hope to continue discussing how the HCI community can build on the challenges of active research and how industry partners can support this process.

REFERENCES

[1] Juan Carlos Alvarez de la Vega, Marta E Cecchinato, and John Rooksby. 2021. "Why lose control ?" A Study of Freelancers ' Experiences with Gig Economy Platforms. In *ACM Conference of Human Factors in Computing CHI 2021*. Yokohama, Japan. https://doi.org/10.1145/3411764.3445305

[2] Juan Carlos Alvarez de la Vega, Marta E. Cecchinato, John Rooksby, and Joseph Newbold. 2021. A Qualitative Exploration of Upwork Freelancers' Work-Life Balance. https://freelbalance.info/{#}about

[3] Amid Ayobi, Paul Marshall, and Anna L. Cox. 2020. Trackly: A Customisable and Pictorial Self-Tracking App to Support Agency in Multiple Sclerosis Self-Care. *Conference on Human Factors in Computing Systems - Proceedings* (2020). https://doi.org/10.1145/3313831.3376809

[4] Amid Ayobi, Tobias Sonne, Paul Marshall, and Anna L. Cox. 2018. Flexible and mindful self-tracking: Design implications from paper bullet journals. *Conference on Human Factors in Computing Systems - Proceedings* 2018-April (2018), 1–14. https://doi.org/10.1145/3173574.3173602

[5] Casey Fiesler and Nicholas Proferes. 2018. "Participant" perceptions of Twitter research ethics. *Social Media+ Society* 4, 1 (2018). https://doi.org/10.1177/2056305118763367

[6] Derek Foster, Ben Kirman, Conor Linehan, and Shaun Lawson. 2017. The role of social media in artisanal production: A case of craft beer. *Proceedings of the 21st International Academic Mindtrek Conference, AcademicMindtrek 2017* 2017-Janua (2017). https://doi.org/10.1145/3131085.3131117

[7] Colin M. Gray, Shruthi Sai Chivukula, and Ahreum Lee. 2020. What Kind of Work Do "Asshole Designers" Create? Describing Properties of Ethical Concern on Reddit. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. ACM, New York, NY, USA, 61–73. https://doi.org/10.1145/3357236.3395486

[8] Danny Harrison and Marta E. Cecchinato. [n.d.]. Bicycle Craft Research. http://www.dbpharrison.com/research/bicycle-craft/

[9] Michael Karlesky and Katherine Isbister. 2016. Understanding Fidget Widgets: Exploring the Design Space of Embodied Self-Regulation. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*, ACM (Ed.). Eindhoven, Netherlands. https://doi.org/10.1145/2971485.2971557

[10] Daniel Lambton-Howard, Robert Anderson, Kyle Montague, Andrew Garbett, Shaun Hazeldine, Carlos Alvarez, John A. Sweeney, Patrick Olivier, and Ahmed Kharrufa. 2019. Whatfutures: Designing large-scale engagements on Whatsapp. *Conference on Human Factors in Computing Systems - Proceedings* (2019), 1–14. https://doi.org/10.1145/3290605.3300389

[11] Daniel Lambton-Howard, Patrick Olivier, Vasilis Vlachokyriakos, Hanna Celina, and Ahmed Kharrufa. 2020. Unplat-formed Design: A Model for Appropriating Social Media Technologies for Coordinated Participation. *Conference on Human Factors in Computing Systems - Proceedings* (2020), 1–13. https://doi.org/10.1145/3313831.3376179

[12] Jay Rainey, Carlos Alvarez de la Vega, Dan Richardson, Daniel Lambton-howard, Tom Bartindale, Shaun Hazeldine, Pamela Briggs, Patrick Olivier, and Kyle Montague. 2020. TalkFutures : Supporting Qualitative Practices in Distributed Community Engagements. In *Designing Interactive Systems*, ACM (Ed.). Eindhoven, Netherlands. https://doi.org/10.1145/3357236.3395531

[13] Inc Reddit. [n.d.]. Reddit, Inc. https://www.redditinc.com