Full Bibliographic Citation:

D. Votipka, M. Nicole Punzalan, S. M. Rabin, Y. Tausczik and M. L. Mazurek, "An Investigation of Online Reverse Engineering Community Discussions in the Context of Ghidra," *2021 IEEE European Symposium on Security and Privacy (EuroS&P)*, 2021.

Link to published (version) of the paper:

This paper has been accepted at the IEEE European Symposium on Security and Privacy, but due to the deadlines, the camera-ready version has been accepted, but not released publicly by IEEE. Therefore, we don't have a link to the final, official copy of the paper. The paper can be found in the accepted papers list (https://www.ieee-security.org/TC/EuroSP2021/accepted.html) on the Euro S&P page. So the reviewers can read the camera-ready version of the paper, we have made it publicly available on our own personal website (https://www.eecs.tufts.edu/~dvotipka/files/papers/VotipkaGhidraEuroSP2021.pdf).

Paper abstract:

Reverse engineering is a complex task. As with many other expert tasks, reverse engineers rely on colleagues and the broader reverse engineering community to provide guidance and develop knowledge necessary to achieve their goals. For example, it is common for reverse engineers to reach out for help to understand and effectively use new tools. Thus far, however, there has been limited investigation of the way knowledge is developed in this community and new tools are adopted. This paper takes a first step toward understanding reverse engineering community dynamics around tool adoption, using the release of the National Security Agency's

Ghidra reverse engineering framework as a point of focus. In this paper, we review discussions about Ghidra to identify what features reverse engineers are most interested in, how reverse engineers develop knowledge about Ghidra together online, and whether these dynamics differ between forums.

In total, we analyze 1590 reverse engineering discussions between 688 reverse engineers over 3 forums (i.e., Twitter, Reddit, and StackExchange). Our results suggest reverse engineers are most interested in features that allow them to customize Ghidra. We also observe limited evidence of collective sensemaking on the forums, with few reverse engineers participating in multiple discussions threads and most acting as either knowledge producers or consumers. Finally, we found that the forums operated similarly, but Twitter was most often used to announce information (e.g., tutorial links, tool overviews, vulnerabilities in Ghidra) and reverse engineers used StackExchange mostly to get support for specific problems. Reddit acted as a middle option. Based on these results, we make recommendations to improve reverse engineering tool development, improve community participation during adoption, and suggest directions for future work.