"Hey Google, Remind me to be Phished"
Exploiting the Google (AI) Assistant Ecosystem for Social Engineering Attacks

Marie Weinz, Saskia Laura Schröer, Giovanni Apruzzese
Liechtenstein Business School, University of Liechtenstein
{*marie.weinz, saskia.schroeer, giovanni.apruzzese*}*@uni.li*

***Abstract*—We showcase how to maliciously exploit a functionality of the Google ecosystem—and, specifically, of Android. By narrating a story, we elucidate how Google may "help" phishers in reaching their goals. We found that Android users who have Google Assistant (which is typically enabled by default in recent smartphones) check their inbox will be "reminded" to carry out duties that are "solicited" in emails that the users themselves have never opened before. From a social-engineering perspective, attackers can send specific phishing emails to Android users, and these users will receive notifications (from Google) "reminding" them that a task is soon due—thereby urging them to "fall for phish." Just imagine: you are going through your day, and you receive a notification on your smartphone saying that "There is an outstanding task which is soon due". Tapping on the notification leads to opening an email containing malicious content (links or attachment). The sense of urgency from the unexpected reminder may lead to overlooking some phishing cues, facilitating social engineering attacks.**

This threat is rooted on the quintessential functionalities of "smart" (AI-based) assistants—which passively analyse our data to improve our digital well-being. Users of these tools must be made aware of this issue to prevent harmful consequences. Therefore, besides describing our discovery and analysing it under a security lens, we also emphasize some practical takeaways for both users and developers. We disclosed our finding to Google, who acknowledged them but stated that no change to their software will be made.

1. Introduction

With nearly 4 billion users worldwide [1], Android is the leading operating system (OS) of modern smartphones, having a market share of 71% [2]. Thanks to its integration with the Google's ecosystem (e.g., GMail), owners of Android devices can benefit from the continuous updates made by one of the world's top tech companies [3]. Among the most recent developments that have substantially enhanced Android users' quality of experience, the *Google Assistant* stands out [4]. Powered by artificial intelligence (AI) [5, 6], Google Assistant monitors the plethora of activities that its users carry out during their daily lives—providing tools and resources (e.g., automatic reminders [7]) that improve the users' well-being [8].

Unfortunately, such a large reservoir of users makes the Android (and, hence the Google Assistant) ecosystem an attractive target for cyberattackers—and, in particular, for *phishers* [9–12]. Indeed, some specific functionalities of Android OS, such as its notification system, can be maliciously exploited to facilitate social engineering attacks—and some of these "security vulnerabilities" have been discussed in prior works [12]. In this paper, we present a novel "vulnerability" of Android which *specifically exploits the Google Assistant ecosystem*.

At a high-level, our discovery is rooted on the fact that Google Assistant *perpetually checks the email inboxes* of their end users. This is done to "help" users, so that if they receive an email stating that, e.g., "a task is due soon", *a notification will be triggered* on their smartphones to warn them. However, the problem is that Google Assistant "blindly trusts" the analysed emails—including those *concealing social engineering attacks* (and, unfortunately, existing phishing email filters can be trivially bypassed [13]). An attacker can thus exploit such a functionality to carry out phishing campaigns, i.e., by using the automatic reminders of Google Assistant as a catalyst to instill a sense of urgency in their victims—who may be more likely to open the email and, e.g., click on a malicious link (Fig. 1).

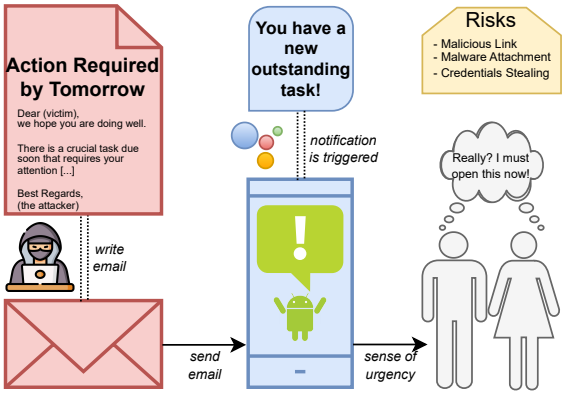

Fig. 1: **Exploiting our discovered vulnerability.** An attacker writes an email stating that "a task is due soon". The email will trigger a notification from Google Assistant, which will remind the users of an outstanding task. The users (i.e., the victims), driven by the sense of urgency, may carelessly open the email and fall for a phishing trap.

⚠ **Why is this a Problem?** We highlight three points that aggravate the "risk" of our discovery:
- the notification is *independent* of the Google Calendar or GMail Android applications;
- the *only requirement* is that users have Google Assistant linked to an email account related to Google;
- the notification is triggered by *an email that is not (nor needs to be) read* by the user.

Put simply, users are suddenly reminded to "do something *now*", but they are oblivious of the reason. This "urgent confusion" may induce users to overlook that they are being targeted by a phishing attack. *Awareness of such a risk* is paramount to protect Android users.

References

[1] BankMyCell. (2024) How many android users are there? global and us statistics. https://web.archive.org/web/20240205082953/https://www.bankmycell.com/blog/how-many-android-users-are-there.

[2] StatCounter. (2024) Mobile operating system market share worldwide. https://web.archive.org/web/20240318202847/https://gs.statcounter.com/os-market-share/mobile/worldwide.

[3] Statista. (2024) Leading tech companies worldwide 2024, by market capitalization. https://web.archive.org/web/20240130130259/https://www.statista.com/statistics/1350976/leading-tech-companies-worldwide-by-market-cap/.

[4] Google. (2020) A more helpful Google Assistant for your every day. https://web.archive.org/web/20201127165358/https://www.blog.google/products/assistant/ces-2020-google-assistant/.

[5] CultOfMac. (2016) The future is AI, and Google just showed Apple how it's done. https://web.archive.org/web/20201108162911/https://cultofmac.com/447898/google-home-google-assistant-siri-ai.

[6] Google. (2023) Assistant with Bard: A step toward a more personal assistant. https://web.archive.org/web/20240222161116/https://blog.google/products/assistant/google-assistant-bard-generative-ai.

[7] Droid-life. (2023) Google assistant reminders go live in google tasks: How to switch. https://web.archive.org/web/2/https://www.droid-life.com/2023/06/01/google-assistant-reminders-go-live-in-google-tasks-how-to-switch/.

[8] T. N. News. (2021) Google assistant now provides ai-powered mental health support to arabic speakers. https://web.archive.org/web/2/https://www.thenationalnews.com/lifestyle/wellbeing/google-assistant-now-provides-ai-powered-mental-health-support-to-arabic-speakers-1.1192581.

[9] S. Aonzo, A. Merlo, G. Tavella, and Y. Fratantonio, "Phishing attacks on modern android," in *ACM CCS*, 2018.

[10] K. Tian, S. T. Jan, H. Hu, D. Yao, and G. Wang, "Needle in a haystack: Tracking down elite phishing domains in the wild," in *IMC*, 2018.

[11] C. Marforio, R. Jayaram Masti, C. Soriente, K. Kostiainen, and S. Čapkun, "Evaluation of personalized security indicators as an anti-phishing mechanism for smartphone applications," in *ACM CHI*, 2016.

[12] A. Ruggia, A. Possemato, A. Merlo, D. Nisi, and S. Aonzo, "Android, notify me when it is time to go phishing," in *IEEE EuroS&P*, 2023.

[13] "State of the phish," https://www.proofpoint.com/it/resources/threat-reports/state-of-phish, ProofPoint, Tech. Rep., 2023.