

Instagram Use and the Well-Being of Adolescents

Using Deep Learning to Link Social Scientific Self-reports with Instagram Data Download Packages

Laura Boeschoten*
l.boeschoten@uu.nl
Utrecht University

Daniel L. Oberski
d.l.oberski@uu.nl
Utrecht University

Irene I. van Driel*
i.i.vandriel@uva.nl
University of Amsterdam

Loes Pouwels
j.l.pouwels@uva.nl
University of Amsterdam

ABSTRACT

Since the introduction of social media platforms, researchers have investigated how the use of such media affects adolescents' well-being. Thus far, findings have been inconsistent [1–3]. The aim of our interdisciplinary project is to provide a more thorough understanding of these inconsistencies by investigating who benefits from social media use, who does not and why it is beneficial for one yet harmful for another [1]. In this presentation, we explain our approach to combining social scientific self-report data with the use of deep learning to analyze personal Instagram archives.

The implementation of the GDPR in 2018 opened up new possibilities for social media research. Each platform is legally mandated to provide its European users with their social media archive in digitally readable format upon request, to which all large platforms currently comply. These data download packages (DDPs) aid in resolving three main challenges in current research. First, the reliability of social media use self-reports suffer from recall bias, particularly among teens [2]. Instagram DDPs provide objective, timestamped insights in Instagram use. Second, previous research has demonstrated that time spent on social media has no or a small relationship with well-being [3]. The Instagram DDPs are an answer to recent calls for knowledge on adolescents' specific activities (posting, messaging, commenting) on social media [3]. Third, DDPs resolve selectivity issues related to research making use of APIs, analyzing public content only, while adolescent use knows an important private component [1].

In a longitudinal study, we invited we invited 388 adolescents (8th and 9th graders of a Dutch high school, $M_{age} = 14.11$, 54% girls) to participate in a panel survey, an experience sample (ESM) and to share their Instagram DDP at the end of both studies. Of this group, 104 Instagram users ($M_{age} = 14.05$, 66% girls) complied to sharing their DDP. As DDPs contain private and third-party content, data managers, ethical committee members and privacy officers have been closely monitoring the research process. Here, we developed a script in Python that anonymizes parts of the DDP by removing

identifiers from images, videos, and text. Other parts of the DDP are pseudonymized, allowing us to for example connect befriended users within the study. During the presentation, we report on this preparation process and the validation of the anonymization and pseudonymization script.

With the combined data-set containing panel survey results, ESM data and Instagram DDPs, we plan to perform a number of analyses, intended to investigate both the possibilities of using DDPs for scientific research and to investigate the well-being of adolescents. First, we plan to investigate the representativeness of the sub-sample that complied with sharing their DDPs. Second, we plan to generate emotional classifications and classifications of contextual factors of the images and text found in the DDPs using Microsoft Azure Cognitive Services and relate this to self-reported trait levels of well-being, derived from both the survey and the ESM. Third, we plan to develop natural language processing and computer vision algorithms using the DDP content as data and self-reported state levels of well-being as labels. By combining deep learning and social science, we aim to understand differences in Instagram use between adolescents who feel happy and those who feel less happy.

CCS CONCEPTS

• **Applied computing** → *Digital libraries and archives*; **Psychology**.

KEYWORDS

adolescents; well-being; data-download package; Instagram

ACKNOWLEDGMENTS

This research is funded by an NWO Spinoza Prize and a Gravitation grant (NWO Grant 024.001.003; Consortium on Individual Development) awarded to P. M. Valkenburg and by an NWO Vidi grant (VI.Vidi.195.152; valid measures derived from incidental data) awarded to D. L. Oberski.

REFERENCES

- [1] Ine Beyens, J Loes Pouwels, Irene I van Driel, Loes Keijsers, and Patti M Valkenburg. 2020. The effect of social media on well-being differs from adolescent to adolescent. *Scientific Reports* 10, 1 (2020), 1–11.
- [2] Nastasia Griffioen, Marieke MJW Van Rooij, Anna Lichtwarck-Aschoff, and Isabela Granic. 2020. A stimulated recall method for the improved assessment of quantity and quality of social media use. *Journal of Medical Internet Research* 22, 1 (2020).
- [3] Amy Orben. 2020. Teenagers, screens and social media: a narrative review of reviews and key studies. *Social psychiatry and psychiatric epidemiology* (2020), 1–8.

*Authors contributed equally

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

ICMI '20 Companion, October 25–29, 2020, Virtual event, Netherlands

© 2020 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-8002-7/20/10.

<https://doi.org/10.1145/3395035.3425185>