Empowering Digital Natives: InstaClone – a Novel Approach to Data Literacy Education in the Age of Social Media

Anna Hartl*
a.hartl@tum.de
Technical University of Munich
Research on Learning and Instruction
Munich, Germany

Elena Starke*
elena.starke@tum.de
Technical University of Munich
Computing Education Research
Munich, Germany

Angelina Voggenreiter*
angelina.voggenreiter@tum.de
Technical University of Munich
Computational Social Science
Munich, Germany

Doris Holzberger
doris.holzberger@tum.de
Technical University of Munich
Research on Learning and Instruction
Munich, Germany

Tilman Michaeli
tilman.michaeli@tum.de
Technical University of Munich
Computing Education Research
Munich, Germany

Jürgen Pfeffer
juergen.pfeffer@tum.de
Technical University of Munich
Computational Social Science
Munich, Germany

ABSTRACT

Social media has become an integral part of the lives of young people, who, despite being regarded as digital natives, lack essential skills in terms of the reflective use of data, thus underscoring the potential of computing education to empower their data literacy. To this end, this paper presents InstaClone, an innovative educational tool for classrooms that allows students to engage with social media platforms within a secure learning environment. With an appearance and functionality resembling Instagram, InstaClone offers a lifelike learning environment in which students can generate data, which are then processed and visualized on integrated data analytics dashboards, facilitating the development of individual data literacy. A case study with year 9 and year 10 students in a K-12 computer science class demonstrated that InstaClone convincingly emulates the real platform and that students benefit from using the tool by developing a deeper understanding of the data collected by social media platforms and the underlying algorithms.

CCS CONCEPTS

• **Human-centered computing** → **Social media**; • **Social and professional topics** → **K-12 education**.

KEYWORDS

social media, data literacy, K-12, computer science education

ACM Reference Format:
Anna Hartl, Elena Starke, Angelina Voggenreiter, Doris Holzberger, Tilman Michaeli, and Jürgen Pfeffer. 2024. Empowering Digital Natives: InstaClone – a Novel Approach to Data Literacy Education in the Age of Social Media. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1 (SIGCSE 2024), March 20–23, 2024, Portland, OR, USA.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3626252.3630839

*Authors contributed equally to this research.

SIGCSE 2024, March 20–23, 2024, Portland, OR, USA

ACM ISBN 979-8-4007-0423-9/24/03…$15.00
https://doi.org/10.1145/3626252.3630839

1 INTRODUCTION

Within our society, there are concerning signs that social media platforms are having a detrimental influence on democracy and public discourse, underscoring the significance of equipping students with an understanding of the underlying mechanisms and algorithms of these platforms [17]. Despite being born in the digital age and regarded as digital natives [2], students often display significant deficits when it comes to critically evaluating information [4], thus highlighting the importance of cultivating their ability to distinguish between reliable and non-reliable sources. As democracy relies on access to accurate and trustworthy knowledge, empowering students with the skills needed to discern truth from misinformation is indispensable for ensuring their active participation in shaping public discourse [17]. Furthermore, to be a responsible member of the community, it is crucial to have the capacity to critically reflect on one's own social media use and to conscientiously navigate its associated challenges [5].

To address the adverse effects of social media use in our democratic society, data literacy is the foundation for profound usage. For this, computing education is essential, as it provides a comprehensive understanding of the mechanisms and technical functionality of social media platforms and thus makes interdependencies assessable. Consequently, students can use social media platforms more reflectively [10] and can develop a firm foundation on which to deal responsibly with social media when actively shaping the future (digital) world [17].

To foster students' data literacy in the context of social media, we developed InstaClone, a unique educational tool for classroom purposes that replicates the appearance and functions of Instagram. By providing a comprehensive and interactive learning experience that incorporates students' own data, this innovative tool aims to enhance students' data literacy skills. To do so, it goes beyond mere imitation via integrated data analytics dashboards that give students a behind-the-scenes view of a social media platform. These dashboards allow students to explore the data collected by InstaClone and to gain insights into the dynamics by, for example, choosing

the newsfeed algorithm or creating advertisements based on a network analysis of their followers. As a result, InstaClone provides hands-on experience of the dynamics possible on social media platforms in the classroom by stepping apart from only hypothetically discussing advantages and disadvantages.

The contributions of this work are:

(1) We provide version 1.0 of InstaClone, an innovative educational tool, that allows students to look behind the scenes of a social media platform and accompanying teaching material for the purpose of fostering data literacy using self-created data. Version 1.0 of InstaClone and the associated teaching material can be downloaded here[1].
(2) A case study generated evidence that the tool is achieving these goals and can be successfully employed in the classroom.

2 RELATED WORK

In developing an interactive tool for teaching data literacy to students, careful consideration of multiple key aspects is essential. These aspects include integrating computer science fundamentals relevant to data literacy, incorporating students' prior knowledge about data literacy, leveraging social media as an educational context, and building upon the foundation laid by existing tools designed to enhance data literacy education.

2.1 Using Social Media in the Classroom

Integrating social media platforms as a teaching tool has gained attention due to its potential to enhance education. However, teachers primarily utilize social media for their own professional development rather than incorporating it into classroom instruction [21]. When social media tools are employed in the classroom, teachers often follow the principle of self-regulated learning to design their classes. As a result, students who engage with social media tools during instruction tend to perform better when compared to those who do not [21].

Among educational tools that utilize social media as a learning context, *TestDrive* [10] and *InstaHub* [12] stand out as noteworthy platforms. Specifically designed for middle school students, *TestDrive* emulates the appearance and experience of a genuine social media platform, thereby providing an interactive learning environment for digital literacy education. Its closed system ensures a safe space for students to explore and develop their digital literacy (without addressing CS concepts), fostering a deeper understanding of online interactions, privacy, and responsible digital citizenship [10]. *InstaHub* is an educational tool designed to cultivate students' database competencies through an immersive learning environment. Within a closed environment, students can delve into the databases underpinning a social media platform or transition from a user's perspective to that of an administrator, consequently gaining a deeper understanding of database management and becoming empowered via the acquisition of valuable skills for navigating the digital landscape [12].

[1]https://info.instaclone.de/

2.2 Data Literacy

Since contact with data occurs in various situations in daily life, "data literacy is increasingly considered as a life skill" [23]. As per our interpretation, data literacy is a fundamental skill grounded in computer science, encompassing data management and data science [14]. Data literacy entails the capacity to "select, clean, analyse, visualise, critique, and interpret data, as well as communicate stories derived from data and integrate data into design processes" [23]. As data already play a considerable role in the lives of teenagers, fostering data literacy skills from an early age, particularly within the school curriculum, has been emphasized in previous research [20, 23]. Considering the competencies students should acquire at school, Grillenberger and Romeike investigated key concepts of data management and ultimately developed a theoretically founded data literacy competency model that outlines competencies relevant to computing education [13, 14]. However, teenagers develop their own perspectives on data, with some associating data primarily with numerical information that can be represented through spreadsheets and diagrams. In contrast, others link data to the broader context of the internet, to digital footprints, and to surveillance [3]. Nevertheless, despite a perceived positive self-concept regarding data literacy skills, students often exhibit little concern for data privacy. Interestingly, students with a negative self-concept concerning their data literacy skills tend to invest more effort in protecting their personal data [6]. Considering the data life cycle, most students demonstrate a basic understanding of the phases occurring at the beginning and end of the cycle. However, there is a considerable knowledge gap concerning the intermediate stages of the data life cycle and the long-term consequences associated with the digital footprint [3]. To address this knowledge gap, selecting teaching contexts, like social media, that can be linked to students' experiences can be beneficial in motivating them to grasp the relevance of the topic [7].

2.3 Fostering Data Literacy

To effectively enhance students' data literacy skills, researchers have identified several strategies: Involving students in collecting their own data fosters a deeper understanding of data literacy concepts [23], especially when the examples are related to students' daily lives [20, 23]. Furthermore, incorporating social media tools, delivering content through workshops, and facilitating hands-on learning experiences have proven to be advantageous for data literacy education [20]. While numerous tools have been developed to promote students' data literacy, each has different purposes and employs various distinct approaches. These tools typically utilize pre-existing data sets rather than allowing students to collect their own data [23]. Some of these tools utilize real-world data to explore visualization possibilities [19] or introduce students to text mining practices such as word count and document analysis [11]. Additionally, several learning platforms offer online courses to both adults and students, combining hands-on activities with real-data analysis, readings, and instructional videos [8, 16, 18].

However, while some of these tools provide opportunities to upload and analyze user-generated data files, the ability to generate and collect one's own data for analysis is often lacking [23].

Moreover, as *TestDrive* focuses on digital literacy and *InstaHub* emphasizes databases, there are currently no existing tools for data literacy education within the context of social media. Hence, there is a pressing need for innovative tools that not only cultivate hands-on data literacy capabilities but also empower students in computing education to assume responsible roles as social media users, ultimately contributing to their active participation in shaping public democratic discourse. By bridging these gaps, InstaClone affords students practical engagement, fosters a profound comprehension of social media's dynamic and influential landscape, and equips students with the data literacy skills needed to meaningfully participate in a democratic society.

3 INSTACLONE

To foster data literacy among students, we developed InstaClone, a lifelike social media tool accompanied by tailored teaching material. Our approach follows existing evidence by prioritizing hands-on experience with self-generated data in real and relevant contexts. In the following, we describe the requirements that influenced the design choices of the tool, the tool itself, and the content of the accompanying teaching material.

3.1 Requirement Analysis

Based on workshops with teachers, we identified several attributes an educational tool should have so that both teachers and students will find it maximally useful. When designing the tool and the teaching material, we aimed to fulfill these requirements:

- **Authenticity**: InstaClone should be designed in a way that closely imitates the daily experiences of students on social media. As many adolescents regularly use the social media platform Instagram [1], the tool should resemble this platform.
- **Usability**: InstaClone should be easy to use by both teachers and students without prior introduction and with minor technical knowledge.
- **Flexibility**: InstaClone should be usable in different settings with different time frames. Teachers should be able to conduct projects with InstaClone either consecutively over several days (e.g., on project days) or in a modular way over several weeks (e.g., within regular lessons).
- **Student-level adaptivity**: InstaClone should be designed for students with different levels of knowledge. Students with less technical expertise should have the option of performing tasks at their own pace. In contrast, students with more technical expertise should have the opportunity to complete additional and more complex exercises at the same time.
- **Class-level adaptivity**: InstaClone should be usable in different subjects and applicable in various teaching contexts to foster data literacy. This means that the InstaClone teaching material should cover the curriculum content from multiple subjects as well as content that occurs across subjects, e.g., misinformation or data security. In addition, in-depth data analysis tasks should be provided, especially for computer science classes.
- **Safety and Security**: InstaClone must be a safe space for students. Data on the system should only be accessible within a project, meaning that students should be able to see, for example, the posts of their classmates but not the data of students from other classes or schools. Teachers should be able to set up and enforce platform rules for their classes and delete inappropriate content created by students.

3.2 Tool Description

InstaClone consists of three major components: the student environment, the dashboard, and the teacher panel. The **student environment** imitates the main functionalities of Instagram. Students can create posts and stories, follow other users, like and comment on the posts of others, update their profile information (e.g., biography, age, interests, or gender), and change their user settings (e.g., language). In addition, we have included functionalities to allow students to assume the role of various social media stakeholders, like developers or business companies. For example, students can create their own targeted advertisements, which are then displayed to other students, or change the algorithms that determine the order of posts in their newsfeed. The student environment is visualized in Figure 1.

The data analytics **dashboard** visualizes user behavior on the platform by showing statistics and diagrams on the post and hashtag distributions, user interests, gender and age distributions, or hashtag co-occurrence networks. This allows students to see what kinds of data social media companies collect and provides room for reflection on how such data might be used or misused. To further motivate the students, the dashboard shows student data that are generated through their tool behavior. The idea of providing a data analytics dashboard to students was inspired by social media research aimed at offering the possibility of performing in-depth data analysis. Figure 2 shows a screenshot of the dashboard.

The **teacher panel** is only accessible to teachers. On this panel, teachers can create new student accounts and can inspect, add, or delete data via simple clicks, consequently giving them access to all data posted on InstaClone at all times. In addition, a functionality to upload data, such as accounts, posts, or advertisements, as CSV-tables, has been included. Among other advantages, this functionality allows teachers to create new student accounts quickly and efficiently. Additionally, we furnished a pre-filled dataset containing profiles of "Influencers" and a collection of posts, thereby affording teachers the flexibility to adapt and upload these materials. This supplement augments the dataset available for analysis with the capacity to tailor it to specific tasks in a predefined manner and enables the integration of bot-like behavior. Furthermore, this approach not only offers an initially more engaging platform for students due to the presence of content but also enables the integration of bot-like behaviors. Likewise, we implemented a functionality to download all data created within the InstaClone environment as CSV-tables. These data can then be imported into other (educational) tools, such as Orange [9], to perform more complex data analysis exercises with students.

We offer access to the tool in two different ways. For teachers with fewer technical resources and/or expertise, we offer online access to the tool. When a teacher signs up for the tool via email, a new project instance, consisting of the three components described above, is created on a server at TUM, and the teacher receives

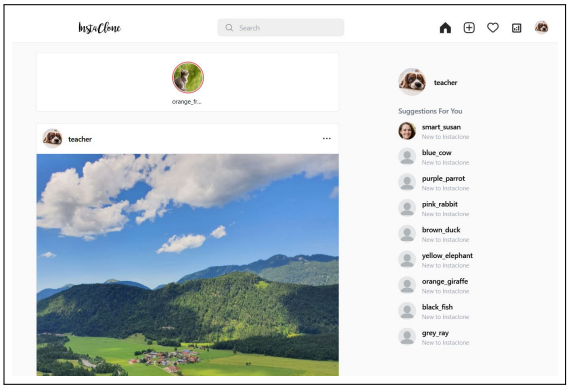

Figure 1: Screenshot of the student environment.

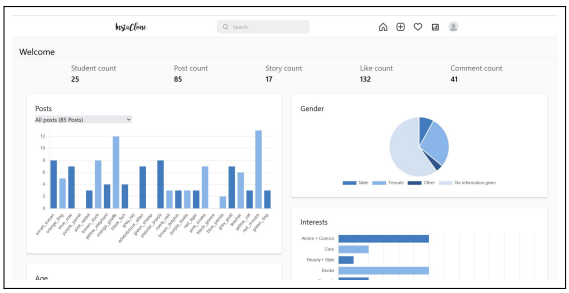

Figure 2: Screenshot of the dashboard.

admin access for this instance. Based on our security concept, the contact person at the university can add or delete a new instance but cannot inspect any data created within the instance, as these data should only be accessible to the teacher and students of a given class. The teacher can then perform all actions mentioned above. For teachers with more technical knowledge and infrastructure, we offer the tool as a docker container, one which can be installed and set up on its own server at the school.

3.3 Teaching Material

Suitable teaching material was developed to support the teacher's instruction using InstaClone in the classroom, including various curricular and socially relevant topics. The teaching material can be used flexibly in different settings. Therefore, it is structured in so-called "micro-units" that can be arranged by the teacher depending on the learning goal or the students' pace. Accordingly, the teaching materials are ideally suited for adaptive teaching and therefore fulfill the requirement of student-level adaptivity. In addition, they also satisfy the criterion of class-level adaptivity as they can be used in different subjects or via cooperation across subjects. In terms of the task format, all chapters include tasks that can be processed by using InstaClone (e.g., "Create new posts on InstaClone, like and comment on existing posts"). Further tasks, like reflective tasks (e.g., "How do you think InstaClone got this information about you?") or participatory tasks (e.g., "Now imagine you are the boss of InstaClone. How would you design your own newsfeed algorithm?"), can be processed directly in the teaching material. Only a few tasks refer to Instagram itself (e.g., "Check out the original Instagram terms of service. Which do you find particularly important?"). In addition to the tasks, the teaching material also contains explanations of, for instance, short definitions of central terms (e.g., filter bubble) or info boxes.

Taking a closer look at the content of the teaching material, the first chapter, *Content Creation*, is necessary for sharing relevant information with the tool and creating data for further analysis. Associated tasks include, for example, "Set up your personal profile on InstaClone (Profile picture, first name, last name, bio...)" or "Think about what you want to share with your future followers." Besides mere content creation as a basis for future data analysis, this chapter also focuses on interaction with other users' content — for example, "Like interesting posts" or "Comment on a post." In the second chapter, *Newsfeed*, students get an overview of relevant postings and are expected to reflect on why they see this content. Additionally, this chapter focuses on misinformation, a socially relevant topic. Students should consider whether they already had contact with misinformation on social media, reflect on what makes misinformation dangerous, and identify misinformation on InstaClone. Chapter 3, *Algorithms*, offers the possibility to try three different newsfeed algorithms in InstaClone. Students should, for example, determine how the algorithms influence their newsfeeds by manipulating them through new postings or likes. Since the project focuses on fostering data literacy competencies, Chapter 4, *Data Analysis*, is the most relevant. First, the students reflect on which data InstaClone has already saved and why it is important for a social media platform to store personal data in general. With a given log file, the students also have the opportunity to analyze the user behavior of three fictional people. Dashboard-related tasks in this chapter allow students to identify fellow students who post frequently, and it is also possible to perform network analytics with hashtag or follower networks. Further, analyzing networks and users on the dashboard makes it possible to post target group-oriented advertisements. Moreover, in-depth teaching materials are available using the data mining software Orange [9] for advanced data analysis. The last chapter, *Data Protection*, deals with tasks regarding personal data, data protection, and privacy. In particular, this is the only chapter that refers to Instagram, where students can engage with their actual privacy settings. Additionally, they should formulate their own terms of use.

4 CASE STUDY

This section describes the participants and the workshop design, followed by the procedure and results of the case study.

4.1 Participants and Workshop Design

The workshops took place in July 2023 within year 9 and year 10 computer science classes at a German high school and were aimed at examining the implementation of the tool in the school context. In addition to testing the authenticity and usability of the tool, another goal was to monitor whether InstaClone motivates students to engage more with data in the context of social media. In total, 45 students participated in the case study (12 female, 32 male, 1 diverse). Eighteen students were in year 9, and 27 students were in year 10. The mean age was 15.79 (SD = 0.86). The workshops

occurred in three different classes and each lasted 90 minutes. After a short introduction, the students logged in to InstaClone. Then, they worked intuitively with the tool, posted pictures, followed classmates, wrote comments, placed personalized advertisements, and evaluated data with the dashboard. Despite having their own InstaClone access, the students in all three classes worked very cooperatively. After one hour of working with InstaClone and the corresponding materials, we held a joint class discussion on several topics, including newsfeed algorithms, personal data, and misinformation. The last 10 minutes of the lesson were used for a short survey.

4.2 Procedure of the Case Study

To evaluate InstaClone in more detail, the students answered self-developed open questions geared toward sharing their experiences with the tool and the teaching material (e.g., "How well did you get on with InstaClone?"). They were also asked about the authenticity (e.g., "Do you think InstaClone is similar to Instagram?") and potential improvements of the tool (e.g., "What would you improve about InstaClone?"). We were also interested in students' activities on social media, especially on Instagram. Therefore, they were asked: "What do you mainly use social media/Instagram for?". In addition, the students rated validated items to the topic "interest in (digital) media" on the ICT from [22] and [15] (e.g., "When you think about your experiences with social media, how much do you agree with the following statements?"). These questions were presented in a 4-point Likert format (1 = "fully agree", 4 = "fully disagree").

4.3 Results

Since, at this time, InstaClone was still a prototype, and the workshops were its first implementation in a classroom setting, experiences and ideas for further improvements to the tool were particularly relevant. Nevertheless, there was an interest in learning more about the students' previous experiences, which social media tools they already used, and how they behaved in a social media setting. As we know from previous research, social media tools are part of students' daily lives. Of the 45 students, 41 agreed with the statement "I like using social media". All but one student reported having previous experiences with social media. A total of 38 students reported using Instagram, followed by Snapchat (n = 35) and TikTok (n = 31). As expected, Facebook and Twitter were not very commonly used in this age group. BeReal and Discord, for example, were mentioned by the students as other social media platforms they use. Of the surveyed students, 35 used social media daily (6 students = min. once per week, 1 student = min. once per month, 1 student = never). Most students agreed that they quickly forget the time when they used social media. Furthermore, the students used social media predominantly to pass the time (n = 39), to exchange ideas with close friends or family (n = 30), to stay informed about the daily news (n = 21), or to receive information about upcoming events (n = 21). Taking a closer look at the activities of the students on Instagram, an identical picture emerged.

As the results were thus far mainly related to students' behavior on social media in general, the focus shifted to the associated InstaClone experience. For this purpose, aspects of the requirement analysis, like authenticity and usability, were taken into account in order to determine whether a realistic and lifelike tool had indeed been developed. With regard to the prototype, the first requirement was authenticity. Therefore, we asked the students: "Do you think InstaClone is similar to Instagram?". From 32 answers, 22 confirmed that InstaClone was similar to Instagram and could be seen as authentic. Another five students even stated that InstaClone and Instagram were very similar: *"Yes, there is definitely a resemblance."* Only one student answered a short *"no."* To gain insights into the usability of InstaClone, the students answered the open-response question "How well did you get on with InstaClone?". From a total of 33 answers, which were predominantly positively formulated, it can be concluded that InstaClone is very self-exploratory to use. While six students answered that they got along very well with InstaClone, 22 responded that it was easy to operate. For example, regarding usability, one student stated: *"I got along very well because you can use it almost like normal Instagram"*. Only one student answered that the usability was bad without further explanation. Moreover, we were interested in the students' experiences working with InstaClone and which aspects they liked the most when using the tool. We received a variety of answers (n = 37) and coded them into 13 categories. The experience students liked the most was posting on the newsfeed (n = 6). Five students mentioned that they enjoyed the lesson. Less frequent categories, comprising four answers each, included students enjoying cooperatively working with other students, enjoying the authenticity of the tool, and appreciating the opportunity to place their own advertisements. The word cloud (see Figure 3) gives an overview of all mentioned categories.

Figure 3: Students' experiences with InstaClone

The major learning goal of InstaClone is to foster students' data literacy through a lifelike social media experience. In this respect, we first asked the students a general question - e.g., whether they became more interested in social media after using InstaClone - and then asked whether they were now more interested in the data used by social media tools. Most of the students responded that the lesson did not increase their interest in social media, likely because it is already a common theme. Nevertheless, InstaClone seemed to demonstrate that the students became more interested in data that are used in a social media context. Twenty students rather agreed on the question of whether InstaClone increased their interest in data, and three students fully agreed. On the other hand, only eight students fully disagreed on this question, and 12 students rather

disagreed. As the duration of the workshop was limited to 90 minutes in each class, the students could not try all parts of the tool and the teaching materials. For this reason, we assumed that we would observe only conditional effects from this brief intervention. However, we asked the students whether they would more consciously use social networks after using InstaClone, to which the majority replied they would not. Nevertheless, we received 26 answers to the question "Did you learn something new about social media?", seven of which were aimed at data literacy. They claimed that they had obtained new insights about data, such as, for example, *"That social media knows more about you than you think"* or *"That everything is stored somewhere"*. Further answers referred to other topics of the teaching material, like misinformation or advertisements. In total, nine students reported that they did not learn something new because of detailed previous knowledge. The last survey question asked whether the students would like to use InstaClone in a classroom setting to learn more about social media. A slider with a 10-step response scale was used for this question. Of 45 responses, 36 had a proficiency between 5 and 10, with the top value range of 9 and 10 being selected particularly frequently (n = 19).

4.4 Classroom Observations

The evaluation that students enjoyed working with InstaClone is identical to our observations. The students were motivated, concentrated on their work, and worked independently with InstaClone. The project members only answered a few short questions, mainly related to functionalities not available on Instagram, such as the dashboard. In addition, they considered what they would like to post and share with their classmates and supported each other in solving problems. The students asked others about their InstaClone name in order to connect with them or comment on newly posted content. Furthermore, we observed that teenagers who were given free access to a social media environment very quickly imitated what they had seen on other platforms like 9gag or Reddit. Despite being informed that all data were visible to the teacher, some students published inappropriate and problematic posts and memes. In addition, we must report that one student was demotivated before the lesson and had no desire to participate in the computer science class. After the first few minutes, during which time the student disturbed the lesson, the student became absorbed with the tool and also seemed very motivated to work with it. Active participation in the collaborative classroom discourse reflected the students' deep engagement with the topic, notably exemplified by their discerning skepticism concerning the personal data amassed by Instagram. The fact that a learning process was initiated was also demonstrated by one student who consulted with us after the lesson on whether we could adjust her privacy settings on Instagram.

5 DISCUSSION

The workshops provided exciting insights into the interactions of students with InstaClone and the accompanying teaching materials. As expected, the students were very familiar with social media and quickly learned how to use the most important functionalities, e.g., publishing posts or creating personalized advertisements. In addition, the students were highly motivated to share their knowledge with classmates and to interact with others on the platform.

At the same time, the workshop revealed approaches for further development of the tool and the teaching materials. First, while the students were very enthusiastic about working with the tool and discovering functionalities independently, multiple students found the printed teaching materials distracting from the online space. They thus did not engage with the exercises and tasks. While the limited duration of the workshop can also explain this behavior, as the students focused more on the tool within this short time frame, it may also explain why some students did not seem to have learned anything new from the workshop. In the future, we plan to develop functionalities that include all exercises within the online system so that students do not have to shift their attention from the "exciting" online space to the more traditional paper-printed exercises, as well as a motivational gamification system, such as rewarding students for completed tasks with virtual tokens. Second, the sometimes inappropriate behavior of students places a great responsibility on teachers working with the tool, as they must closely monitor what is happening on the platform and respond quickly to inappropriate content. While teachers can already view and delete all content within their InstaClone instance, in the future, we will develop features that allow for easier detection of inappropriate content, such as a feature that will enable students to report content to the teacher (remarkably, the workshop students explicitly asked for this function). Furthermore, to counter inappropriate behavior, the teacher could ask the students to jointly develop their own InstaClone Community Guidelines and delete any of their own posts that violate these guidelines. Third, students showed less interest in exercises that asked them to perform an action compared to activities that gave them the freedom to create something themselves. In the future, we will therefore place greater emphasis on creation and discovery tasks. For example, while there is already a task involving learning how the newsfeed algorithm on the platform works, we will add a functionality that will allow students to design their own newsfeed algorithm for InstaClone. Moreover, we plan to develop more in-depth data analysis assignments that will permit students to delve deeper into the role of a data analyst.

6 CONCLUSION

With InstaClone, we developed a social media tool for educational purposes that allows teachers to conduct hands-on classroom activities to foster students' data literacy. We additionally designed accompanying teaching material that can be used flexibly in the classroom. An initial case study of students who tested the prototype for authenticity and usability showed that the students enjoyed working with the tool. Further, InstaClone helped the students achieve the learning goal of acquiring greater insights into data literacy. Based on the students' feedback, we will further develop and continuously test the tool in various classroom settings.

7 ACKNOWLEDGMENTS

We thank all students who were part of our developer team, Leonard Husmann, Tim Kruse, Asli Kök, and Murat Demir, for their excellent work. This work was funded and supported by the Reboot Social Media Lab of the TUM Think Tank.

REFERENCES

[1] Monica Anderson, Emily A Vogels, Andrew Perrin, and Lee Rainie. 2022. Connection, creativity and drama: Teen life on social media in 2022.

[2] Sue Bennett, Karl Maton, and Lisa Kervin. 2008. The 'digital natives' debate: A critical review of the evidence. *British Journal of Educational Technology* 39, 5 (2008), 775–786. https://doi.org/10.1111/j.1467-8535.2007.00793.x

[3] Leanne Bowler, Amelia Acker, Wei Jeng, and Yu Chi. 2017. "It lives all around us": Aspects of data literacy in teen's lives. *Proceedings of the Association for Information Science and Technology* 54, 1 (2017), 27–35. https://doi.org/10.1002/pra2.2017.14505401004

[4] Joel Breakstone, Mark Smith, Sam Wineburg, Amie Rapaport, Jill Carle, Marshall Garland, and Anna Saavedra. 2021. Students' Civic Online Reasoning: A National Portrait. *Educational Researcher* 50, 8 (2021), 505–515. https://doi.org/10.3102/0013189X211017495

[5] David Buckingham. 2007. Digital Media Literacies: rethinking media education in the age of the Internet. *Research in comparative and international education* 2, 1 (2007), 43–55.

[6] Yu Chi, Wei Jeng, Amelia Acker, and Leanne Bowler. 2018. Affective, Behavioral, and Cognitive Aspects of Teen Perspectives on Personal Data in Social Media: A Model of Youth Data Literacy. In *Transforming Digital Worlds*, Gobinda Chowdhury, Julie McLeod, Val Gillet, and Peter Willett (Eds.). Springer International Publishing, Cham, 442–452.

[7] Aayushi Dangol and Sayamindu Dasgupta. 2023. Constructionist Approaches to Critical Data Literacy: A Review. In *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference* (Chicago, IL, USA) *(IDC '23)*. Association for Computing Machinery, New York, NY, USA, 112–123. https://doi.org/10.1145/3585088.3589367

[8] data.europa academy | Open Data courses. 2023. https://data.europa.eu/en/academy [Accessed: 2023-08-11].

[9] Janez Demšar, Tomaž Curk, Aleš Erjavec, Črt Gorup, Tomaž Hočevar, Mitar Milutinovič, Martin Možina, Matija Polajnar, Marko Toplak, Anže Starič, et al. 2013. Orange: data mining toolbox in Python. *The Journal of Machine Learning Research* 14, 1 (2013), 2349–2353.

[10] Dominic DiFranzo, Yoon Hyung Choi, Amanda Purington, Jessie G. Taft, Janis Whitlock, and Natalya N. Bazarova. 2019. Social Media TestDrive: Real-World Social Media Education for the Next Generation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300533

[11] Catherine D'Ignazio and Rahul Bhargava. 2016. DataBasic: Design principles, tools and activities for data literacy learners. *The Journal of Community Informatics* 12, 3 (2016), 83–107.

[12] Julian Dorn. 2023. https://instahub.org/ [Accessed: 2023-08-11].

[13] Andreas Grillenberger and Ralf Romeike. 2017. Key Concepts of Data Management: An Empirical Approach. In *Proceedings of the 17th Koli Calling International Conference on Computing Education Research* (Koli, Finland) *(Koli Calling '17)*. Association for Computing Machinery, New York, NY, USA, 30–39. https://doi.org/10.1145/3141880.3141886

[14] Andreas Grillenberger and Ralf Romeike. 2018. Developing a Theoretically Founded Data Literacy Competency Model. In *Proceedings of the 13th Workshop in Primary and Secondary Computing Education* (Potsdam, Germany) *(WiPSCE '18)*. Association for Computing Machinery, New York, NY, USA, Article 9, 10 pages. https://doi.org/10.1145/3265757.3265766

[15] David Kastberg, Gordon Murray, David Ferraro, Carlos Arieira, Shep Roey, Saida Mamedova, and Yuqi Liao. 2020. *Technical Report and User Guide for the 2016 Program for International Student Assessment (PISA) Young Adult Follow-Up Study. (NCES 2021-020.).* U.S. Department of Education. Washington, DC: National Center for Education Statistics, Washington, DC, USA.

[16] School of Data Evidence is Power. 2023. https://schoolofdata.org/ [Accessed: 2023-08-11].

[17] Jonathan Osborne, Daniel Pimentel, Bruce Alberts, Douglas Allchin, Sarit Barzilai, Carl Bergstrom, Janet Coffey, Brain Donovan, Kari Kivinen, Anastasia Kozyreva, and Sam Wineburg. 2022. *Science Education in an Age of Misinformation.* Stanford Universtiy, Stanford, CA, USA.

[18] The Data Literacy Project. 2023. http://www.thedataliteracyproject.org/ [Accessed: 2023-08-11].

[19] Lindsay Reiten and Susanne Strachota. 2016. Promoting statistical literacy through tuva. *The Mathematics Teacher* 110, 3 (2016), 228–231.

[20] Chantel Ridsdale, James Rothwell, Mike Smit, Michael Bliemel, Dean Irvine, Dan Kelley, Stan Matwin, Brad Wuetherick, and Hossam Ali-Hassan. 2015. Strategies and Best Practices for Data Literacy Education Knowledge Synthesis Report. https://doi.org/10.13140/RG.2.1.1922.5044

[21] Baruch Schwarz and Galit Caduri. 2016. Novelties in the use of social networks by leading teachers in their classes. *Computers & Education* 102 (2016), 35–51. https://doi.org/10.1016/j.compedu.2016.07.002

[22] Mario Vennemann, Birgit Eickelmann, Amelie Labusch, and Kerstin Drossel. 2021. *ICILS 2018# Deutschland. Dokumentation der Erhebungsinstrumente der zweiten Computer and Information Literacy Study.* Waxmann, Münster; New York.

[23] Annika Wolff, Daniel Gooch, Jose J. Cavero Montaner, Umar Rashid, and Gerd Kortuem. 2016. Creating an understanding of data literacy for a data-driven society. *The Journal of Community Informatics* 12, 3 (2016), 9–26. https://doi.org/10.15353/joci.v12i3.3275