big data and
cognitive computing

MDPI

Article

GO-E-MON: A New Online Platform for Decentralized Cognitive Science

Satoshi Yazawa [1,*], Kikue Sakaguchi [2] and Kazuo Hiraki [1]

[1] Graduate School of Arts and Science, The University of Tokyo, Meguro-ku, Tokyo 153-8902, Japan; khiraki@idea.c.u-tokyo.ac.jp
[2] KOMEX (Komaba Organization for Educational Excellence), The University of Tokyo, Meguro-ku, Tokyo 153-8902, Japan; sakaguchi@fye.c.u-tokyo.ac.jp
* Correspondence: yazawa@ardbeg.c.u-tokyo.ac.jp

Abstract: Advances in web technology and the widespread use of smartphones and PCs have proven that it is possible to optimize various services using personal data, such as location information and search history. While considerations of personal privacy and legal aspects lead to situations where data are monopolized by individual services and companies, a replication crisis has been pointed out for the data of laboratory experiments, which is challenging to solve given the difficulty of data distribution. To ensure distribution of experimental data while guaranteeing security, an online experiment platform can be a game changer. Current online experiment platforms have not yet considered improving data distribution, and it is currently difficult to use the data obtained from one experiment for other purposes. In addition, various devices such as activity meters and consumer-grade electroencephalography meters are emerging, and if a platform that collects data from such devices and tasks online is to be realized, the platform will hold a large amount of sensitive data, making it even more important to ensure security. We propose GO-E-MON, a service that combines an online experimental environment with a distributed personal data store (PDS), and explain how GO-E-MON can realize the reuse of experimental data with the subject's consent by connecting to a distributed PDS. We report the results of the experiment in a groupwork lecture for university students to verify whether this method works. By building an online experiment environment integrated with a distributed PDS, we present the possibility of integrating multiple experiments performed by different experimenters—with the consent of individual subjects—while solving the security issues.

Keywords: online experiment platform; distributed personal data store; wearable sensors

check for updates

Citation: Yazawa, S.; Sakaguchi, K.; Hiraki, K. GO-E-MON: A New Online Platform for Decentralized Cognitive Science. *Big Data Cogn. Comput.* **2021**, *5*, 76. https://doi.org/10.3390/bdcc5040076

Academic Editor: Carson K. Leung

Received: 11 November 2021
Accepted: 8 December 2021
Published: 13 December 2021

1. Introduction

Personal data, such as location information and web search history, are the "oil field" of modern data-intensive science [1] and the foundation of the modern online economy [2]. Currently, such information is collected by various applications and is used to optimize services and experiences [3]. However, since these data are collected by different services and companies, it is impossible to combine and analyze all obtained data, which may prevent the realization of potentially innovative services and research. One reason for this is that handling this information requires consideration of individual privacy and legal aspects [4–6]. A distributed personal data store (PDS) has been proposed as a solution to these problems [7,8].

The use of technologies that ensure the security of personal data, such as a distributed PDS for the management of laboratory data, can also be expected to improve the soundness of research procedures in the field of behavioral science. Traditional academic research in which data management and hypothesis generation are the exclusive responsibility of the experimenter is inherently dangerous in the sense that the significance level criterion based on classical hypothesis testing depends on the experimenter's initial hypothesis formulation. If a researcher collects many types of data, the possibility of obtaining a statistic that is significant at the 5% level by chance is also increased. If the researcher wants to handle multiple tests properly, the results that can be claimed depend on "how the independent variables were set (and reported) by the original hypothesis" and at what point the data collection was discontinued. The problem of a "reproducibility crisis" of research results exists in all academic disciplines [9], but it is considered particularly serious in the fields of medicine and psychology. A survey report [10] found that the reproducibility of psychology papers published in top journals is at best 50%, and this has shocked the research community.

Many factors affect the results of research on living organisms and human subjects, and it is impossible to control fully the necessary conditions in a single study, especially in the social sciences [11]. Although the research design strongly depends on the experimenter's expectations, popular hypotheses about topics of interest to many people are periodically replaced, and novel (though less certain) theories are cited as often as those that stand up to more rigorous testing [12]. The main ways to deal with such trust-shaking situations are to (1) disclose hypotheses and plans for data analysis in advance [13], (2) release raw data to withstand follow-up testing by third parties; and (3) shift to open data science (based on large-scale data) that pools acquired data. It has been pointed out that the problem with the conduct outlined in (1) is that it ultimately depends on the researcher's sense of ethics [14]. In addition, as new data acquisition and analysis techniques are being developed at an unprecedented rate, limiting the use of hypotheses and analysis techniques that have already been tested at the time of the research design will reduce the field's competitiveness, and basic research will lag behind commercial and engineering research that does not require such effort. While the call for data disclosure ((2) above) is plausible, there are many data that cannot be obtained if disclosure is a prerequisite, such as medical information, social discrimination, and information that may lead to personal identification. Furthermore, even when a meta-analysis is conducted to verify the reproducibility of previous studies, simple integration is likely to lead to false negatives [15,16].

In biology, it has become mainstream to use bioinformatics and machine learning to refine hypotheses based on the large-scale accumulation of whole genome, expression, and brain function measurement data ((3) above). However, in psychology and education, the data acquisition method is relatively simple, and it is not difficult for people to interpret the data, which in turn raises privacy concerns and tends to limit the discussion to that outlined in (1) and (2) above. If laboratory experiments can be conducted online and users can accumulate their data longitudinally on the PDS, they can share the data after anonymizing such data based on their own intentions. If they can use analysis services to make sense of the data, this would allow them to integrate and take advantage of valuable big data with their consent, which may lead to the honing of useful theories that do not depend on the intentions of experimenters or the trends of academic theories.

In this paper, we propose "GOod Experiment for Mankind Online" (GO-E-MON), which is a novel combination of online experimental tools and a distributed PDS, to conduct laboratory experiments online and allow users to contribute to the acquisition, management, and utilization of their own experimental data. The platform, which has both the task delivery capability of an online experiment platform and the ability to store experimental data in the experimenter's and subject's own storage using a distributed PDS, not only facilitates experiments such as long-term tests to measure cognitive functions and emotions [17–19], but also has the potential to realize more advanced experiments, combining various physiological data such as electroencephalography (EEG) and anthropometry over a long period with the subject's consent. Enabling individuals to manage their own experimental data, MYND [20]—an EEG experiment service provider—uses a consumer-grade EEG monitor that allows individuals to measure and manage their own data. Although MYND is designed for EEG and EEG-related experimental tasks, this platform enables each experimenter to design new feedback programs freely, with the ability to link any measurement device and deliver any experimental task. It is also possible for subjects to understand the value of such data and provide their own experimental data to third-party experimenters as they choose, and this can lead to a virtual pool of experimental big data.

1.1. Online Experiment Platform

Online experiments using the World-Wide Web have a long history, starting with the experiment by Johnson-Laird et al. in 1999 [21], and recently, open-source frameworks and software [22–24] have made it easier to conduct online experiments. Nevertheless, the implementation of online experiments is a time-consuming process for experimenters; it includes writing the program, deploying the program on the machine, and paying attention to security updates. Therefore, instead of deploying their own servers, experimenters can choose to use web services for online experiments. For example, Amazon Mechanical Turk [25] is a service developed to ask people to perform tasks that are difficult for computers to perform, but it can be used for online experiments by registering questionnaires and tasks that test cognitive abilities as tasks. PsyToolkit [26] is an online laboratory tool that specializes in cognitive psychology experiments and can be used for questionnaires and other tasks that are well known in cognitive psychology experiments, such as the N-back task. These can only be performed with questionnaires or tasks that ask for cognitive performance and are difficult to integrate with data from electroencephalographs, heart rate monitors, and so on, as used in laboratory experiments.

As an alternative to various sensors such as electroencephalographs and heart rate monitors used in existing laboratory experiments, currently available consumer-grade wearable sensors can be used. For example, watch-type wearable sensors can be used to obtain information about sleep and pulse waves, and the relationship between these and cognitive and physical performance has been investigated [27,28]. In addition, portable electroencephalographs are available, and experiments have been conducted using these devices [29,30]. By combining these sensors in an online experimental environment, a wide range of experiments can be performed, and new insights can be obtained by integrating these experimental data. Gorilla [31] allows experimenters to write scripts freely, and it is expected to work with these sensors. However, for all of the online experiment services mentioned above, the experimental data are stored in a specific server managed by the service provider. The experimental data obtained by these devices are extremely sensitive, and centralized management on a specific server involves various risks, such as leaks due to cracks.

1.2. Personal Data Stores

As platforms that allow individuals to manage their own data securely, distributed PDSs are attracting attention, and researchers and platformers are exploring and developing them. To utilize personal data, such as location information, search history, call history, and so on, openPDS/SafeAnswers [7] by Pentland et al. realized a method for processing data that uses personal data safely by executing a small code provided by the service provider in the user's environment—a code that performs a certain calculation from personal data and returns only the result to the service provider instead of sharing the personal data. This method is very secure because the database containing the personal data is closed within the user's own device. Another example is Personary [8] by Hashida, which is one of the distributed PDSs. Keeping the data encrypted in the storage of the public cloud service with a secret key that can only be used within the user's own application achieves both availability in the public cloud and security. Personary uses a data model in which data are managed in units of channels, and each channel has a timeline in which items with timestamps are stored. The user's channel and items are encrypted with different symmetric keys and can be disclosed to another user. The users can become "friends" with each other and share a public key corresponding to their private key. The user encrypts the symmetric key of the channel that he/she wants to share with the public key and transfers it to the user with whom they want to share, and that user can then decrypt the shared channel using the channel's key decrypted by his/her private key. Because the channel contains the symmetric key of the items in it, the user can also decrypt the items in it. Both applications focus on how to store and manage the accumulated data safely but do not address how to collect the experimental data obtained in experiments and the distribution methods, which are discussed in this paper.

If users' various experimental data obtained from online experimental services, such as answers to questionnaires, task scores, their own EEG data from wearable devices, and cognitive task scores can be stored in their own PDS, with their consent they can share their data with researchers and service providers. In addition, by storing the experimental data in the experimenter's PDS instead of the centralized online experiment environment, a decentralized management of the data in the online experiment environment can be achieved, thereby reducing the risk of data leakage.

2. Materials and Methods

The goal of GO-E-MON is to add the function of storing experimental data in the distributed PDS to the online experiment platform, realizing the safe collection of experimental data by experimenters, and the management of experimental data by the subjects themselves, which was not previously possible.

GO-E-MON is provided as a Software as a Service-type (SaaS) service that provides an application server with online experiment distribution capabilities. The experimental data obtained from the experiments on GO-E-MON are not permanently stored in the application server, but are sent immediately to the PDS of the experimenter and the PDS of the subject (Figure 1).

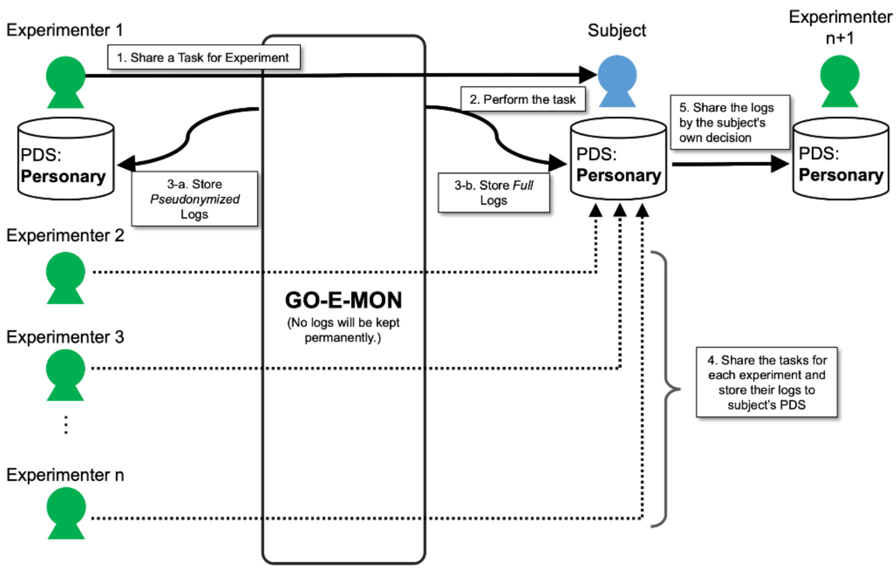


Figure 1. Overview of GO-E-MON: Experimenters can deliver the task to subjects through GO-E-MON. Each experimenter and subject have their own PDS (Personary), and GO-E-MON stores the experimental data in each PDS. Subjects have their own experimental data for the experiments that they have participated in and can provide such data to another experimenter as they please.

GO-E-MON uses Personary [8] as a distributed PDS. Experimenters and subjects can receive the experimental data recorded by GO-E-MON by sharing the experimental data storage channel from their own personal accounts to GO-E-MON. GO-E-MON does not add any information, such as email address or user ID, that can match the subject with the experimental data, so that the experimenter cannot easily identify his own experimental data against the experimental data of other experimenters. Instead, GO-E-MON generates a different pseudonym for each user for each experiment, appends this pseudonym to the experimental data, and stores the data in the PDS of the experimenter. Then, after sending the data to the PDS of the experimenter and the subject, the data are deleted from GO-E-

MON. This mechanism prevents large-scale data leaks due to intrusion into the centrally managed application server, and makes it difficult for the data of the experimenters to be merged with other data if the data are leaked.

The experimental data stored in the subject's PDS can be managed and utilized by the subject at his/her own discretion. For example, the data can be used to provide a log of the subject's past task execution for an experiment that asks about long-term changes in cognitive ability. For subjects to manage and operate these experimental data properly, they need to understand fully the value of the data. Therefore, we conducted a lecture in which students actually received their own experimental data as subjects in the PDS and analyzed their data.

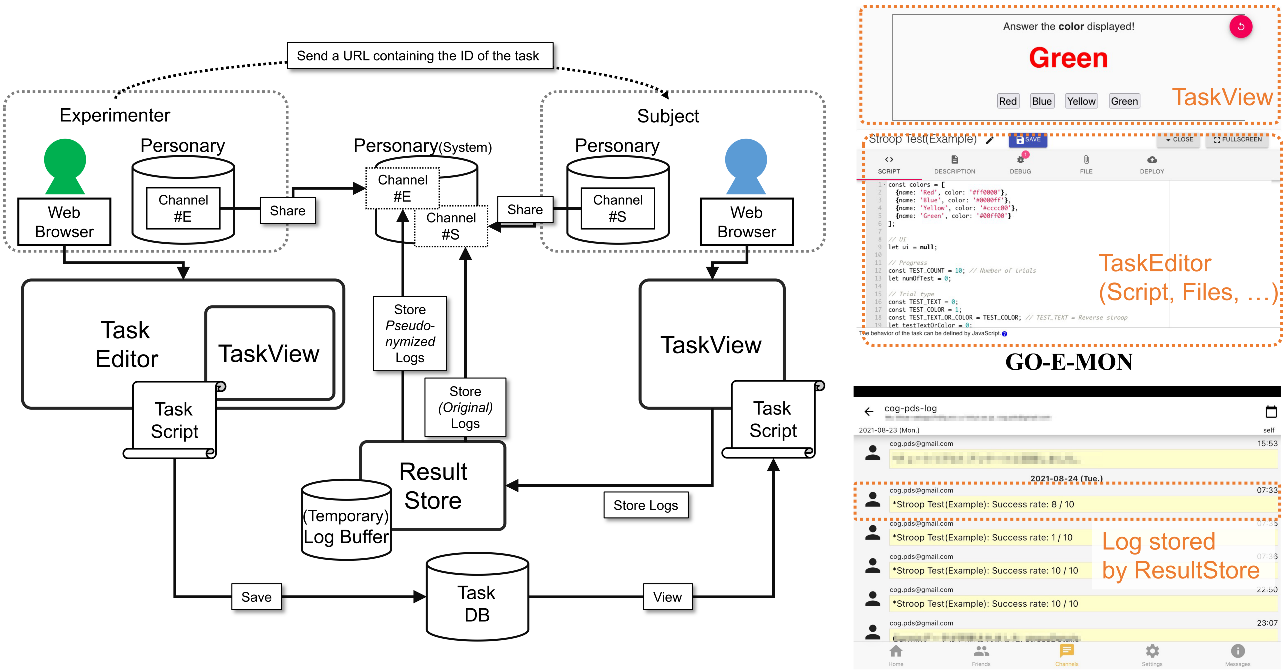
In this study, we developed a prototype service based on the GO-E-MON concept and used it in a groupwork-style lecture on cognitive science to confirm that experimenters can deliver tasks and obtain results, and to investigate whether the subjects can understand the value of the data in the PDS by analyzing it (the service is available at https://goemon.cloud, accessed on 11 November 2021).

2.1. Architecture

GO-E-MON consists of a TaskEditor for writing and managing online experiment codes, a TaskView for executing the created tasks, and a ResultStore for sending the experimental data stored by the TaskView to subjects and experimenters through Personary (PDS; Figure 2). When an experimenter creates a task using the TaskEditor, the task script is saved in the task DB. The experimenter checks the behavior of the task created with the TaskEditor using the TaskView, and then provides the subjects with a URL containing the ID of the task. When the subject opens the URL, the TaskView executes the task based on the task script corresponding to the task ID. When the task is completed, the experimental data are sent to the ResultStore. The ResultStore sends the pseudonymized experimental data to the experimenter and the complete experimental data to the subject. The ResultStore has a Personary account for writing the data, and each experimenter and subject issues a friend registration to the system's Personary account, so that the account can write experimental data to each GO-E-MON-dedicated channel. The system's Personary account can write experimental data for each GO-E-MON channel. If the user does not share the channel with the system's Personary account, GO-E-MON temporarily holds the data in the temporary log buffer of the ResultStore and outputs the experimental data when it detects that the corresponding channel has been shared. There is no timeout, and the ResultStore will continue to hold data until the user's channel is shared.

The experimenter can choose not to receive the execution result at all or to receive the pseudonymized result at the time of task distribution. If the experimenter chooses to receive pseudonymized results, the ResultStore generates a random pseudonym ID that uniquely corresponds to the task ID and GO-E-MON user ID, and the pseudonym ID is appended to the experimental data to pseudonymize the experimental data instead of the GO-E-MON user ID that can be used to uniquely identify the user regardless of the task. Each user can associate his/her Personary account with a user on GO-E-MON on the settings page.

The behavior of the task is defined using JavaScript (ECMAScript) [32]. By enabling the execution of arbitrary scripts, the experimenter can use existing libraries [33] and code, such as an integration with Bluetooth devices [34], although knowledge of scripting is required. Once the experimenter logs into GO-E-MON and creates a new task, he or she can edit the script using the TaskEditor, which runs on a web browser. The TaskEditor also has a file uploader that allows users to upload resources, such as JavaScript, style sheets [35], image files, and so on, and use them in the task. Using the file uploader, experimenters can also upload and use frameworks written in JavaScript, such as jsPsych [21] and WebBCI [24].



System Architecture **Personary**

Figure 2. System Architecture and User Interface of GO-E-MON: The experimenter can define the behavior of a task as a script through the TaskEditor and distribute the URL to the subjects. Once the task is executed, the execution result is saved as experimental data in the Personary accounts of the experimenter and the subject. The data on the experimenter's side is assigned a pseudonym ID, which can be used to identify the user in the task.

The task script consists of the construction processing of the user interface on the TaskView in the web browser and the event processing that occurs at the user interface. The script can store the necessary experimental data in variables and call the API of GO-E-MON at the time when the task is completed to indicate transmission of the experimental data. Task scripts that require users to enter their names and addresses should be prohibited by the terms of use to prevent the addition of any personal identifiers to the experimental data.

2.2. Case Studies

As a case study, 19 students who chose the groupwork-style lecture in the "First Year Seminar" for first-year university students at the University of Tokyo were asked to participate in two experiments using GO-E-MON. Explanations were provided and written consent for participation in the experiments was obtained from the students themselves (all university students were over 18 years old) based on the ethical form 246-17.

The first was an experiment to collect video viewing behavior using GO-E-MON in a video viewing lecture. The percentage of online lectures has increased during the COVID-19 pandemic, making video-based lectures important. It has also been reported that the use of Massive Open Online Courses (MOOCs), which are mainly video-based lectures, has increased rapidly [36]. However, statistics have shown that the completion rate of MOOCs with a large number of registered students is approximately 10% [37], and it is important to analyze the video viewing records and examine any problems with the contents [38,39]. In this experiment, a lecture was conducted using the video viewing behavior recording task created by GO-E-MON to confirm if it was possible to aggregate the experimental data as the experimenter using Personary. Secondly, we provided the students with a number memory task using GO-E-MON, a Garmin Vivosmart 4 wearable sensor [40], and a MUSE-Classic portable electroencephalograph [41], and asked them to perform a group experiment on the topic of memory and brain activity. We analyzed the data, and made a presentation. There are some examples of group lectures on data science [42,43], and in this lecture we verified whether the students can analyze their own data using an online experimental environment by conducting such a lecture using their own data.

Sample Tasks for the Case Studies

As sample tasks, we showed the screen shots of a task for recording video viewing behavior and a task for memorizing numbers implemented in GO-E-MON (Figure 3). The theme of the lecture was human memory, and as part of the lecture, a lecture video on memory techniques was divided into three sessions. The video viewing behavior was recorded using the YouTube IFrame API [44], and the changes in the playback-state of the video player running on the browser were recorded as experimental data and sent to the ResultStore. The number memory task consisted of memorizing 60-digit numbers and repeating them using a keyboard or by touch operation. The jQuery [45] library was used to construct the appearance of the task in JavaScript and to control the data input by users. The subjects were asked to set up the Personary account using their own Google Workspace account for their university organization and to designate the GO-E-MON Personary account as a friend before performing the task. Thereafter, the subjects can store the experimental data obtained by the GO-E-MON task in their own Personary accounts and analyze their own data in the data analysis environment described below.

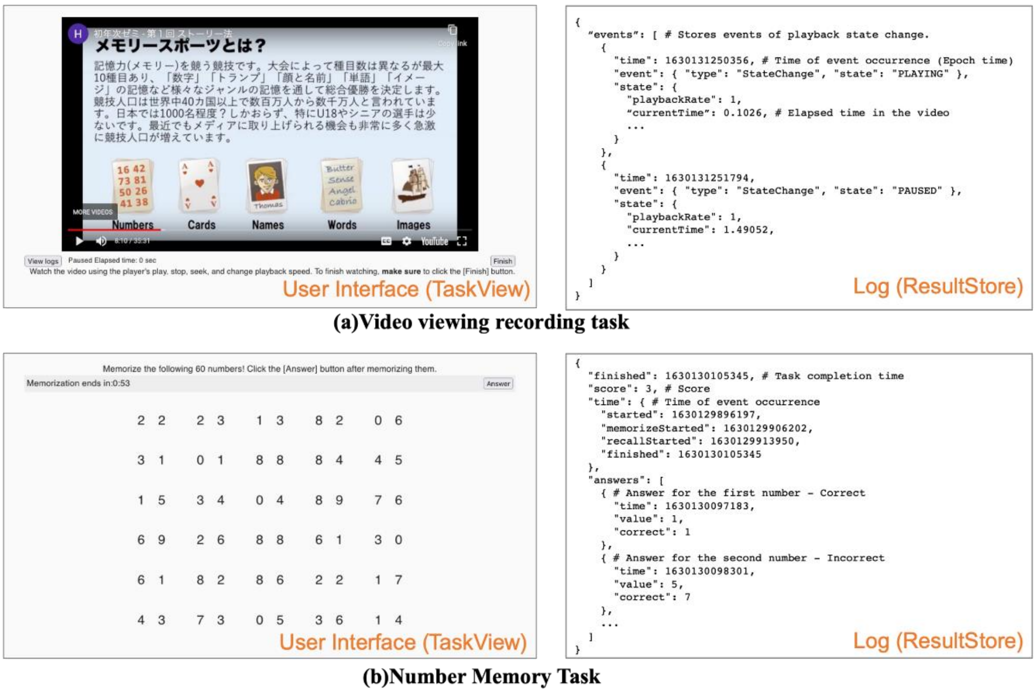


Figure 3. Screenshot of the execution of sample tasks: For the case study, we prepared two tasks: (**a**) a video viewing task and (**b**) a number memory task. Task (**a**) is a lecture video on the theme of memory techniques, and the subject's viewing behavior, such as playing, stopping, and changing the playback speed, is recorded with a timestamp; (**b**) is a task to memorize and repeat 60-digit numbers within a specific timeframe. The number sequence to be memorized is randomly generated, and the content of the answer, identification of an answer as correct or incorrect, and the timing of the answer are recorded together with a timestamp.

We also prepared applications to acquire data from the Garmin vivosmart 4 wearable sensor and the MUSE-Classic electroencephalograph for analysis of the sensor data (Figure 4). For the Garmin vivosmart 4, we implemented a service that stores the data obtained from the Garmin vivosmart 4, including sleep records and heart rate sensor values, in the subject's own Personary through the Garmin Health API [46]. EEG Explorer [47] was used to acquire data from the MUSE-Classic electroencephalograph. The EEG Explorer runs in a web browser and can be connected to the MUSE-Classic through the Web Bluetooth API. The obtained EEG data are stored in the memory of the subjects' PC and can be saved as a file in the subjects' local file system. A task in GO-E-MON can also connect to the MUSE-Classic via Web Bluetooth API, but because of the development time of the task, we decided to use the existing EEG Explorer.

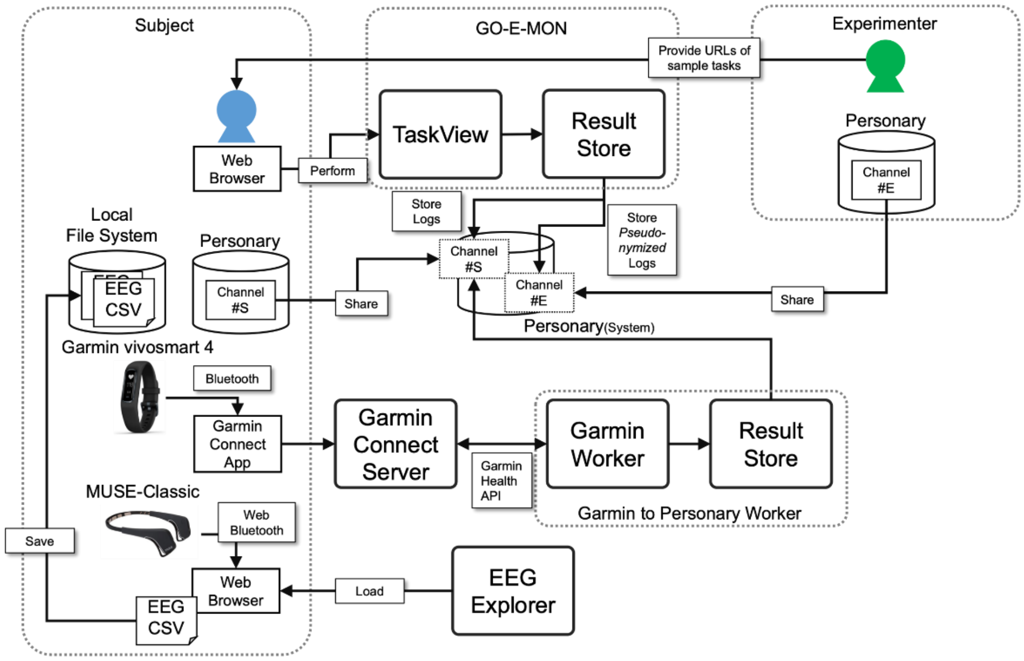


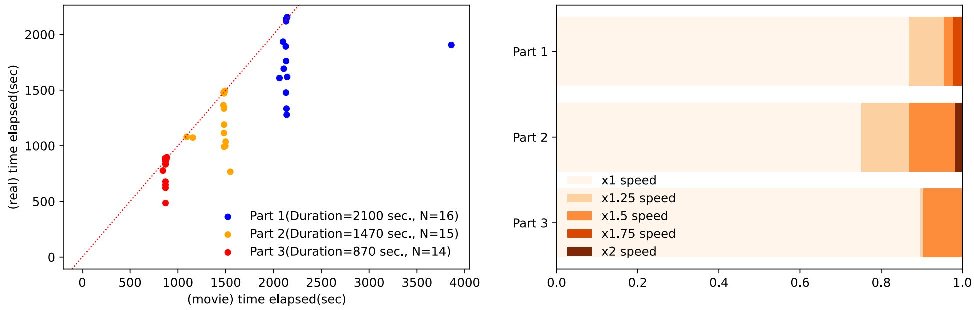
Figure 4. Data flow of the sample tasks: the experimental data of each task created by GO-E-MON and various measurement data by Garmin vivosmart 4 are stored in the subject's Personary account, and the EEG data generated by the MUSE-Classic is stored in the local file system.

The subjects were provided with a computing environment using JupyterHub [48] as a data analysis environment, which can be used in a web browser and configured to allow the subjects to log in with their university accounts. In the JupyterHub, the execution environment of the Jupyter Notebook that can be connected to the Personary account is pre-configured, and the subject follows the instructions in the form of a Jupyter Notebook [49] placed for explanation, and sets up the connection with his or her own Personary account. By following the instructions, the subject can immediately use the analysis examples for each task and sensor prepared in advance by the experimenter. To access the MUSE-Classic data files, we assumed that the files would be uploaded to the JupyterHub environment for analysis. Numpy [50], SciPy [51], and Pandas [52] were installed in the JupyterHub environment for analysis. These libraries can help each group analyze the data on their own, since many documents and technical information are shared on the Internet.

3. Results

3.1. Collection of Video Viewing Behaviors: Collection of Experimental Data by the Experimenter

Three lectures on the theme of memory techniques were provided online as video lectures. The videos of each lecture were provided as separate GO-E-MON tasks, and the URLs of each task were distributed to the students, who watched them for 40 min, 30 min, and 20 min. After each lecture, we confirmed that GO-E-MON sent the video viewing record with a pseudonym ID to the experimenter's Personary account. An example of the data analysis is shown in Figure 5.



(a) Video duration watched and actual time spent **(b) Ratio of playback rate during video viewing**

Figure 5. Example of collecting video viewing behaviors: In the examples of data analysis of video viewing behavior above, (**a**) shows the relationship between the duration of the video watched and the actual time spent watching the video, and (**b**) shows at what playback speed the video was played as a percentage of the elapsed time in the video. The figure shows that many subjects used the seek function or playback speed change function to watch the video in a shorter time than the actual time.

Figure 5a plots the relationship between the duration of the video watched by each subject and the actual time spent watching the video. For each lecture session, there were absentees, and the number of subjects who were able to obtain the video differed. Because the video length is different for each lecture session, the X-axis is concentrated on the video length of each session. However, there are variations in the Y-axis, and the figure shows that the actual time is shortened by using the seek function to advance the playback position or change the playback speed. In Part 1, the total duration of the video watched by one subject was approximately twice as long as the length of the video, because the subject rapidly forwarded through the entire video once and then played it again from the beginning. Figure 5b shows the playback speed at which the video was played as a percentage of the elapsed time of the viewed video. The figure reveals that in the first session, there were few students who played the video at a playback speed of 1.5 times or more, but in the second and third sessions, the percentage of subjects who watched the video at a playback speed of 1.5 times or more increased. In addition, even when the playback speed was changed, the audio was played back with the appropriate pitch processing [53]. In a study using viewing behavior data from Stanford's Lagunita Platform [54], among 21,835 participants, 4345 changed the playback speed when playing any of the videos, and the median value for changing the playback speed was between 1.25 and 1.5 times, indicating that the viewing behavior was generally the same. In this figure, only the entire video is analyzed, and by analyzing the viewing record in more detail, it is possible to identify the parts where many subjects increased the playback speed and analyze the parts that the subjects decided they did not need to watch in real-time, such as the redundant parts of the lecture.

Because each video was created as a different task, a pseudonym ID for the user was generated for each video and assigned to the experimental data. This pseudonym ID was generated randomly and differed for each experiment, such that it cannot be easily assigned to a user. However, all the videos played by the user are stored in the Personary account of each subject. Therefore, by sharing their own channels with other experimenters, the subjects could also check the changes in behavior at the individual level. Without using a dedicated lecture platform such as MOOCs, we confirmed that it was possible to collect video viewing logs in the PDS of the experimenter by delivering the experimental script to the subjects as a GO-E-MON task.

3.2. Experimental Data Collection and Analysis by Students: Collection and Analysis by the Subjects Themselves

After the video lecture, the students were divided into groups of four or five, and groupwork was conducted in four groups. After that, two lectures were delivered to explain how to acquire and analyze the experimental data, and the final presentation was held one week later. The presentations of each group and the types of data used are presented in Table 1. We confirmed that each student group, except for Group 3, was able to analyze the data stored in Personary.

Table 1. Presentations by student groups and types of data used.

	Video Viewing Task	Number Memory Task	Garmin Vivosmart 4	MUSE-Classic (EEG) [1]	What to Analyze
Group 1		✔		✔	Relationship between brain waves and performance on a number memory task after waking and before bedtime
Group 2		✔	✔	✔	Relationship between brain waves and heart rate and performance in a number memory task
Group 3				✔	Relationship between EEG and the execution of their original multi-digit calculation task
Group 4			✔		Relationship between sleep and stress

[1] Only EEG data are stored locally on PCs; other data are stored in Personary through GO-E-MON and shared among group members.

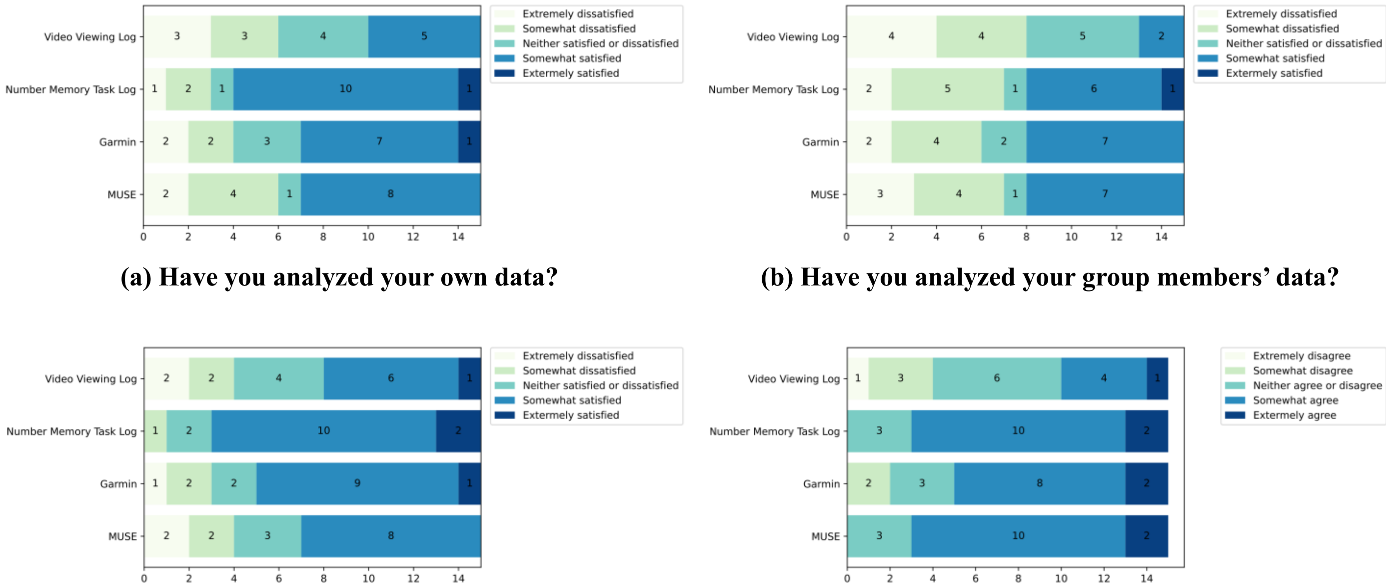
We confirmed that each student group was able to analyze EEG data stored as CSV files in the local file system, as well as the Garmin data and experimental data of the number memory task stored in the Personary. Although not all students were able to analyze the data because of the groupwork style, and it was thought that students who were good at programming and data analysis contributed significantly, we showed that it was possible to realize a lecture in which the subjects analyze their own data in Personary.

In addition, at the end of the lecture, we administered a questionnaire for each data set (Table 2), and the results are shown in Figure 6.

Table 2. Questionnaire after the lecture.

No.	Questions	Answers
1	Please answer whether you feel that you were able to analyze your own data in this lecture.	0: Extremely dissatisfied 1: Somewhat dissatisfied 2: Neither satisfied or dissatisfied 3: Somewhat satisfied 4: Extremely satisfied
2	Please answer whether you feel that you were able to analyze the group members' data in this lecture.	
3	Please answer whether you feel you have understood each of the data obtained in this lecture.	
4	Please answer whether you feel that the data obtained in this lecture are important or not.	0: Extremely disagree 1: Somewhat disagree 2: Neither agree or disagree 3: Somewhat agree 4: Extremely agree

Many subjects responding to the questionnaire stated that they were able to analyze and understand the data (Table 1) which they considered important, but not the video viewing data. The video viewing data were also deemed to be relatively incomprehensible and of low importance, but we think this was because it was not directly connected to the theme of the groupwork. We believe that this result could change if themes that require video viewing records are assigned. We confirmed that the subjects could understand and analyze the data using their own experimental data in Personary and the EEG data obtained in the local file system.



(a) Have you analyzed your own data?

(b) Have you analyzed your group members' data?

(c) Have you understood the data?

(d) Do you feel the data is important?

Figure 6. Questionnaire Results: After the last lecture, a questionnaire (Table 2) was administered, to which 15 students responded. The majority of the participants answered that they could analyze, understand, or find the data important except for the video viewing record, while some participants could not analyze, did not understand, or found the video viewing record unimportant.

4. Discussion

The results of the experiment showed that GO-E-MON, an online experiment environment, can be used in collaboration with a distributed PDS, not only for the experimenter to conduct conventional experiments in which the experimenter collects experimental data, but also for the subjects to collect and utilize the experimental data themselves. In this experiment, although the EEG data were stored locally and separately, it was easy to integrate the data acquisition applications within the GO-E-MON service and consolidate the data into Personary. In this section, we focus on the security of the GO-E-MON service.

4.1. The Importance of the Data Transmission Function to the Experimenter

Within the GO-E-MON service, we implemented a function for sending experimental data to the experimenter's Personary channel from ResultStore in a pseudonymized form. However, it is possible for the experimenter to collect data by storing the data only in the subject's Personary channel and then sharing his or her own channel with the experimenter through Personary, without sending the experimental data to the experimenter's Personary channel. Therefore, in this section, we consider whether it is necessary to send data to the experimenter's Personary account. Currently, to use Personary, users need to have access to a public cloud, especially a Google account, and create and manage their own storage using the Personary application on their own smartphone or PC. However, it is not always the case that the subjects have a Google account [55], and although the ownership rate of smartphones and PCs is high in Japan, it is possible that the subjects do not have one [56], and it may be difficult to ask some subjects to operate Personary. Therefore, we believe that the capability to transmit data to the experimenter is necessary for the application of GO-E-MON in experiments.

Because there is a risk that the experimental data may be leaked because of the inappropriate management of the experimenters' Personary account, we consider it necessary to have data processing options other than pseudonymization as implemented in this project. For example, only statistical information such as the number of times a task has been performed or the number of unique users who have performed a task is sent to the experimenter, and the experimental data itself is shared by the subject, or only a certain number of the same data is transferred, as in k-anonymization [57]. Other than using GO-E-MON in laboratory experiments, it can be used in lectures (the teacher collects the state and behavior of students in class), as in this case study, or individuals can conduct publicly available tasks on their own to measure their cognitive abilities. In situations where GO-E-MON can be applied, future studies could determine what type of aggregation method should be adopted to achieve both convenience and security.

4.2. Countermeasures against Host Cracking

GO-E-MON uses the system's Personary account to write to Personary, and this account information is stored in the application server. When the application server receives the experimental data, it identifies the Personary account of the experimenter and the subject and writes the data in the appropriate channel. Therefore, if the application server is cracked, there is a risk that the experimental data will be referenced before it is sent to Personary. The retention period of the experimental data temporarily held in the application server is as short as a few minutes, and we believe that the risk of long-term experimental data leakage is small. To reduce this risk further, it is possible to reduce the exposed data by distributing the application servers to multiple locations or by making the retention period of the experimental data in the application servers as short as possible. In the traditional centralized data management application server model, host cracking is a significant risk, but by storing data in a distributed PDS, the risk of host cracking is drastically reduced. The data of accounts that have not been set up with a Personary account will be kept in the temporary storage area of the ResultStore until an account is set up, without any timeout. Therefore, it is necessary to delete the data after a certain period. We also plan to set a timeout with considerations such as making it so that the experimenter cannot start delivering the task without connecting the Personary, and asking the subject to declare in advance if they do not intend to connect the Personary.

In addition, there is a risk that if the system's Personary account is cracked, the GO-E-MON channel of all friends can be referenced; as such, strict management of the Personary account is essential. However, if the write-only privilege can be provided in the personal account, it can be applied to the system account to greatly reduce the risk of data leaks, although it is possible to disrupt data integrity by writing a large amount of garbage data. The Personary uses public cloud storage to store data and uses the user's private key to protect the data, which has a passphrase attached to it; therefore, if the public cloud account is stolen, the data cannot be decrypted without the passphrase. However, with the current implementation of Personary, data can be deleted only by the public cloud account, so we plan to provide a mechanism to duplicate the data so that it can be restored from another backup, even if it is deleted.

4.3. Safety of the Experimental Script

In this experiment, no execution restrictions, such as sandboxing [58], were implemented for the task scripts created by users in the GO-E-MON. The intention was to not limit the utilization of devices and sensors [33,34]. As a result, malicious scripts such as cryptocurrency mining [59] and illegal data acquisition [60] may be distributed. We believe that such scripts can be automatically detected to some extent using static detection techniques for malicious codes [61], and testing environments [62]. It may also be useful to implement a feature that allows experimenters to publish their experiment scripts. This platform allows for the distribution of experiment scripts in an executable form with data storage only for another experimenter and subject. In this way, it is possible for another experimenter to conduct a replication experiment using the published experiment script, or for a non-researcher individual to use the script to measure his or her own cognitive abilities.

In addition, GO-E-MON allows the subjects to acquire by themselves the same data that the experimenters are acquiring, so that by checking the data themselves they can verify that the experimenters do not deviate from the experimental description, as in the lecture described earlier. If a task script illegally extracts information and sends it to a server other than GO-E-MON, data that the subject cannot know will be available; as such, task scripts running on GO-E-MON will be prohibited from communicating with servers other than those allowed by GO-E-MON. In this way, the experimenter must always send the same experimental data to the subject's Personary account, as he/she sends to his/her own Personary account. This way, the subject can always confirm what type of experimental data the experimenter has obtained. Furthermore, by standardizing the format of the experimental data [63,64], it can be expected that the experimental data will be actively disclosed to the subjects to ensure their consent for the experiment.

4.4. Paradigm Shift in Behavioral Big Data Accumulation and the Need for Digital Citizenship Education

Until now, IT mega platformers have established a revenue model in which they provide useful tools to users while acquiring and accumulating online behavioral data and sharing analysis data with general service providers to earn advertising revenue. However, the increasing volume of information that can be collected from users online, as well as the increasing power of machine learning to analyze it, has made it possible to identify individuals simply by being connected to the Internet, as seen in web fingerprinting and device fingerprinting [65–67]. However, although such valuable behavioral data can be obtained, services that allow users to utilize them to understand their own characteristics and improve their learning efficiency have not been sufficiently explored.

Individuals' widespread use of PDSs, such as GO-E-MON, to store various behavioral data, will facilitate open data science in the behavioral sciences and avoid replication crises. Furthermore, users can retain the convenience of receiving recommendations on their own PDS application without passing their behavioral history directly to the mega platformer, as was previously provided through cookies. Users will be able to choose recommendations that are useful to them. This strategy is essential for providing personalized education, for example, in the field of education, which deals with a large amount of sensitive data.

We believe that the sharing model of GO-E-MON is one way to implement the right to data portability [68], which is an important element in personal data protection. However, while users themselves can regain the right to retain and manage their own behavioral data, the decision regarding which service providers to provide data to and to what extent is left to the users' judgment. To expand the user base, digital citizenship education needs to be updated so that users can understand their rights to manage data and make profit and loss decisions when sharing with others.

5. Conclusions

With the advent of various portable measurement devices and the improvement of technologies that allow the uses of PC camera images in browsers, it is becoming easier to realize an online experiment environment. These environments are provided by servers connected to the Internet, and the risk of data leakage is always involved. There is also the possibility that innovative research can be conducted by integrating the results of several different experiments, but data connectivity and portability among various platforms have not been established, and data aggregation is not easy from the perspective of protecting the subjects' information. By building an online experiment environment combined with distributed PDSs, GO-E-MON solves these security issues and presents the possibility of conducting a study that integrates multiple experiments conducted by different experimenters with the consent of the individual subjects.

However, GO-E-MON must address several future challenges. Future tasks include: (1) the consideration of the modes of data acquisition by the experimenters to establish a method of acquiring data in a way that excludes personal information from experimental data, such as acquiring only statistical information or anonymizing data, and verifying the scripts made by the experimenter, which do not acquire any personal identification information; (2) improvement of usability to verify that it is possible for a person with limited knowledge to build an experimental environment by using publicly available codes (e.g., EEG data acquisition, real-time data visualization, etc.); (3) performance evaluation of the architecture to verify whether it is possible to ensure sufficient throughput and provide real-time performance when multiple users handle data with a large volume per time, such as EEG data; (4) consideration of the flow of obtaining the agreement when conducting the experiment, for example, in this experiment, we obtained consent forms from subjects face-to-face, not on GO-E-MON, but it is possible to obtain consent online using the functions of Personary; (5) standardization and visualization of the experimental data format, for example, in this study, we analyzed the subjects themselves using JupyterHub, but it is necessary to consider the standardization and visualization of experimental data so that end users can also visualize the data they have obtained in a way that they can understand; (6) establishing a secure way to process data aggregated from distributed PDSs, such as building a service that obtains experimental data from a large number of people, analyzes it, and returns the results to individuals in the form of recommendations, and so on—it is important for subjects to gain complete understanding of the data and choose which of their own data to provide. It is necessary for service providers to collect data from a large number of users in order to return useful recommendations; however, there must be a mechanism for explaining and promoting the subjects' understanding to obtain consent, such as whether to allow the use of one's own data only for the purpose of making recommendations for oneself within the PDS under the user's control, or whether to allow the aggregated data to be stored by the service provider after anonymization and abstraction.

Author Contributions: Conceptualization: K.H.; software and visualization: S.Y.; investigation: K.S., K.H.; writing the paper: S.Y., K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by CREST of JST, grant number JPMJCR18A4.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the University of Tokyo (subject code 246-17; 19 April 2021).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The service can be accessed at https://goemon.cloud (accessed on 11 November 2021).

Acknowledgments: We thank Kashiwakura S., Suzuki K., and Aoki T. for their support during our experiments. Also supported by Cross-ministerial Strategic Innovation Promotion Program (SIP).

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Lazer, D.; Pentland, A.S.; Adamic, L.; Aral, S.; Barabasi, A.-L.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; et al. Social science: Computational Social Science. *Science* **2009**, *323*, 721–723. [CrossRef] [PubMed]
2. Schwab, K.; Marcus, A.; Oyola, J.; Hoffman, W.; Luzi, M. Personal Data: The Emergence of a New Asset Class. Available online: https://www.weforum.org/reports/personal-data-emergence-new-asset-class (accessed on 16 August 2021).
3. Almuhimedi, H.; Schaub, F.; Sadeh, N.; Adjerid, I.; Acquisti, A.; Gluck, J.; Cranor, F.L.; Agarwal, Y. Your Location has been Shared 5,398. Times! A Field Study on Mobile App Privacy Nudging. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; Association for Computing Machinery: New York, NY, USA, 2015; pp. 787–796. [CrossRef]
4. Eagle, N.; Macy, M.; Claxton, R. Network diversity and economic development. *Science* **2010**, *328*, 1029–1031. [CrossRef] [PubMed]
5. de Montjoye, Y.A.; Hidalgo, C.A.; Verleysen, M.; Blondel, V.D. Unique in the crowd: The privacy bounds of human mobility. *Sci. Rep.* **2013**, *3*, 1376. [CrossRef] [PubMed]
6. Narayanan, A.; Shmatikov, V. Robust de-anonymization of large sparse datasets. In Proceedings of the 2008 IEEE Symposium on Security and Privacy (sp 2008), Oakland, CA, USA, 18–22 May 2008; pp. 111–125.
7. de Montjoye, Y.A.; Shmueli, E.; Wang, S.S.; Pentland, A.S. openpds: Protecting the Privacy of Metadata through SafeAnswers. *PLoS ONE* **2014**, *9*, e98790. [CrossRef]

8. Hasida, K. Personal life repository as a distributed PDS and its dissemination strategy for healthcare services. In Proceedings of the 2014 AAAI Spring Symposium Series Big Data Becomes Personal: Knowledge into Meaning, Palo Alto, CA, USA, 24 March 2014; pp. 10–11.
9. Baker, M. 1500 Scientists lift the lid on reproducibility. *Nature* **2016**, *533*, 452–454. [CrossRef]
10. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science* **2015**, *349*, aac4716. [CrossRef]
11. Camerer, C.F.; Dreber, A.; Holzmeister, F.; Ho, T.H.; Huber, J.; Johannesson, M.; Kirchler, M.; Nave, G.; Nosek, B.A.; Pfeiffer, T.; et al. Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nat. Hum. Behav.* **2018**, *2*, 637–644. [CrossRef] [PubMed]
12. Yang, Y.; Youyou, W.; Uzzi, B. Estimating the deep replicability of scientific findings using human and artificial intelligence. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 10762–10768. [CrossRef]
13. Center for Open Science. OSF. Available online: https://osf.io/ (accessed on 16 August 2021).
14. Ioannidis, J.P. Anticipating consequences of sharing raw data and code and of awarding badges for sharing. *J. Clin. Epidemiol.* **2016**, *70*, 258–260. [CrossRef]
15. Doleman, B.; Williams, J.P.; Lund, J. Why most published meta-analysis findings are false. *Tech. Coloproctology* **2019**, *23*, 925–928. [CrossRef]
16. Hajek, T.; Kopecek, M.; Alda, M.; Uher, R.; Höschl, C. Why negative meta-analyses may be false? *Eur. Neuropsychopharmacol.* **2013**, *23*, 1307–1309. [CrossRef]
17. Trampe, D.; Quoidbach, J.; Taquet, M. Emotions in everyday life. *PLoS ONE* **2015**, *10*, e0145450. [CrossRef]
18. Fett, A.J.; Velthorst, E.; Reichenberg, A.; Ruggero, C.J.; Callahan, J.L.; Fochtmann, L.J.; Carlson, G.A.; Perlman, G.; Bromet, E.J.; Kotov, R.; et al. Long-term changes in cognitive functioning in individuals with psychotic disorders: Findings from the Suffolk County Mental Health Project. *JAMA Psychiatry* **2020**, *77*, 387–396. [CrossRef] [PubMed]
19. Cragg, L.; Kovacevic, N.; McIntosh, A.R.; Poulsen, C.; Martinu, K.; Leonard, G.; Paus, T. Maturation of EEG power spectra in early adolescence: A longitudinal study. *Dev. Sci.* **2011**, *14*, 935–943. [CrossRef]
20. Hohmann, M.R.; Konieczny, L.; Hackl, M.; Wirth, B.; Zaman, T.; Enficiaud, R.; Grosse-Wentrup, M.; Schölkopf, B. MYND: Unsupervised Evaluation of Novel BCI Control Strategies on Consumer Hardware. *arXiv* **2020**, arXiv:2002.11754v1.
21. Johnson-Laird, P.N.; Savary, F. Illusory inferences: A novel class of erroneous deductions. *Cognition* **1999**, *71*, 191–229. [CrossRef]
22. de Leeuw, J.R. jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behav. Res. Method* **2015**, *47*, 1–12. [CrossRef]
23. Peirce, J.; Gray, J.R.; Simpson, S.; MacAskill, M.; Höchenberger, R.; Sogo, H.; Grosse-Wentrup, M.; Schölkopf, B. PsychoPy2: Experiments in behavior made easy. *Behav. Res. Method* **2019**, *51*, 195–203. [CrossRef]
24. Stegman, P.; Crawford, C.; Gray, J. WebBCI: An electroencephalography toolkit built on modern web technologies. *Lect. Note Comput. Sci.* **2018**, *10915*, 212–221. [CrossRef]
25. Amazon Mechanical Turk, Inc. Amazon Mechanical Turk. Available online: https://www.mturk.com/ (accessed on 16 August 2021).
26. Stoet, G. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teach. Psychol.* **2017**, *44*, 24–31. [CrossRef]
27. Yazawa, S.; Yoshimoto, H.; Hiraki, K. Learning with Wearable Devices reveals Learners' Best Time to Learn. In Proceedings of the 2018 2nd International Conference on Education and e-Learning, Bali, Indonesia, 5–7 November 2018; Association for Computing Machinery: New York, NY, USA; pp. 87–92. [CrossRef]
28. Okano, K.; Kaczmarzyk, J.R.; Dave, N.; Gabrieli, J.D.; Grossman, J.C. Sleep quality, duration, and consistency are associated with better academic performance in college students. *NPJ Sci. Learn.* **2019**, *4*, 16. [CrossRef]
29. Poulsen, A.T.; Kamronn, S.; Dmochowski, J.; Parra, L.C.; Hansen, L.K. EEG in the classroom: Synchronised neural recordings during video presentation. *Sci. Rep.* **2017**, *7*, 43916. [CrossRef] [PubMed]
30. Sawangjai, P.; Hompoonsup, S.; Leelaarporn, P.; Kongwudhikunakorn, S.; Wilaiprasitporn, T. Consumer Grade EEG measuring sensors as research tools: A review. *IEEE Sens. J.* **2020**, *20*, 3996–4024. [CrossRef]
31. Anwyl-Irvine, A.L.; Massonnié, J.; Flitton, A.; Kirkham, N.; Evershed, J.K. Gorilla in our midst: An online behavioral experiment builder. *Behav. Res. Method* **2020**, *52*, 388–407. [CrossRef] [PubMed]
32. Ecma International. ECMAscript® Specification Suite. Available online: https://www.ecma-international.org/ (accessed on 27 July 2021).
33. Papoutsaki, A.; Sangkloy, P.; Laskey, J.; Daskalova, N.; Huang, J.; Hays, J. WebGazer: Scalable webcam eye tracking using user interactions. In Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI), New York, NY, USA, 9 July 2016; pp. 3839–3845. Available online: http://cs.brown.edu/people/apapouts/papers/ijcai2016webgazer.pdf (accessed on 12 December 2021).
34. Web Bluetooth Community Group. Web Bluetooth Draft Community Group Report, 21 July 2021. Available online: https://webbluetoothcg.github.io/web-bluetooth/ (accessed on 27 July 2021).
35. W3C. Cascading Style Sheets Specification. Available online: https://www.w3.org/Style/CSS/Overview.en.html (accessed on 27 July 2021).
36. Impey, C.; Formanek, M. MOOCS and 100 days of COVID: Enrollment surges in massive open online astronomy classes during the coronavirus pandemic. *Soc. Sci. Humanit. Open* **2021**, *4*, 100177. [CrossRef] [PubMed]

37. Jordan, K. MOOC Completion Rates: The Data. Available online: http://www.katyjordan.com/MOOCproject.html (accessed on 15 August 2021).
38. Kim, J.; Guo, P.J.; Seaton, D.T.; Mitros, P.; Gajos, K.Z.; Miller, R.C. Understanding in-video dropouts and interaction peaks in online lecture videos. In Proceedings of the First ACM Conference on Learning @ Scale Conference (L@S '14), New York, NY, USA, 4 March 2014; pp. 31–40. [CrossRef]
39. Liang, J.; Li, C.; Zheng, L. Machine learning application in MOOCs: Dropout prediction. In Proceedings of the 11th International Conference on Computer Science & Education (ICCSE), Nagoya, Japan, 23–25 August 2016; pp. 52–57. [CrossRef]
40. Garmin Japan. Garmin Vívosmart®4. Available online: https://www.garmin.co.jp/products/intosports/vivosmart-4-black-r/ (accessed on 27 July 2021).
41. InteraXon, Inc. Muse 2: Brain Sensing Headband. Available online: https://choosemuse.com/muse-2/ (accessed on 27 July 2021).
42. Donoghue, T.; Voytek, B.; Ellis, S.E. Teaching creative and practical data science at scale. *J. Stat. Data Sci. Educ.* **2021**, *29* (Suppl. 1), S27–S39. [CrossRef]
43. Millman, K.J.; Brett, M.; Barnowski, R.; Poline, J.B. Teaching computational reproducibility for neuroimaging. *Front. Neurosci.* **2018**, *12*, 727. [CrossRef] [PubMed]
44. Google Developers. Youtube IFrame Player API. Available online: https://developers.google.com/youtube/iframe_api_reference (accessed on 28 July 2021).
45. Open JS Foundation. jQuery: The Write Less, Do More, Javascript Library. Available online: http://jquery.com/ (accessed on 28 July 2021).
46. Garmin Ltd. Garmin Health API. Available online: http://developer.garmin.com/garmin-connect-api/overview/ (accessed on 28 July 2021).
47. Shaked, U. EEG Explorer. Available online: https://github.com/urish/eeg-explorer (accessed on 28 July 2021).
48. Project Jupyter Team. JupyterHub. Available online: https://jupyterhub.readthedocs.io/ (accessed on 28 July 2021).
49. Kluyver, T.; Ragan-Kelley, B.; Pérez, F.; Granger, B.; Bussonnier, M.; Frederic, J.; Kelley, K.; Hamrick, J.; Grout, J.; Corlay, S.; et al. Jupyter Notebooks—A publishing format for reproducible computational workflows. In *Positioning and Power in Academic Publishing: Players, Agents and Agendas*; IOS Press: Amsterdam, The Netherlands, 2016; pp. 87–90. [CrossRef]
50. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J. Array programming with NumPy. *Nature* **2020**, *585*, 357–362. [CrossRef]
51. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Method* **2020**, *17*, 261–272. [CrossRef]
52. McKinney, W. Pandas: A Foundational Python Library for Data Analysis and Statistics. *Python High Perform. Sci. Comput.* **2011**, *14*, 1–9.
53. Powale, P. YouTube Engineering and Developers Blog Variable Speed Playback on Mobile. Available online: https://youtube-eng.googleblog.com/2017/09/variable-speed-playback-on-mobile.html (accessed on 15 August 2021).
54. Lang, D.; Chen, G.; Mirzaei, K.; Paepcke, A. Is faster better? A study of video playback speed. In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (LAK '20), Frankfurt, Germany, 23–27 March 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 260–269. [CrossRef]
55. ZDNET. Google: G Suite Now Has 2 Billion Users. 13 March 2020. Available online: https://www.zdnet.com/article/google-g-suite-now-has-2-billion-users/ (accessed on 28 July 2021).
56. Statistics Bureau, Ministry of Internal Affairs and Communications of Japan. Diffusion of the Internet Ministry of Internal affairs and Communications in Japan, Statistical Handbook of Japan. Available online: https://www.stat.go.jp/english/data/handbook/c0117.html (accessed on 27 July 2021).
57. Sweeney, L. k-anonymity: A model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl. -Based Syst.* **2002**, *10*, 557–570. [CrossRef]
58. Terrace, J.; Beard, S.R.; Katta, N.P. JavaScript in JavaScript (js.js): Sandboxing Third-Party Scripts. In Proceedings of the 3rd USENIX conference on Web Application Development (WebApps'12), Boston, MA, USA, 13 June 2012; p. 9. Available online: https://www.usenix.org/conference/webapps12/technical-sessions/presentation/terrace (accessed on 12 December 2021).
59. Huang, D.Y.; Dharmdasani, H.; Meiklejohn, S.; Dave, V.; Grier, C.; McCoy, D.; Savage, S.; Weaver, N.; Snoeren, A.C.; Levchenko, K. Botcoin: Monetizing stolen cycles. In Proceedings of the Network and Distributed System Security Symposium (NDSS), San Diego, CA, USA, 23 February 2014. [CrossRef]
60. ReversingLabs. Groundhog Day: NPM Package Caught Stealing Browser Passwords. 21 July 2021. Available online: https://blog.secure.software/groundhog-day-npm-package-caught-stealing-browser-passwords (accessed on 28 July 2021).
61. Curtsinger, C.; Zorn, B.; Seifert, C. ZOZZLE: Fast and Precise in-Browser JavaScript Malware Detection. In Proceedings of the 20th USENIX Secur. Symp. (USENIX Secur. 11), San Francisco, CA, USA, 12 August 2011; Available online: https://www.usenix.org/conference/usenix-security-11/zozzle-fast-and-precise-browser-javascript-malware-detection (accessed on 12 December 2021).
62. Musch, M.; Wressnegger, C.; Johns, M.; Rieck, K. Web-based Cryptojacking in the wild. *arXiv* **2018**, arXiv:abs/1808.09474.
63. Soldatova, L.N.; King, R.D. An ontology of scientific experiments. *J. R. Soc. Interface* **2006**, *3*, 795–803. [CrossRef] [PubMed]

64. Ocheja, P.; Flanagan, B.; Ogata, H. Connecting decentralized learning records: A blockchain based learning analytics platform. In Proceedings of the 8th International Conference on Learning Analytics and Knowledge (LAK '18), Sydney, Australia, 7–9 March 2018; Association for Computing Machinery: New York, NY, USA, 7 March 2018; pp. 265–269. [CrossRef]
65. Nikiforakis, N.; Kapravelos, A.; Joosen, W.; Kruegel, C.; Piessens, F.; Vigna, G. Cookieless monster: Exploring the ecosystem of web-based device fingerprinting. In Proceedings of the 2013 IEEE Symposium on Security and Privacy, Berkeley, CA, USA, 19–22 May 2013; pp. 541–555. [CrossRef]
66. Formby, D.; Srinivasan, P.; Leonard, A.; Rogers, J.; Beyah, R. Who's in control of your control system? Device fingerprinting for cyber-physical systems. In Proceedings of the 2016 Network and Distributed System Security Symposium, San Diego, CA, USA, 21 February 2016. [CrossRef]
67. Merchant, K.; Revay, S.; Stantchev, G.; Nousain, B. Deep learning for RF device fingerprinting in cognitive communication networks. *IEEE J. Sel. Top. Signal Process.* **2018**, *12*, 160–167. [CrossRef]
68. Publications Office of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of such Data, and Repealing Directive 95/46/EC (General Data Protection Regulation). Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504&qid=1532348683434 (accessed on 28 July 2021).