Privacy Measures for
The Big Data Show

Version 0.8, 17th April 2023

**Contents**

[1. Introduction 2](#_Toc51926064)

[2. Organisational Structures 3](#_Toc51926065)

[2.1. TBDS Ethics Committee 3](#_Toc51926066)

[2.3. Civic Digits Team 4](#_Toc51926068)

[2.4. Two Tails 4](#_Toc51926069)

[3. Stakeholders 4](#_Toc51926070)

[3.1. Internal Stakeholders 4](#_Toc51926071)

[Co-Producers 4](#_Toc51926072)

[Partners 4](#_Toc51926073)

[3.2. External Stakeholders 4](#_Toc51926074)

[4. Policy 4](#_Toc51926075)

[4.1. Child Protection 4](#_Toc51926076)

[4.2. Privacy Statement 5](#_Toc51926077)

[4.3. Terms of Service for the app 5](#_Toc51926078)

[4.4. IP and Licensing 5](#_Toc51926079)

[Confidentiality 5](#_Toc51926080)

[4.5. Project Evaluation 5](#_Toc51926081)

[Quantitative evaluation 5](#_Toc51926082)

[Qualitative evaluation 6](#_Toc51926083)

[4.6. Social Media 6](#_Toc51926084)

[4.7. Customer Relationships Management (CRM) of Data for schools 6](#_Toc51926085)

[5. Functionality and Data Handling of the app 6](#_Toc51926086)

[5.1. System Data Flow 6](#_Toc51926087)

[5.2. App Permissions 7](#_Toc51926088)

[5.3. App Data Collection 9](#_Toc51926089)

[5.4. Live looking over the shoulder of other audience members 9](#_Toc51926090)

[Privacy Risk Register 11](#_Toc51926091)

[Appendix A: Security measures for the App “Super Swipe” and its control system 13](#_Toc51926092)

[Appendix B. The Big Data Show Website Security and Privacy 15](#_Toc51926093)

# Introduction

**Due to the Covid-19 pandemic, the plans for The Big Data Show were cancelled in March 2020. This document updates ‘Privacy Measures’ v.05, which was written for the live show, to reflect the fact that the show is now being made available online.**

This document has been prepared and approved by the Ethics Committee, taking on advice and guidance from key partners including Police Scotland. It presents the privacy and data protection measures adopted by *The Big Data Show* (TBDS).

TBDS will be delivered through a website [https://thebigdatashow.online](https://thebigdatashow.online/) (“the website”). The website has been designed and developed by Foxdog studio and the new version of the show will also involve the use of a smartphone app (“the app”), called *Super Swipe.* The app is being designed and implemented with the help of Two Tails Studio, a specialist in games and software development. Documents outlining the security measures for the app and the website are attached to this document in the Appendices.

Up to 30 shows will be presented in Perth, Glasgow and Edinburgh through the website and app. In-school workshops will be delivered to each participating school when it is safe to do so.

The producers of TBDS aim to improve the data literacy and cyber resilience of pupils who take part in the project and particularly to inspire young women and people of colour to consider careers in digital technologies.

Table . Outcomes for: children, teachers, education system, government, Civic Digits

|  |  |  |
| --- | --- | --- |
| **Short-term outcomes (immediate):**Awareness of cyber issues is raisedEthical hacking careers are better understoodLearners achieve SCQF creditTeachers feel supported to deliver curriculumCurriculum delivery is recognised as being creatively enhancedFunders’ requirements are fulfilled | **Medium-term outcomes (e.g., after six months):**Cyber awareness is raised Children have progressed onto further cyber-related/digital or STEM learning opportunities as a result of TBDS Interest in cyber/digital/STEM careers is increasedA model of creative delivery of STEM learning is available, with evaluationTeachers have explored more opportunities to bring creative digital/STEM learning into the curriculum (eg cross-departmentally within the school)Scotland’s progressive approach in creatively teaching cyber/digital awareness is recognised | **Long-term outcomes (after a year):**This approach is common and growing in Scotland’s schoolsChildren are on the path to further cyber/digital/STEM learningThe model has run successfully outwith ScotlandWe have concrete evidence of an increase in learning and achievement and aspiration from the children in this areaWe have evidence of increased capacity in schools and between schools and other organisations to deliver curriculum in a creative way. |

This risk assessment has been written by Dr Clare Duffy, Artistic Director of Civic Digits with Rupert Goodwins, co-writer of *The Big Data Show* and Creative Technology Officer of Civic Digits

This is a dynamic document and will be reviewed on at least a bi-annual basis.

# Organisational Structures

## TBDS Ethics Committee

|  |  |  |
| --- | --- | --- |
| **Person** | **Affiliation** | **Role** |
| (NC) Dr. Natalie Coull | Abertay University | Ethical Hacking Skills Advisor |
| (DS) Daniel Sellers | Scottish Government | Cyber Resilience Learning and Skills Policy Adviser |
| (MW) Mary Wilson | Police Scotland | Cybercrime Harm PreventionPartnerships, Prevention and Community Wellbeing |
| (HP) Hilary Philips | Youth Link | Senior Development Officer (Digital Youth Work) |
| (CT) Cheryl Torano | cyberQuarter | Business Development Manager at Abertay cyberQuarter. |
|  |  |  |

## Civic Digits Team

|  |  |
| --- | --- |
| **Person** | **Role** |
| (CD) Clare Duffy | Creative Director, Civic Digits |
| (JT) Janet Tam | Operations Manager |
| (RG) Rupert Goodwins(JG) John Grieve | Creative Technical OfficerProject Manager |

##

Civic Digits Board

|  |  |
| --- | --- |
| **Person** | **Role and expertise** |
| Suzy Glass | Chair Producer/Facilitator |
| Liz Green | Digital Youth Work: Youth Link |
| Ciara MitchellAileen Appleyard | Cyber Advocacy Director: Quorum CyberHead of Account Team (College of Arts Humanities and Social Sciences at Edinburgh Innovations University of Edinburgh |

## Two Tails

Director David Mitchell: Delivery of the app for use in live performance and workshops.

# Stakeholders

## Internal Stakeholders

### Producers

* Civic Digits C.I.C

### Partners

* cyberQuarter Abertay University
* QA (Corporate Sponsors)
* SSE (Corporate Sponsors)
* TBDS Ethics Committee (Details given above)

## External Stakeholders

* Schools: approximately 20 across Dundee
* Jenni Mackay. Education Support Officer, Digital Learning & Technologies | Dundee City Council
* Current funders: Scottish Government. Cyber Resilience Unit, QA, SSE. Applications in with 9 trades and DigitalXtra
* Creative Informatics. The University Edinburgh.

# Policy

## Child Protection

Clare Duffy is a registered STEM Ambassador and therefore is a member of Disclosure Scotland’s Protecting Vulnerable Groups (PVG) Scheme. Civic Digits are currently looking into working with an umbrella organisation to provide PVG checks for all employees working with young people.

## Privacy Statement

Civic Digits has a Privacy Statement which conforms with the requirements of the

General Data Protection Regulation (GDPR) as enacted into UK law by the Data Protection Act 2018. Civic Digits will hold data about teachers, schools and public audience members who consent to it and will not share any of this personal data with any third parties (except where required to by law).

Civic Digits will share anonymised data about audience and participant numbers, age groups, and geographical location of performances for research and evaluation purposes, for example with funding bodies such as Creative Scotland and with academic institutions. The Privacy Statement is available on a website before the workshops begin. ‘Plain language’ Privacy Statements will also be given out to audiences at the end of each show to make it clear that no personal data has been seen or collected by Civic Digits though the process of delivering the show.

## Terms of Service for the app

Civic Digits have written a statement of Terms of Service for using the app, with the guidance and final approval of the Ethics Committee. The aim of the Terms of Service is to replicate as much as possible the typical experience of a user downloading and using an app from Google Play or the Apple App Store. We also want to use this document to show how the language and form of these terms can be very hard to understand by including absurd and/or silly statements. These absurd statements will be referred to in the workshop/TBDS and there will be further suggestions of ways to explore them as part of our challenge to the users to think critically about their relationship with the makers of apps.

## IP and Licensing

The intellectual property for the app is shared equally between Civic Digits C.I.C and RG. Civic Digits C.I.C and RGpropose to release any source code for the app that does not already fall under previous licences under the [GPL-3.0 License.](https://www.gnu.org/licenses/gpl-3.0.en.html) This source code will either be available via the website or by email request.

### Confidentiality

All participants will be informally asked to agree not to disclose the way this project and performance works. Much of the delight and magic of the experience is based on testing what the participants do in a typical online situation. Alerting the audience to ‘the game’ would genuinely spoil the experience. However, we will not require participants to sign a non-disclosure agreement.

## Project Evaluation

### Quantitative evaluation

The project will offer pupils the opportunity to achieve the SCQF-credit rated award certificates. This data will be managed by the participating schools. Civic Digits will have a code that relates to a school register and will have no access to pupils’ personal data.

This ‘Introduction to Cyber Resilience and Digital Citizenship’ course will involve participation in two workshops, TBDS and a test. Over 60% pass mark.

Workshop leaders will ask participants to fill in an evaluation form at the beginning and end of the day. This data will be anonymous and will be processed by Prof. Alasdair Rutherford (BRAW DATA)

### Qualitative evaluation

The workshop leaders delivering the shows will create a ‘show report’ for every day in the roll out at the cyberQuarter. This will reflect the whole team’s assessment of how the participants responded to each element of the workshops. This will not include any personal data from participants.

Teachers will be asked to give qualitative evaluation of the delivery of TBDS programme on the day. This will be recorded as part of the ‘show report’.

## Social Media

TBDS has a Twitter, Instagram and Facebook account. Social media platforms sell, use and/or share data which can identify individuals and threaten their privacy and security for financial and potentially other reasons. Civic Digits’ aim is that the work we do will contribute to the debate about the role and responsibility of us as citizens and the organisations who offer these platforms to know about how they use (personal) data and how that affects privacy and security online and offline.

## Customer Relationships Management (CRM) of Data for schools

Civic Digits will keep a CRM-style database of the schools and teachers it has contact with to maintain communication throughout the project. In some cases, this will include personal telephone numbers for teachers. Consent to hold this data will be requested with booking instructions.

#

# Functionality and Data Handling of the app

For the role-out May/June and September 2023 in Dundee the participants will be p7s. Because of the age of participants, we have agreed with Jenni Mackay at Dundee council that Civic Digits will provide all the phones for participants, to ensure there is no question about 1) equality of access to the experience 2) data security being a concern.

The phones will be secure and have updated software.

It is our experience that participants are likely to take photos with phones they are given and it will be part of the workshop leaders process to delete all photos at the end of the day.

## System Data Flow

An overview of the data flow is shown in the accompanying diagram. In this diagram, the app described elsewhere in this document is referred to as “the Phone App”.

The Control App does not ever interact directly with the Phone App, but sends data to the Server. The Server stores this data in a database. During a TBDS performance, the Phone App connects regularly with the Control Application to ‘ask’ it what scene it should be displaying, whether it should be playing a sound effect or triggering any other ‘magic tricks’. The app also sends a game score to the server and receives a high score table back, where the phone is known to the server by a long, random key that exists only for the duration of the show and for this single purpose: nothing else about the phone or its user is known to the server software. Data extracted from the database by the Server is sent to the Phone App as a text object (in JSON format).

RG as CTO wrote a security design specification. In the appendix to this document Two Tails have answered each point of this specification. RG will also have eyes on the code for the app.

## App Permissions

The app will ask the user to allow access to data such as GPS. However, the app will not in fact be able to collect this data, either in the core code or via Android data access permissions, as the functionality required to collect this data will not be programmed into the app. According to the Android documentation,[[1]](#footnote-1) these are permissions

to access sensitive user data (such as contacts and SMS), as well as certain system features (such as camera and internet). Depending on the feature, the [Android] system might grant the permission automatically or might prompt the user to approve the request.

We consider it unethical to allow our app to have the functionality to request data access in this way, since if there was a malicious act on the phone, the data could be exploited.

When first downloaded, the app asks for permissions to access storage and network connections.

This is a standard for Apple and Google Play and beyond our ability to change. However, our app will not in fact access these areas on the user’s phone. If we made this clear before the show, we believe this would destroy the potential learning impact of the experience. For this reason, we will distribute a simple and ‘Plain English’ Privacy policy/declaration to all pupils after the show.

## App Data Collection

No user data is transferred from the user’s phone via the app except a single high score figure known by a long, random, temporary key shared between the phone app and the server. The only information that will be recorded from the user in the app is their GPS location. This data will be stored locally (within the user app file structure on the user's phone) and never leaves the phone.  This GPS data will be deleted as soon as it has been used in the show.

For players of *Super Swipe* who do not attend the show, the app will be updated at the end of each tour. Every update or re-installation automatically deletes previously collected data.

All analytic data recording has been disabled for the app itself. Some analytic data may be recorded by the Operating System (Android or iOS) or by the devices themselves (for example the Apple App Store may record when the application was downloaded) but this is standard for all phone apps, is outside of our control and is compliant UK Data Protection legislation.

There will be a competition at the start of the show to find ‘The Swipe Champion’. This will be announced by a message sent to the highest scorer and will not require any personal data to be seen or collected.

## Live looking over the shoulder of other audience members

There is the risk that an audience member can see data ‘over the shoulder’ of a class mate when the GPS data of the user is presented to them in the form of a map of where the user was when they played *Super Swipe*. The map is specifically designed to show the user where they were, but to make it difficult for a glance to reveal very much information. Each point on the map is visited for only a few seconds and there are no place identifies such as street names or organisation addresses (for example, an LGBT youth club). The resolution of the map is not particularly fine grained, each view of the map representing approximately 0.25 km.

Privacy Risk Register

Note: Risks are ranked on a 1—5 scale. Exposure is an estimate of the combined effect of Impact and Likelihood.

|  |  |  |  |
| --- | --- | --- | --- |
| **Risk description** | **Inherent Privacy Risk** | **Options for avoiding or mitigating this risk** | **Risk Owner** |
| **Impact**  | **Likelihood** | **Exposure** |  |  |
| Reputational risk to Civic Digits if there is a perceived risk (e.g., via press or social media) of privacy invasion. | **Mild (2)** | **Mild (2)** | **Mild (2)** | Ensure demonstrable adherence with data protection legislationAsk a highly qualified ethics committee to support the project and think through all of the risks. | **CD** |
| Spoof versions of the app are installed on the Apple and/or Android app stores and allow malicious code to be installed on users’ phones. | **Low (1)** | **Low (1)** | **Low (1)** | Ensure clear branding and make it as fool proof as possible for TBDS participants to only download the correct appOn a par with any other app in the ecosystem. It is difficult to assess impact as a spoof app could do anything it liked, but TBDS presents no more risk here than do other apps. | **CD/RG** |
| Faulty implementation of the app allows personal data to be transferred to the system Server. | **Low (1)** | **Low (1)** | **Low (1)** | **Server is protected by strong passwords, and is only physically accessible to authorised members of the team.**TBDS does not collect any personal data on the phone beyond some GPS tracking. The architecture of the system, the configuration of the server and the functionality of the back-end system do not collect or store any data of any kind during operation, so it's not as if there's a pathway which could malfunction - there is no pathway.  | **RG/CD** |
| Insecurities with and on the server allow third parties to tamper with the Server routines to send malicious payloads to the user app. | **Low (1)** | **Low (1)** | **Low (1)** | **The server is a standard regularly patched currently supported Linux instance hosted on AWS and managed by a reputable hosting company, with whitelisted 2FA admin access to the hosting service and SLL access to the instance itself.**  | **CD/RG** |
| Insecure data transfer between user app and Server is breached and malicious payloads are transferred to the user’s phone. | **Mild (2)** | **Mild (2)** | **Mild (2)** | No user data is transferred between the server and the mobile app, and there is no pathway for malicious data onto the app. There is a risk that an unknown vulnerability in the app could be exploited by a custom-designed attack, but that's true for all apps. As the total expected audience for the app is small, compared to most apps, and the audience is mostly school children with little economic or political impact. It is considered unlikely that it would repay the resources required to attack it.  | **CD/RG** |
| Malicious code on user’s phone is able to access private data stored on the app’s local file store. | **Low (1)** | **Low (1)** | **Low (1)** | The only private data stored is recent GPS tracking, which is hard to exploit in any harmful fashion for our audience, and thus presents little motivation to any malicious attacker to invest the resources in extracting it.  | **CD/RG** |
| Malicious software on one user's phone attacking another phone connected to the server | **Mild (2)** | **Mild (2)** | **Mild (2)** |  All phone traffic to and from the server is encrypted and server inputs are bounds checked. Phones cannot see each other’s traffic, and the chances of a successful malicious attack that grnts root access t the server is considered low due to the simplicity of the installation and the use of standard security precautions. Server instances are Ubuntu LTS releases with current patches. | Malicious software on one user's phone attacking another phone connected to the server |

Appendix A: Security measures for the App “Super Swipe” and its control system

1. All communications controlled or specified by TBDS are encrypted. This includes game-server, control\_app-server, and game-mapping\_service. All communication between app, server and control app uses HTTPS/TLS encryption.
2. All communications protocols between game-server and control\_app-server are explicitly documented by Two Tails and disclosed to TBDS. These include which APIs are being used, data structures, message sizes, timing, and transaction descriptions. APIs implement suitable security practices to ensure user access is restricted where necessary, for example, e.g. by using API keys.
3. Write endpoints (POST/PUT/PATCH) used by the control app require a valid API key to access. Requests with an invalid API key are rejected. API keys are managed by a server admin. API keys are only valid for a certain period of time before automatically expiring. Expired API keys cannot be reused.
4. Geo-mapping service API and throughput requirements.

Mapbox free tier is available for up to 25,000 monthly active users, which gives plenty of room for TBDS. See <https://www.mapbox.com/pricing/>

1. Overall network traffic per user, average and peak, and an indication of server CPU/memory requirements for 1.5K users.

Outside the show, the app makes no requests to the server. During a show, on average the app makes a request to the server every 5 seconds. This request is typically around 800 bytes including request and response.

Therefore, over a 10-minute period, one user would typically use around 93.75KB of bandwidth (includes in and out). 1.5k users would use 137.32MB.

The server software itself is very lightweight, making use of a single PHP instance and a MySQL database. A small virtual hosted server with 2G of RAM and a standard CPU configuration will be more than capable of handling the requirements for the show.

1. All data structure handling is bounds-checked and explicit handling provided for malformed or out-of-bounds messages. Any malformed requests to the server are rejected.
2. Control app, game authentication, DOS attack mitigation

DOS (Denial of Service) protection is difficult as there is no single easy solution and a proper solution requires an experienced server administrator to put in place and monitor.

1. All code for the app, control app and server is hosted securely on Bitbucket and access is granted to members of Civic Digits.

Appendix B. The Big Data Show Website Security and Privacy

**17th Jul 2020**

* The server is a virtual machine provided by Linode running Arch Linux. Only the directors of Foxdog have access to the Linode account and their passwords are stored by Lastpass.
* The virtual machine is dedicated to the site and is not used for any other purpose.
* Access to the server is restricted to SSH with key-based authentication. Password authentication and root login are disabled.
* Administrative access on the server is protected by sudo with password authentications.
* The server is updated weekly to apply security and stability fixes.
* All communication with the server is protected by TLS. Our certificate is provided by Let’s Encrypt and the web server’s TLS layer is configured by certbot.
* Streaming to the server is protected by a password, which is only shared with the operator.
* The remote control interface is protected by a password, which is only shared with the operator.
* The website’s admin panel is password protected by Meteor’s authentication system. The password is only shared with operators of the site.
* An audit of client-server communication points was conducted to ensure each has the appropriate authentication checks. No modifying (e.g., changing show name) operation can be performed without authentication.
* For non-administrative use of the site, no identifying information is stored on the client or server.
* All information entered by administrators (show names, portraits, etc.) should be considered public and contain no identifying information.
* Removing a show also removes all uploaded images associated with that show.
1. https://developer.android.com/guide/topics/permissions/overview [↑](#footnote-ref-1)