Information Sheet

Participant Information Sheet

Project Title: **Usability of data disclosures**

What is the study about?

We invite you to participate in a remote user study about 'data disclosures' organised by academic researchers at the University of Cambridge.

Some regulations give people the right to access information that organisations hold on them. However, organisations differ in the way they return and present this information to individuals. This research compares two ways in which this information (termed a **disclosure**) is presented.

Do I have to take Part?

Participation is entirely voluntary. This information sheet has been written to help you decide if you would like to take part. If you do decide to take part, you will be free to withdraw at any time without prejudice and without providing a reason, now or in the future. All questions are optional. After completing the study, you can contact either of the researchers at the email addresses listed on the consent form (next page) if you would like your data to be deleted, though this can only be done up to the point where the study is published and the anonymised dataset is made available online (so that other researchers can scrutinise and verify our findings).

What would I be required to do?

You will be asked to look at an interface presenting a user's data from a hypothetical company and attempt a selection of tasks (note that this data has been created by the researchers and is not a real disclosure containing your data). After doing so, you will be asked to complete a questionnaire about what you think of that interface (the way in which the data was presented to you). This process will then be repeated with a second interface, and you will be asked to complete the same scenarios and questionnaire but for the second interface. Finally, you will be asked a few questions about your experience with these interfaces.

It is important to note that you are not being evaluated in this study; rather, we are evaluating the two interfaces. It may not be obvious how certain tasks can be completed. The study should take around 15 minutes, and you will be given an MTurk code on completion. Enter this code into MTurk to receive credit for taking part.

Will my participation be anonymous and confidential?

Yes. Any identifiable information will be kept confidential, secure, and will be deleted after a year. The answers collected during this study will be pseudonymised (e.g. you will be renamed to 'participant 1', etc.) and kept separately from identifiable data. If your answers to any of the questions contain information that could potentially identify you (e.g. locations, names), these will be changed to protect your identity.

Your pseudonymised answers will be collated with other participants, and this dataset will be made available online following the publication of this research. We may choose to quote some of your answers in the publication, however, these quotes will be anonymous and will not contain any identifiable information.

Are there any potential risks to taking part?

We have not identified any specific risks associated with participating in this study.

Questions

If you have any questions, please feel free to contact either researcher at the details provided on the consent form (next page). You can do so before starting, during, or after completing the study.

Consent and Approval

This research proposal has been scrutinised and been granted Ethical Approval through the university ethical approval process.

What should I do if I have concerns about this study?

A full outline of the procedures governed by the University Research Ethical Committee is available at: https://www.research-integrity.admin.cam.ac.uk/research-ethics/university-research-ethics-committee.

Consent Form

Consent Form

We ask that you to consider the following points before proceeding. Completing this form confirms that you are an adult over 18 years of age and are happy to participate in this user study. If you have any questions before, during, or after the study, please don't hesitate to email us via the contact details below.

Please check each statement concerning the collection and use of the research data:

- I have read and understood the participant information sheet.

- I have been given the opportunity to ask questions and have had them answered to my satisfaction.

- I understand that my participation is voluntary and that I am free to withdraw at any time without giving a reason.

- I understand that all questions are optional.

- I understand that my data will be pseudonymised (including the removal of any names or specific locations mentioned).

- I understand that answers I give may be included in a research publication in a way which is non-identifiable.

- I understand that non-identifiable data from this user study will be collated with the answers of other participants and the dataset will made available online for research transparency purposes.

☐ I have read and agree with each of the above statements, and I agree to take part in the study.

Participation in this research is completely voluntary and your consent is required before you can participate. If you decide at a later date that you would like your data to be destroyed, we will delete all of your data. However, this will only be possible up to the point when the dataset is made available online.

Researcher Contact Information

Chris Norval: chris.norval@cst.cam.ac.uk

Jat Singh: jatinder.singh@cst.cam.ac.uk

Introduction

In this study we ask you to evaluate two "**disclosure interfaces**". A disclosure interface represents the way in which data is sent back to you and presented when you request it from an organisation.

For each disclosure interface, you will be asked to use the interface to uncover particular information. The purpose of these activities is to familiarise yourself with the interface. Do not worry if you are unsure how to uncover the information and answer the question.

The organisation in question is a hypothetical music and fitness app that runs on your phone. You can use this app to record your exercise sessions and statistics, monitor health information, and have the app stream recommended music to you while you exercise. The app presents users with various advertisements.

Importantly, we are not asking you to evaluate the application itself, but rather, only the way in which the data is presented (the disclosure interface) through screenshots. In particular, we are interested in your views on navigating that which is presented.

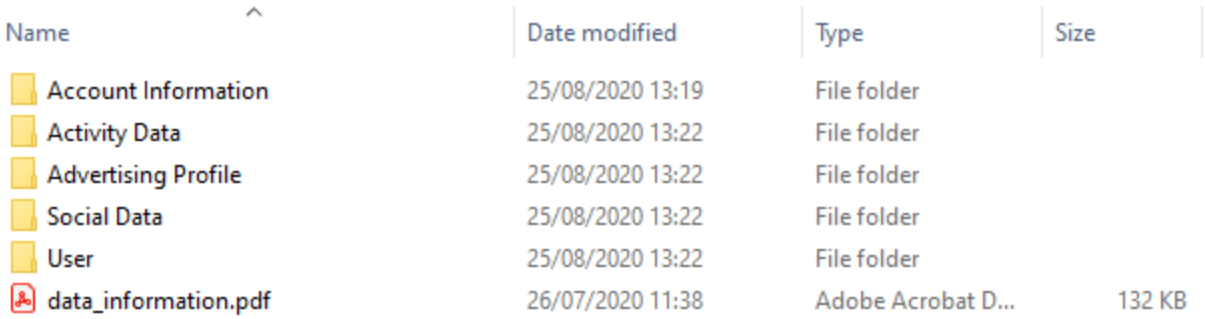
Control

We will now present you with the File-Based Disclosure Interface.

File-Based Disclosure Interface

You approach the organisation responsible for this app and request a copy of the data they have about you.

They respond, sending a copy of your data back to you in the following way:



Name	Date modified	Type	Size
Account Information	25/08/2020 13:19	File folder	
Activity Data	25/08/2020 13:22	File folder	
Advertising Profile	25/08/2020 13:22	File folder	
Social Data	25/08/2020 13:22	File folder	
User	25/08/2020 13:22	File folder	
data_information.pdf	26/07/2020 11:38	Adobe Acrobat D...	132 KB

A description of the data is provided in data_information.pdf:

Understanding my data

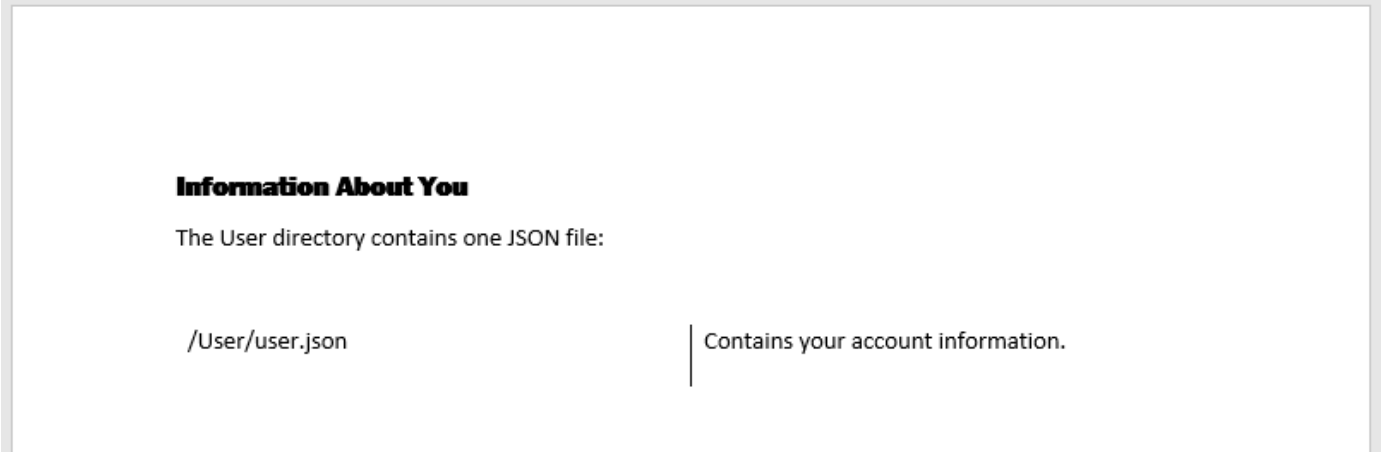
When you request your data, you will receive several files in JSON format. JSON stands for JavaScript Object Notation, and is a structured, commonly used format that is capable of being understood by both computers and human beings. Each directory contains a different type of personal data as described below:

Information about you	Information that could personally identify you. Includes: account name, name, email address, date of birth, location, height, weight
Account information	Information about your account settings. Includes: Privacy settings, email settings
Activity data	Records of your exercise sessions and music streaming activity. Includes: Exercise session, GPS data, distance, pace, duration, heart rate, songs streamed
Social data	Records of your social interactions with other users. Includes: Friend requests, comments, clubs, events
Advertising profile	Information collected and derived for advertising purposes. Includes: Advertising profile, inferences, adverts shown, adverts clicked

We will now ask you to attempt a few tasks in relation to the disclosure. Do not worry if you do not know how to complete these tasks.

Scenario 1: User Information

A description of the files is shown as:



Information About You

The User directory contains one JSON file:

/User/user.json	Contains your account information.

You open the file associated with your account information.

user.json:


```
{
  "id": 123456789,
  "first_name": "David",
  "last_name": "Smith",
  "created_at": 1527183643,
  "email": "d.smith@cam.ac.uk",
  "birthday": "1988-05-10",
  "gender": "M",
  "language": "en_GB",
  "last_sign_in_at": 1597519305,
  "last_sign_in_ip": "92.240.34.127",
  "locale": "Cambridge, UK",
  "login": "david-smith",
  "tos_accepted_at": 1583433560,
  "tos_accepted_version": 2,
  "updated_at": 1588783374,
  "height": 1.85,
  "weight": 70
}
```


On what date did you accept the Terms of Service (ToS)? (leave blank if unable to answer)

How certain are you of your answer?

Very certain	Uncertain	Couldn't answer
○	○	○

Scenario 2: Activity Data

A description of the files is shown as:

Activity Data

The Activity Data directory contains two JSON files, and one GPX file for each route associated with an exercise:

/Activity Data/streaming_history.json	Contains information about your music streaming activities.
/Activity Data/exercise_history.json	Contains information about your exercise activities.
/Activity Data/GPS/{exercise id).gpx	Contains GPS information about this particular exercise.

You open a selection of files relating to your usage of the app.

streaming_history.json:

```
[
  {
    "endTime": 1595780985,
    "artistName": "Bob Dylan",
    "trackName": "Like a rolling stone",
    "msPlayed": 23455
  },
  {
    "endTime": 1595781231,
    "artistName": "Audioslave",
    "trackName": "Be yourself",
    "msPlayed": 246375
  },
  {
    "endTime": 1595781483,
    "artistName": "Frightened Rabbit",
    "trackName": "Get out",
    "msPlayed": 252138
  },
  {
    "endTime": 1595781683,
    "artistName": "The Strokes",
    "trackName": "Is this it",
    "msPlayed": 19976
  },
  {
    "endTime": 1595781914,
    "artistName": "Blur",
    "trackName": "Coffee and TV",
    "msPlayed": 231476
  }
]
```


exercise_history.json:

```
[
  {
    "start_time": 1595782159,
    "created_at": 1595782169,
    "updated_at": 1595782169,
    "distance": 4790,
    "duration": 1795,
    "average_speed": 6.14,
    "longitude": -2.9782898,
    "latitude": 56.455082,
    "duration_per_km": 358814,
    "gps": "GPS/936224.gpx",
    "id": 936224,
    "sport_type": "Running"
  },

...
```


/GPS/936224.gpx:

```
<?xml version="1.0" encoding="UTF-8"?>
<gpx creator="FitMusic App" version="1.1" xmlns="http://www.topografix.com/GPX/1/1"
xmlns:gpxtpx="http://www.garmin.com/xmlschemas/TrackPointExtension/v1"
xmlns:gpxx="http://www.garmin.com/xmlschemas/GpxExtensions/v3"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.topografix.com/GPX/1/1 http://www.topografix.com/GPX/1/1/gpx.xsd
http://www.garmin.com/xmlschemas/GpxExtensions/v3 http://www.garmin.com/xmlschemas/GpxExtensionsv3.xsd
http://www.garmin.com/xmlschemas/TrackPointExtension/v1
http://www.garmin.com/xmlschemas/TrackPointExtensionv1.xsd">
<metadata>
<name>fitmusic_20160310_1339</name>
<copyright author="FitMusic">
<year>2020</year>
<license>http://www.fitmusic.com</license>
</copyright>
<link href="http://www.fitmusic.com"><text>fitmusic</text>
</link>
<time>2016-03-10T12:39:15.000Z</time>
</metadata>
<trk>
<name>fitmusic_20160310_1339</name>
<link href="http://www.fitmusic.com/sport-sessions/56e16af57fa48rf5bb9584d3"><text>Visit this link to
view this activity on fitmusic.com</text>
</link>
<trkseg>
<trkpt lat="56.455081939697266" lon="-2.978289842605591">
<ele>62.0</ele>
<time>2016-03-10T12:39:15.000Z</time>
</trkpt>
<trkpt lat="56.45508575439453" lon="-2.9783031940460205">
<ele>62.0</ele>
<time>2016-03-10T12:39:15.000Z</time>
</trkpt>
<trkpt lat="56.455081939697266" lon="-2.978330373764038">
<ele>62.0</ele>
<time>2016-03-10T12:39:17.000Z</time>
</trkpt>
<trkpt lat="56.45500183105469" lon="-2.9784152507781982">
<ele>57.0</ele>
<time>2016-03-10T12:39:19.000Z</time>
</trkpt>
<trkpt lat="56.45493698120117" lon="-2.9785897731781006">
<ele>55.0</ele>
<time>2016-03-10T12:39:22.000Z</time>

...
```


What was the distance you ran on this particular session? (leave blank if unable to answer)

How certain are you of your answer?

Very certain	Uncertain	Couldn't answer
○	○	○

Scenario 3: Advertising Information

A description of the files is shown as:

Advertising Profile

The Advertising Profile directory contains five JSON files:

/Advertising Profile/inferences.json	Lists the inferences deemed relevant for your account.
/Advertising Profile/user_ad_profile.json	Details the assignment of each inference to your account.
/Advertising Profile/advertisers.json	Shows you information about each advertiser that has advertised to you.
/Advertising Profile/adverts.json	Shows you information about each advert that has been shown to you, and why.
/Advertising Profile/advert_displays.json	Lists information about when each advert was shown to you, and whether you clicked it or not.

You open a selection of files relating to your advertising profile on the app.

inferences.json:

```
{
  "inferences": [
    {
      "id": 43756,
      "name": "Low Mood",
      "details": "May be suffering from low mood.",
      "type": "Wellbeing"
    },
    {
      "id": 38852,
      "name": "Fast Food",
      "details": "May enjoy purchasing fast food.",
      "type": "Food and Drink"
    },
    {
      "id": 92995,
      "name": "Cambridge",
      "details": "May be interested in events and organisations based in Cambridge, UK.",
      "type": "Location"
    },


...
```


user_ad_profile.json:

```
{
  "inferences": [
    {
      "inference_id": 43756,
      "date": 1587081600,
      "criteria": [
        {
          "exercise_frequency": "running",
          "trend": "decrease",
          "date_start": 1586390400,
          "date": 1460764800
        },
        {
          "streaming_frequency": "downbeat music",
          "trend": "increase",
          "date_start": 1586390400,
          "date": 1460764800
        }
      ]
    },
    {
      "inference_id": 38852,
      "date": 1582502400,
      "criteria": [
        {
          "proximity_fast_food_venue": "true",
          "measurement": "heart_rate",
          "trend": "increase",
          "date": 1582416000
        }
      ]
    },
    {
      "inference_id": 92995,
      "date": 1580688000,
      "criteria": [
        {
          "location": "Cambridge, UK"
        }
      ]
    },

...
```

advertisers.json:


```
{
  "advertisers": [
    {
      "id": 245132,
      "name": "Serenity Yoga Centre",
      "text": "Britain's premium yoga centre for inner wellness",
      "logo": "/advertisers/245132.jpg",
      "url": "www.serenityyoga.com"
    },

...
```


adverts.json:


```
{
  "adverts": [
    {
      "id": 734109,
      "advertiser_id": 245132,
      "audio_advert": "/audio_adverts/734109.wav",
      "image_advert": "/audio_adverts/734109.png",
      "advert_url": "www.serenityyoga.com/signup"
      "inference_id": [
        43756,
        92995
      ]
    },

...
```


advert_displays.json:

```
{
  "advert_displays": [
    {
      "advert_id": 734109,
      "shown": [
        {
          "time": 1587048920,
          "clicked": "no"
        },
        {
          "time": 1587050089,
          "clicked": "no"
        }
      ]
    },

...
```


Why was an advert for Serenity Yoga Centre shown to you? (leave blank if unable to answer)

How certain are you of your answer?

Very certain	Uncertain	Couldn't answer
○	○	○

File-Based Disclosure Interface

We will now ask you to evaluate the File-Based Disclosure Interface based on your experiences.

To what extent do you agree with the following?

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I would like this style of disclosure interface to become more widely used.	○	○	○	○	○
I found the disclosure interface unnecessarily complex.	○	○	○	○	○
I thought the disclosure interface was easy to use.	○	○	○	○	○
I think that I would need the support of a technical person to be able to use this disclosure interface.	○	○	○	○	○
I found the various functions in this disclosure interface were well integrated.	○	○	○	○	○
I thought there was too much inconsistency in this disclosure interface.	○	○	○	○	○
I would imagine that most people would learn to use this disclosure interface very quickly.	○	○	○	○	○
I found the disclosure interface very cumbersome to use.	○	○	○	○	○

to use.

I felt very confident using the disclosure interface.	○	○	○	○	○
I feel that I would need to learn a lot of things before I could get going with this disclosure interface.	○	○	○	○	○

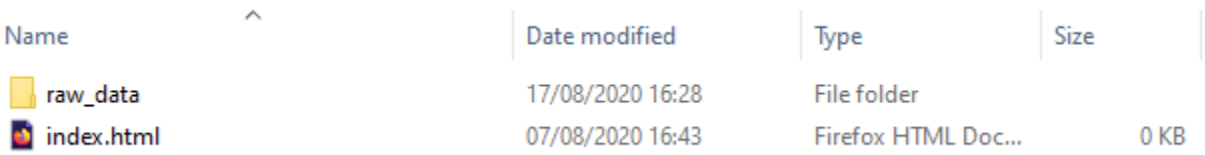
Modified

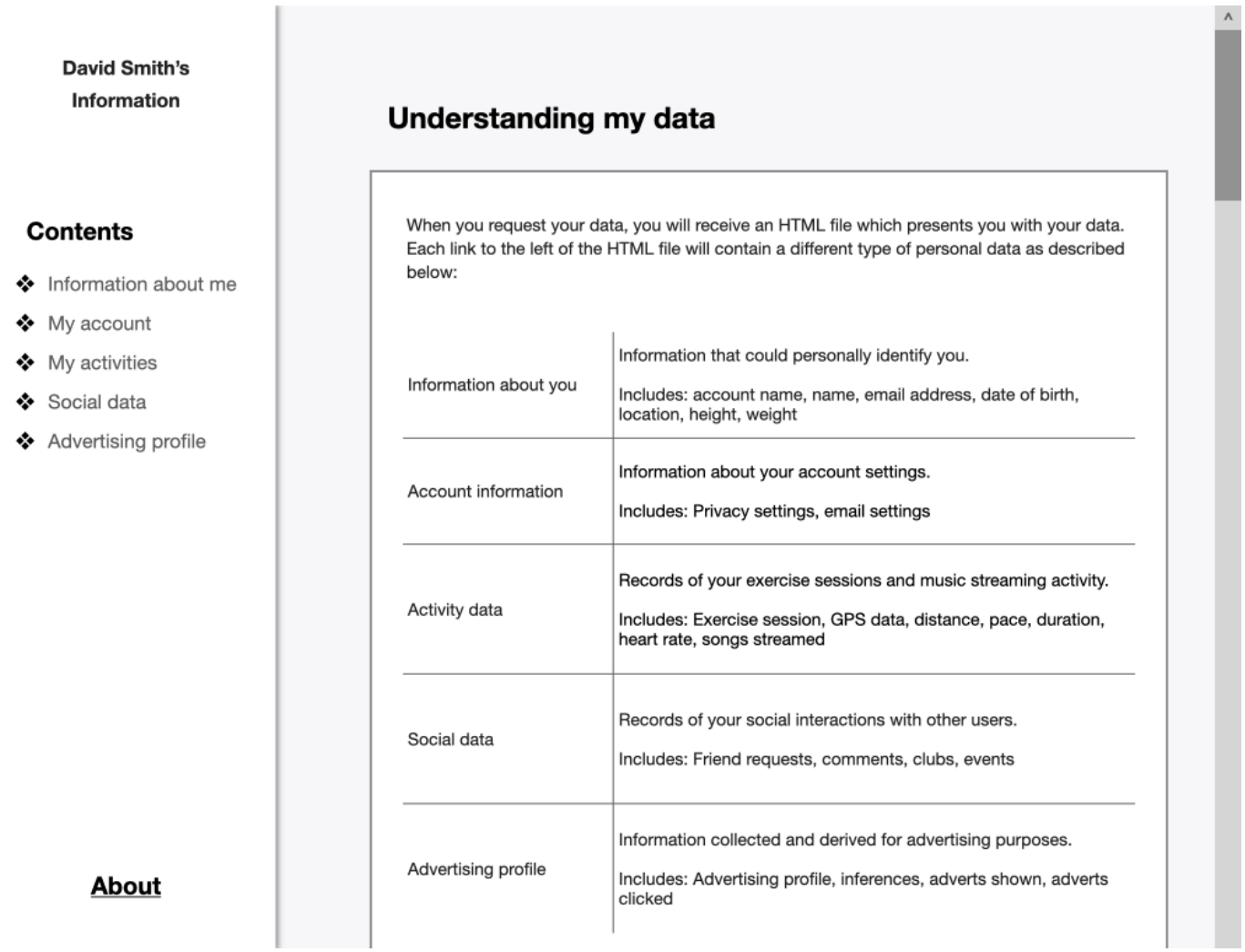
We will now present you with the Browser-Based Disclosure Interface.

Browser-Based Disclosure Interface

You approach the organisation responsible for this app and request a copy of the data they have about you.

They respond, sending a copy of your data back to you in the following way:

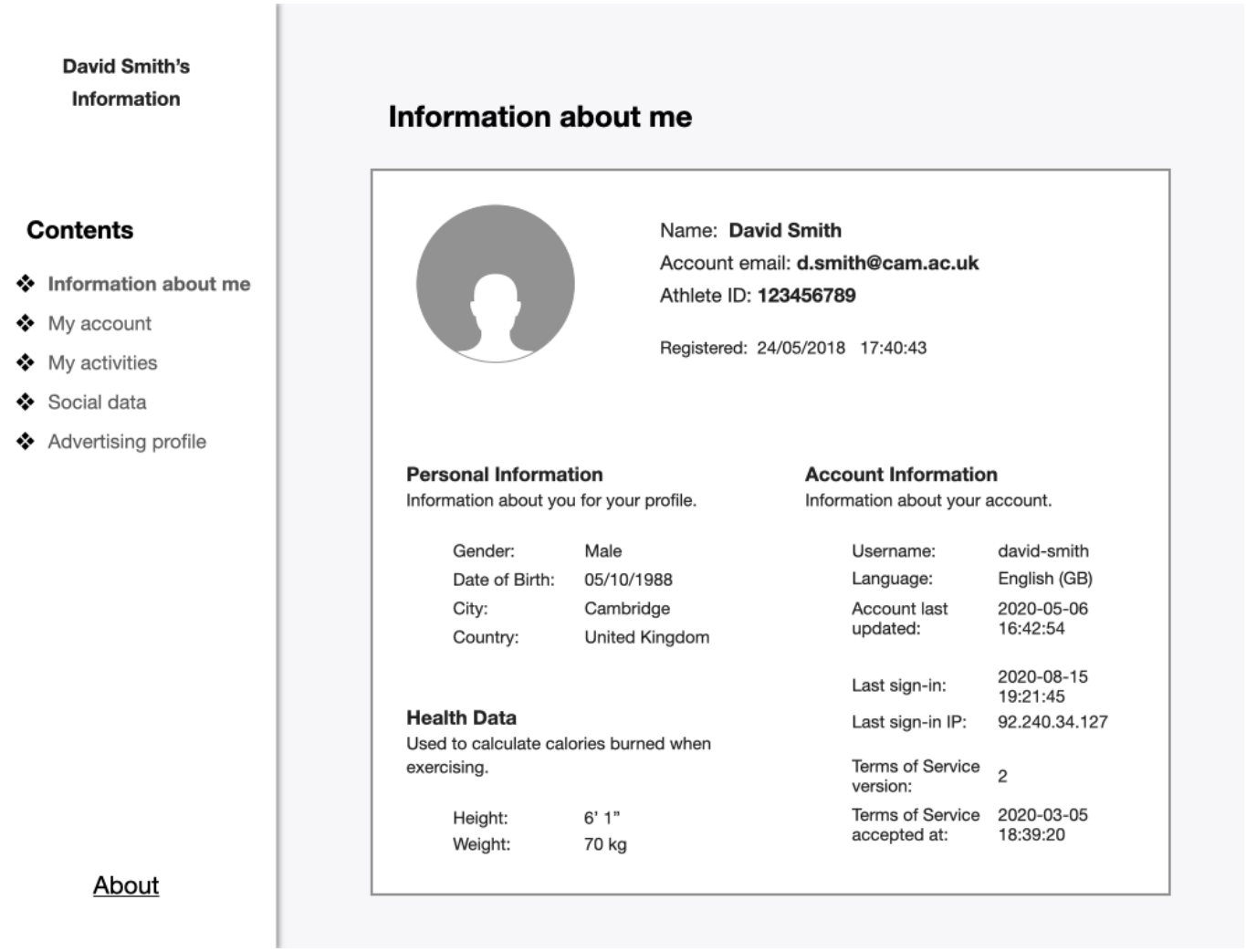




We will now ask you to attempt a few tasks in relation to the disclosure. Do not worry if you do not know how to complete these tasks.

Scenario 1: User Information

You click on 'Information about me'.



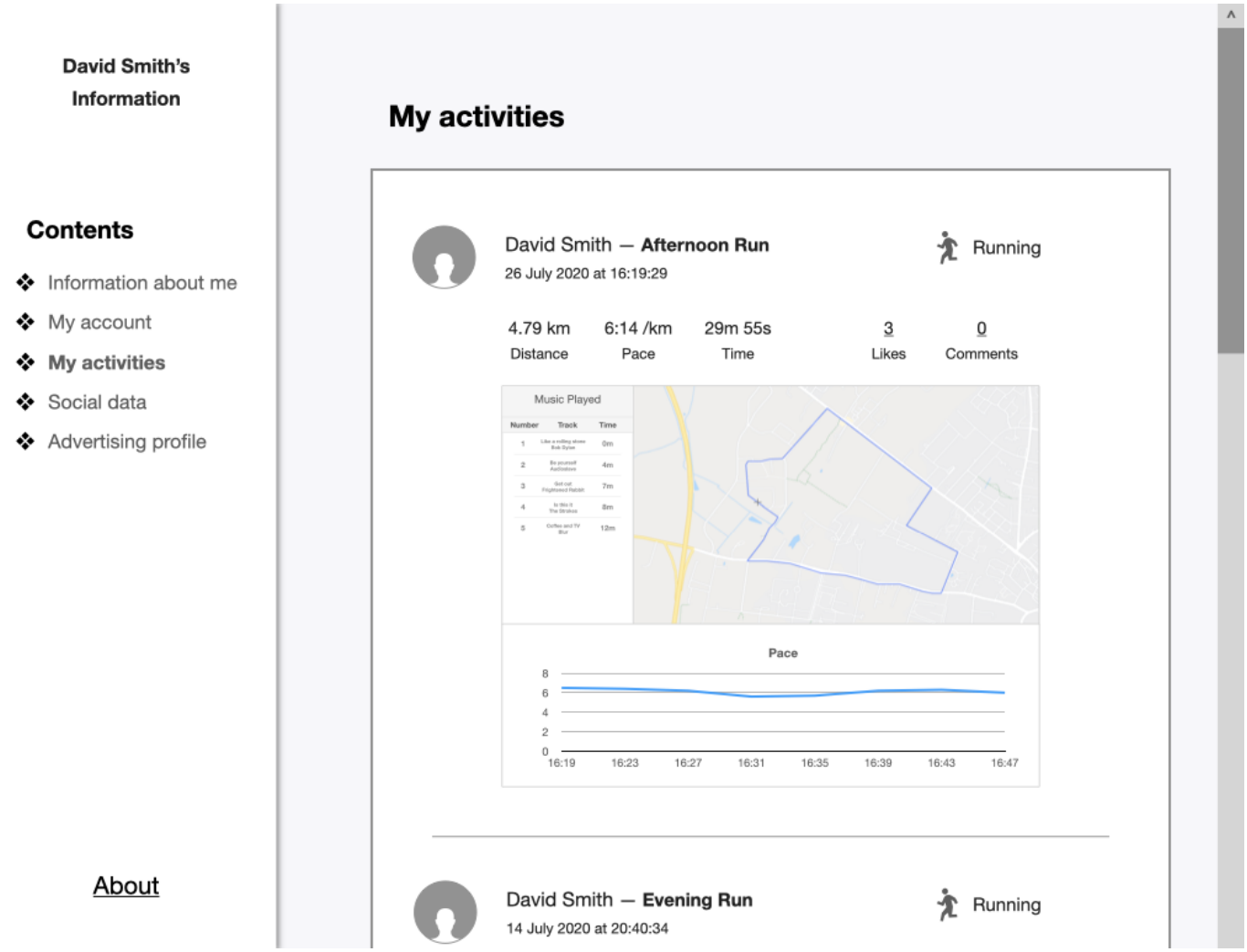
On what date did you accept the Terms of Service (ToS)? (leave blank if you're unable to answer)

How certain are you of your answer?

Very certain	Uncertain	Couldn't answer
○	○	○

Scenario 2: Activity Data

You click on 'My activities'.



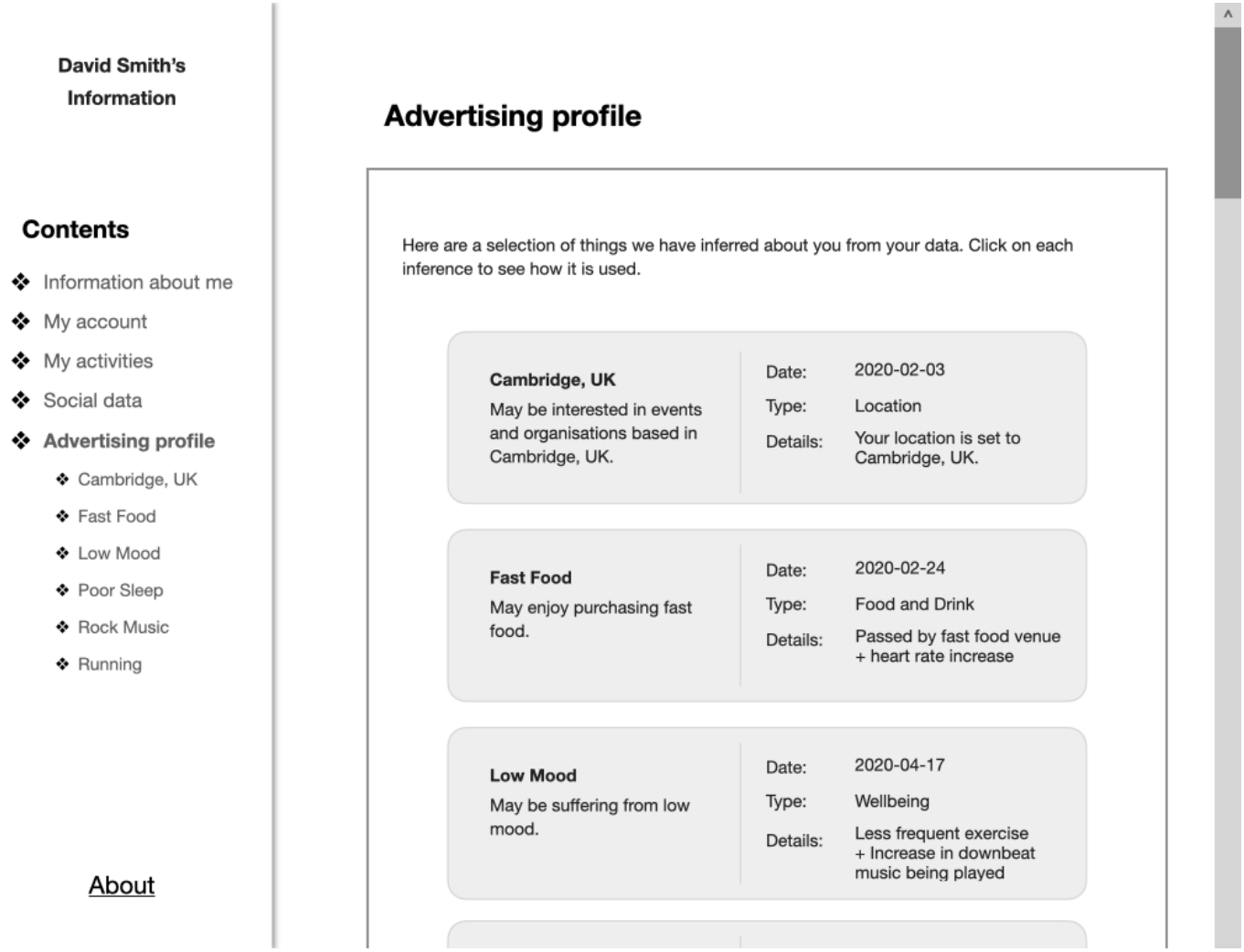
What was the distance you ran on this particular session? (leave blank if unable to answer)

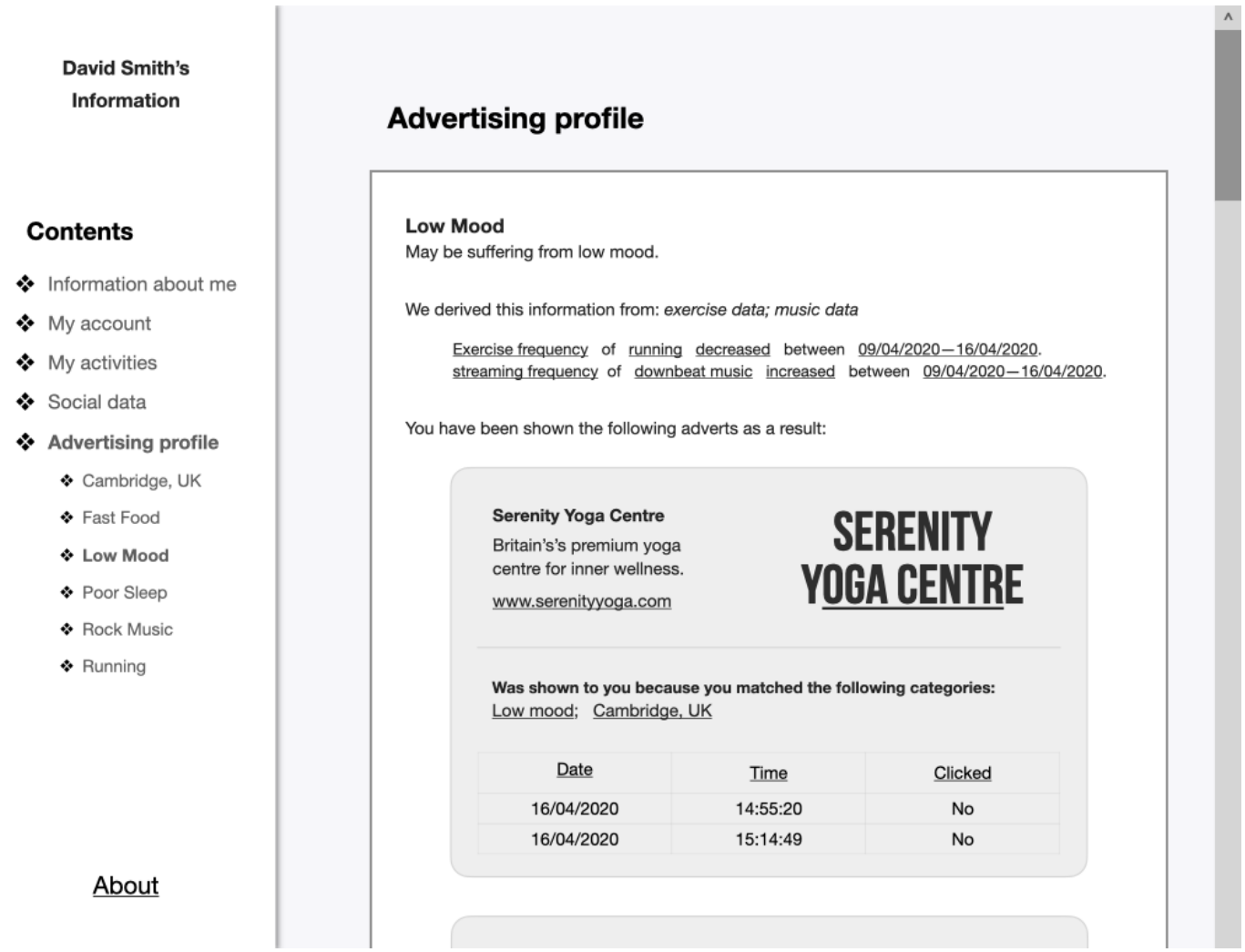
How certain are you of your answer?

Very certain	Uncertain	Couldn't answer
○	○	○

Scenario 3: Advertising Information

You explore the 'Advertising profile'.





Why was an advert for Serenity Yoga Centre shown to you? (leave blank if unable to answer)

How certain are you of your answer?

Very certain	Uncertain	Couldn't answer
○	○	○

Browser-Based Disclosure Interface

We will now ask you to evaluate the Browser-Based Disclosure Interface based on your experiences.

To what extent do you agree with the following?

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I would like this style of disclosure interface to become more widely used.	○	○	○	○	○
I found the disclosure interface unnecessarily complex.	○	○	○	○	○
I thought the disclosure interface was easy to use.	○	○	○	○	○
I think that I would need the support of a technical person to be able to use this disclosure interface.	○	○	○	○	○
I found the various functions in this disclosure interface were well integrated.	○	○	○	○	○
I thought there was too much inconsistency in this disclosure interface.	○	○	○	○	○
I would imagine that most people would learn to use this disclosure interface very quickly.	○	○	○	○	○
I found the disclosure interface very cumbersome to use.	○	○	○	○	○

to use.

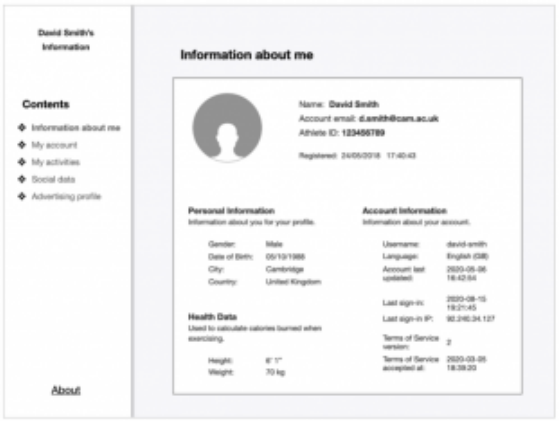
I felt very confident using the disclosure interface.	○	○	○	○	○
I feel that I would need to learn a lot of things before I could get going with this disclosure interface.	○	○	○	○	○

Ending Questions

Disclosure Preferences

We will now ask you some questions about your experiences with the two disclosure interfaces, and disclosures more generally.

File-based Disclosure Interface



Browser-based Disclosure Interface

Which disclosure interface did you prefer?

- ○ Browser-based Disclosure Interface
- ○ File-based Disclosure Interface

When using the disclosure interfaces, what did you like or find helpful?

When using the disclosure interfaces, what did you dislike or find unhelpful?

What advice would you give a designer of a disclosure interface in order to make it more user-friendly?

Reflecting back on the tasks, do you have any further thoughts you'd like to

share?

Please order the reasons for which you would be interested in obtaining your data from an organisation (where 1 is most interested):

General interest

Find a particular bit of information

Concern or distrust of the organisation

Legal reasons

Archival purposes

To transfer that data to another service

Verify correct information

Ensure that no unwanted data is being stored

Other

If you have an 'Other' in mind for the previous question, please elaborate

Have you ever requested your data from an organisation?

- ○ Yes
- ○ No

Please list the key reason(s) motivating you to request your information

Please comment on usefulness, and what you liked about the presentation of the data provided by the organisation(s)

Please commented on what you disliked about the presentation of the data provided by the organisation(s)

Demographic

We will now ask you some demographic questions to help us to better understand our sample of participants. Please keep in mind that all questions are optional.

How would you evaluate your knowledge of the following?

	No knowledge	Some knowledge	Good level of knowledge	Advanced knowledge	Expert knowledge
Technical expertise	○	○	○	○	○
Knowledge of the GDPR	○	○	○	○	○

Gender:

Age:

- ○ 18 - 29
- ○ 30 - 39
- ○ 40 - 49
- ○ 50 - 59
- ○ 60+
- ○ Prefer not to say

Country

- ○ UK
- ○ Other

Which country do you live in?

MTurk Code

Thank you for taking part. **Here is your MTurk code: ${e://Field/Random%20ID}**

Copy this value to paste into MTurk.

When you have copied this value, please click the next button to complete the study.