
Assessment Template	TfL Data Protection Impact Assessment 3.0
Template Version	1
Status	Approved
Assessment Workflow	TrustArc Default Workflow
Assessment Description	This relates to a previous assessment (PID00100004). Following on from our successful Wi-Fi data insights pilot, we are now developing this processing into a 'business as usual' approach.
TfL business areas	London Underground Technology and Data
TfL subsidiaries	London Underground Ltd Transport for London
Assessment Label	Journey data Customer data Wi-Fi
Created By	Rebecca Florence
Assessment Owner	Lee McGirr (Data Governance Manager)
Approvers	Richard Bevins (Head of Information Governance and Data Protection Officer) Simon Guild (Head of Privacy and Data Protection) Rebecca Florence (Privacy Adviser)
Respondents	Lee McGirr (Data Governance Manager)

Questions and Answers: *Will TfL be processing personal data?*

1 *Below are some categories of personal data routinely processed by TfL. Please select all of those that will be processed in connection with this project or initiative. If it involves the processing of a category of personal data not included as an option, please select "Other" and provide further details.*

Status No Issue

Response Media Access Control ("MAC") address

Risk -

Comments & Attachments

📎 When a device such as a smartphone or tablet has Wi-Fi enabled, the device will continually search for a Wi-Fi network to connect to. When searching for a Wi-Fi network, the device sends out a probing request which contains an identifying number specific to that device known as a Media Access Control (MAC) address.

If the device finds a Wi-Fi network that is known to the device, it will automatically connect to that network. If the device finds unknown networks, it will list these in the user's device settings allowing them to decide whether to connect to one of them.

When a device is near one of our station Wi-Fi access points and Wi-Fi is enabled, the device will send a probing request to connect. This will be received by TfL's Wi-Fi network, even if the device does not subsequently connect.

If a device has not signed up to use the free Wi-Fi provided on the London Underground network, it's an un-authenticated device. Most modern devices send out a randomly generated MAC address to prevent unknown routers identifying the device which is known as spoofing.

We will not process un-authenticated devices for the purposes described within this DPIA. We will remove un-authenticated devices from the data we will be analysing as soon as possible after receipt.

If the device has been signed up for free Wi-Fi on the London Underground network, the device will disclose its genuine MAC address. This is known as an authenticated device. We process authenticated device MAC address connections (along with the date and time the device authenticated with the Wi-Fi network and the location of each router the device connected to). This helps us to better understand how customers move through and between stations - we look at how long it took for a device to travel between stations, the routes the device took and waiting times at busy periods.

We do not collect any other data generated by a device. This includes web browsing data and data from website cookies.

Individuals can choose to de-register from using the Wi-Fi on our network. Alternatively, individuals can turn their Wi-Fi off, turn their device off or put their device on airplane mode. All individuals will be informed via posters and the TfL website how they can opt-out. All MAC addresses will be pseudonymised using an irreversible process. This prevents TfL from being able to identify an individuals MAC address and we are unable to single out a specific device.

Questions and Answers: *Will TfL be processing personal data?*

2 *Will information which falls within any of the following special categories of sensitive personal data be processed as a result of this project or initiative?*

Status No Issue

Response No

Risk -

Questions and Answers: *Personal Information Custodian approval*

3 *Do you have approval from the relevant Personal Information Custodian(s) to proceed with this project or initiative?*

Status No Issue

Response Yes

Risk -

3.1 *Please confirm the name(s) and job title(s) of the Personal Information Custodian(s) who approved this project or initiative? Please attach appropriate evidence, for example an email approval or signed Statement of Work (SOW)/Project Initiation Document (PID).*

Status No Issue

Response Yes, Lauren Sager Weinstein as TfL's Chief Data Officer. This project was been approved at Customer Experience Gate 0 by the Technology and Data (T&D) Senior Management Team (SMT) in July 2016. Since the pilot, the project has been through the T&D Gate 2 receiving additional scrutiny and approval.

Risk -

Questions and Answers: *Overview and business case*

4 *Please provide a brief description of this project or initiative and attach a simple data flow diagram below. You can also attach a copy of your business case, statement of works (SOW) or project initiation document (PID).*

Status No Issue

Response In 2016, TfL carried out a pilot collection of depersonalised Wi-Fi connection data from devices using the station Wi-Fi at 54 Tube stations in central London. We found that:

- Wi-Fi data can help to understand the paths people take in stations, the platforms and lines they use, the routes they take when they have many options and the interchanges they make
- Aggregated data can show which sections of the network are crowded, at what times and how this changes in response to events and network alterations
- Data can be used with analytical tools and services to improve the way we run and plan our network, and provide customers with more detailed information

We want to use technology to provide better information to our customers. Wi-Fi connection data can give us a better understanding of crowding and collective travel patterns so we can improve services and information for customers. We expect the benefits to be as follows.

- We will be able to give customers better information to help them plan their journeys and avoid congestion. For example, we plan to use aggregated Wi-Fi connection data to show the relative 'busy-ness' of London Underground stations, in near real time. We will make this available on our website
- We will be able to manage disruptions and events more effectively, deploy staff to best meet customer needs and ensure a safe environment
- We will be able to make better transport planning decisions - for example about timetables, station designs and major station upgrades
- By understanding how many customers we have and how they move around stations we will be able to maximise revenue from the companies which advertise on our poster sites and those who rent retail units on our property. This money can be invested in improving our services.

Oyster and Contactless Payment Card ticketing data helps us understand where customers enter and exit the London Underground network (as well as any intermediate validations). But it does not tell us the platforms and lines they are using, the stations they interchange at or how they navigate around our stations. The nature of the Tube network means that people can take many different options for their journeys.

TfL has previously used paper surveys - but these are expensive, only provide a snapshot of travel patterns on the day of survey and are unable to provide a continuous flow of information. Depersonalised Wi-Fi connection data provides more accurate real time information for improving our services.

5

Please provide an overview of the benefits to TfL and its stakeholders (including customers and/or employees) and explain how those benefits outweigh any potential impact on the privacy of the individuals whose data will be processed. Include any regulatory, operational or commercial benefits.

Status No Issue

Response

From our data analysis during the pilot, we have been able to conclude that;

- Wi-Fi data can help us understand the paths customers take in stations, the platforms and lines they use, which route they take when many options exist and where they interchange
- The aggregated data can show which sections of our network are crowded, at what times, and how this changes in response to events and network alterations
- We can use this data to power analytical tools and services that can improve the way we run and plan our network, and can provide our customers with much more detailed information

In view of the clear benefits to us and our customers, we are now planning to formally roll out network-wide Wi-Fi data collection, so we can better understand travel patterns, provide enhanced information for our customers, and improve our planning and operations.

Risk

-

6

Please provide details of any previous DPIA (or Privacy Impact Assessment) carried out in relation to any elements of this project or initiative. You can also attach a copy below.

Status No Issue

Response

Data Protection Impact Assessment completed for the pilot of this project and accompanies this DPIA on our website.

Risk

-

Questions and Answers: *Stakeholders*

7 *Has everyone directly involved in the design and delivery of this project or initiative completed TfL's ["My role in privacy and data protection"](#) eLearning course?*

Status No Issue

Response Yes

Risk -

8 *From the list provided, please identify all of the relevant internal stakeholders for this project or initiative.*

Status No Issue

Response Information Governance | Cyber Security and Incident Response Team (CSIRT) | Legal | Press Office | Tech & Data Digital | Commercial Development | Customer Communications & Technology

Risk -

9 *From the list provided, please identify all of the relevant external stakeholders for this project or initiative.*

Status No Issue

Response Trade unions | Customers and/or members of the public | Consumer groups (including London TravelWatch) | Civil liberties groups and privacy campaigners | Greater London Authority (GLA) | Information Commissioner's Office

Risk -

9.1 *Please select the type of consultation exercise conducted with customers and/or members of the public. Please also attach any supporting evidence/outcomes (research reports etc).*

Status No Issue

Response Focus groups

Risk -

Comments & Attachments

📎 For the pilot, TfL conducted a number of focus groups to help understand how our customers felt about this type of processing. TfL formally wrote to 6 external and high profile privacy campaign groups asking them if they would like to meet with TfL to discuss our project. Only 1 responded where a meeting was set up to discuss the pilot and how TfL may implement the capture of this data going forward. TfL provided the group with additional information after the meeting, but there was no follow up from the group.

TfL has fully engaged with the Information Commissioners Office during the pilot and during the project for full rollout.

10 *Have you already notified all of the internal and external stakeholders identified in your responses to questions eight and nine, about this project or initiative?*

Status No Issue

Response Yes

Risk -

11 *Have any of the internal or external stakeholders for this project or initiative expressed any concerns or reservations about the way in which TfL is intending to process personal data?*

Status No Issue

Response Yes – During the follow up discussions after the pilot, Privacy groups and Assembly members recommended that for future rollout that the posters explicitly state how to opt-out. We have actioned this recommendation. Owing to design constraints (eg size considerations) of the permanent signs that will be displayed at stations, full details of each of the opt-out options cannot be included on those – but they do explain that you can opt-out by turning off a device’s Wi-Fi. A link to the dedicated webpage for individuals to find out more information is included on permanent signs. This webpage informs of the three options as to how individuals can opt-out.

Risk -

Questions and Answers: *Suppliers and third parties*

12 *As part of this project or initiative, will any external service provider(s) be involved in processing personal data on behalf of TfL?*

Status No Issue

Response No

Risk -

13 *Will personal data be disclosed to any third party organisation(s) on a routine basis as a result of this project or initiative? Please select all that apply from the list below.*

Status No Issue

Response None

Risk -

Comments &**Attachments**

 TfL will not share any personal data or individual pseudonymised device Wi-Fi Connection data with any 3rd party.

Questions and Answers: *Benchmarking*

14 *Are you aware of any other TfL business area(s) currently undertaking similar processing of personal data?*

Status No Issue

Response No

Risk -

Issue Description -

15 *Have you identified any other organisation(s) currently undertaking similar processing of personal data?*

Status No Issue

Response Yes

Risk -

15.1 *Please specify which organisation(s) and tell us if you have already engaged with them to discuss this project or initiative.*

Status No Issue

Response We are aware that a number of retail centers and airports currently collect and process Wi-Fi connection data.

Risk -

16 *Are you aware of any external data privacy or security standards, guidelines or codes of practice which are relevant to this project or initiative?*

Status No Issue

Response Yes

Risk -

16.1 *Please provide details and include a hyperlink (if available online) or attach copies.*

Status No Issue

Response ICO's Wi-Fi location analytics guidance
<https://ico.org.uk/media/1560691/wi-fi-location-analytics-guidance.pdf>

Risk -

Questions and Answers: *Fairness and transparency*

17 *How will individuals be notified that their personal data is being processed as a result of this project or initiative? For example, using a privacy notice presented at the point of collection (on an online or paper form), signage, an audio announcement, a webpage, etc. Please attach copies of any relevant drafts/designs.*

Status No Issue

Response We are committed to being open and transparent with customers on how we use data. A range of communication messages and channels will be used to ensure customers are aware of what data we are collecting and why we are collecting the data. The following channels will be used:

- 1) audio announcements at every station participating - similar to that of CCTV announcements
- 2) signage at every station participating including entrances and interchanges which will include a link to the below mentioned webpage
- 3) dedicated webpage for individuals to find out more information which will be hosted on our privacy pages
- 4) social media announcement which directly informs millions of individuals of this processing
- 5) brief social media teams and our Contact Centre staff to enable them to respond to any questions customers may ask
- 6) an article within the Metro newspaper for launch but also consider a follow up article informing customers of the benefits of the project
- 7) station staff will be fully informed of the processing and will advise customers to seek more information using the dedicated webpage
- 8) press office will coordinate media enquiries

Risk -

18 *Which of the following processing conditions will apply to this project or initiative?*

Status No Issue

Response Article 6 (1)(e) of the GDPR - processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller;

The specified purposes that we intend to process pseudonymised data for are identified in the DPIA as: transport planning, prioritising investment, assessing effectiveness of poster sites and retail units and providing travel planning and information to the travelling public.

TfL is a statutory body created by the Greater London Authority (GLA) Act 1999. This Act gives the Mayor of London a general duty to develop and apply policies to promote and encourage safe, integrated, efficient and economic transport facilities and services to, from and within London.

The GLA Act contains specific powers to provide information to the public to help them to decide how to make use of public passenger transport services (s162) and to provide or secure the provision of public passenger transport (s173), as well as a broadly scoped power to do such things and enter into such transactions as are calculated to facilitate, or are conducive or incidental to, the discharge of any of its functions (Schedule 10 (1)).

Further miscellaneous powers are set out in Schedule 11, which include the power to “spend such sums as it considers reasonable in the exploitation of commercial opportunities arising from the activities it carries on in the discharge of its functions” (Schedule 11 (13)(c)) and also under Schedule 11 (22)(1) which states “Transport for London may do anything which appears to it to be practicable and desirable for the purpose of promoting—

(a) research into matters affecting, or arising out of, the exercise of the functions of Transport for London or any of its subsidiaries, and

(b) the exploitation of the results of any research into any such matter (whether or not promoted by Transport for London) and of anything resulting from any idea affecting, or arising out of, the exercise of any of those functions.”

Poster advertising has long been a feature on the London Underground and we generate a sizeable amount of the money required to discharge our statutory functions from advertising, property rental, and property sales. We take advantage of commercial development opportunities such as 'click and collect' at stations as well as third-party sponsorship for the Cycle Hire and the Emirates Air Line. Thus we consider that the effective use of poster sites and retail units is necessary for the funding of public transport in London.

Risk -

19 *As a result of this project or initiative, will TfL be processing personal data for the purposes of profiling individuals (eg to segment or categorise them based on predetermined criteria)?*

Status No Issue

Response No

Risk -

20 *Please provide (or attach) details of the process or mechanism which will allow individuals to access their own personal data (which will be processed as a result of this project or initiative).*

Status No Issue

Response TfL will be using an irreversible, one way pseudonymisation process which will prevent TfL from being able to identify an individuals Media Access Control (MAC) address and we will be unable to identify a specific device from the data. Therefore, because we are unable to single out a specific device, we will be unable to process requests for access to data on a specific device.

Risk -

21 *Will there be a process or mechanism which will allow individuals to transfer their personal data to another service provider (eg a train operating company or public transport authority) in a machine readable format?*

Status No Issue

Response No

Risk -

Comments & Attachments  TfL will be using an irreversible, one way pseudonymisation process which will prevent TfL from being able to identify an individuals Media Access Control (MAC) address and therefore, we will be unable to identify a specific device from the data.

22 *Please provide (or attach) details of the process or mechanism which will allow individuals to restrict or prevent the on going processing of their personal data (which will be processed as a result of this project or initiative). For example, opt-out/unsubscribe preferences or the ability to ask for their personal data to be deleted via a dedicated email address or online form.*

Status No Issue

Response Individuals can choose to de-register from using the Wi-Fi on our network. Alternatively, individuals can turn their Wi-Fi off, turn their device off or put their device on airplane mode. All individuals will be informed via posters and the TfL website how they can opt-out.

Risk -

23 *Are there any plans to combine personal data processed as part of this project or initiative with externally sourced data sets (eg bought-in marketing lists or mobile phone tracking data provided by telecoms service providers)?*

Status No Issue

Response No

Risk -

Questions and Answers: *Data minimisation and disposal*

24 *What will be done to ensure that personal data processed as a result of this project or initiative, is relevant and necessary for its stated purpose(s)? For example, cleansing the data to remove any irrelevant content not required to achieve the stated purpose(s).*

Status No Issue

Response We will only be processing data that is necessary for the project. We will be using an irreversible, one way pseudonymisation process which will prevent TfL from being able to identify an individuals Media Access Control (MAC) address and therefore, we will be unable to identify a specific device from the data.

Risk -

25 *Will any techniques be used to minimise the amount of personal data being processed as a result of this project or initiative? For example, hashing, tokenising, aggregating or pseudonymising it?*

Status No Issue

Response Yes

Risk -

25.1 *Please describe the data minimisation techniques or methodologies you will be using; and the categories of personal data to which they will be applied.*

Status No Issue

Response We will be using an irreversible, one way pseudonymisation process which will prevent TfL from being able to identify an individuals Media Access Control (MAC) address and therefore, we will be unable to identify a specific device from the data. The pseudonymisation process has been approved by our Cyber Security team.

Risk -

26 *Have you already identified the retention periods which will be applied to the personal data which will be processed as a result of this project or initiative?*

Status No Issue

Response Yes

Risk -

26.1 *Please specify the retention periods which will apply.*

Status	No Issue
Response	<p>Data is being retained for scientific, historical research and statistical purposes. Pseudonymised Wi-Fi connection data will be held for two (2) years.</p> <p>After this retention period, journeys will be aggregated to ensure individual journeys will no longer be held. The parameters of the aggregation will be confirmed at a later date.</p> <p>A local data retention schedule will be created and approved by Information & Records Management and the Privacy and Data Protection team which will be made available to customers.</p>
Risk	-

Questions and Answers: *Data accuracy and quality*

27 *What will be done to ensure that the personal data processed as part of this project or initiative is as accurate as possible (for example a validation process and/or allowing individual data subjects to update their details using an online account)?*

Status No Issue

Response As data is automatically collected, we are unable to validate or verify the data with the data subject. Given the controls we will have in place, TfL will be unable to identify an original MAC address and therefore, unable to identify a data subject from the data collected.

We will conduct an analytics task to understand station level volumes and volumes by time of day to understand how representative our data collection is in terms of the number of unique Wi-Fi connections versus the number of customers we know use our stations.

Risk -

Questions and Answers: *Data storage and security*

28 *How will personal data processed as a result of this project or initiative be stored? Please select all of the relevant locations.*

Status No Issue
Response Private cloud
Risk -

29 *Where will personal data be physically located whilst being processed (including stored)?*

Status No Issue
Response Country which the European Commission has deemed as having adequate levels of protection
Risk -

29.2 *Please specify the country or countries in which personal data will be processed (including stored).*

Status No Issue
Response The UK and the Netherlands.
Risk -

30 *Who will have access to personal data processed as a result of this project or initiative? Please provide the names of relevant roles and teams (within TfL and any external service providers or partner organisations identified in Question 12), not the names of specific individuals.*

Status No Issue
Response Access to pseudonymised data will be highly restricted using role based access and to only individuals within Data and Analytics that require access to the data. These roles within Data and Analytics may include;
 Data Science team
 Governance & Architecture team
 Development stream
 Operational support team
 Each individual would have completed TfL's Privacy and Data Protection training within the last 12 months with access constantly being monitored.

 TfL will not share any personal data or individual pseudonymised device Wi-Fi Connection data with any 3rd party. Only aggregated outputs will be available to anyone outside of Data & Analytics.
Risk -

31 *Will individuals with access to personal data as a result of this project or initiative be subject to any screening or vetting? For example, Disclosure and Barring Service ("DBS") or financial probity checks.*

Status No Issue

Response All staff would have been subject to TfL's pre-employment checks.

Risk -

32 *How will access to the personal data processed as a result of this project or initiative be restricted and controlled? Please select all that apply.*

Status No Issue

Response User activity audit trails | Read-only system access | Documented joiners, movers and leavers process | Password protected user accounts

Risk -

33 *If an individual (including the data subject themselves; a colleague from a business area or service provider that doesn't normally have access to the data; or any other third parties) requests access to personal data processed as a result of this project or initiative; will there be a procedure in place to verify their identity and ensure that they are authorised to see it?*

Status No Issue

Response No

Risk -

Comments & Attachments  TfL will not share any personal data or individual pseudonymised device Wi-Fi Connection data with any 3rd party. Only aggregated outputs will be shared.

34 *What other measures and controls will be in place to protect personal data processed as a result of this project or initiative? Please summarise the relevant safeguards under each of the following three headings: **Physical** - access (eg secure office space and storage cabinets); clear desk policy; confidential waste disposal arrangements; etc **Technological** - encryption; anti-virus; firewalls; intrusion detection; Data Loss Prevention ("DLP"); etc **Organisational** - documented policies and procedures; information asset registers; retention and disposal schedules; training; etc*

Status No Issue

Response Physical - controlled access to TfL buildings including access controls to each floor.
Technological - all data will be encrypted end to end and at rest.
Organisational - full engagement with Information Governance, Cyber Security and the Information Commissioners Office. Documented retention and disposal schedule will be created.

Risk -

35 *Have you discussed your project or initiative with the TfL Cyber Security and Incident Response Team (CSIRT)?*

Status No Issue

Response Yes

Risk -

35.1

Please provide (or attach) a summary of any feedback and/or recommendations provided by CSIRT. Please also attach any cyber security appraisal or gap analysis documentation issued by CSIRT in the context of this project or initiative.

Status No Issue

Response The Cyber Security team have been fully involved and approved our pseudonymisation / tokenisation process and will continue to be involved with the project once all designs have been formalised.

Risk -

Questions and Answers: *Assurance and complaints*

36 *Is there a documented process for addressing, escalating and resolving privacy and data protection related complaints from individuals whose data will be processed as a result of this project or initiative?*

Status No Issue
Response Yes
Risk -

36.1 *Please describe (or attach a copy of) the relevant process for dealing with privacy and data protection related complaints.*

Status No Issue
Response Individuals will be informed they can make a complaint to TfL's Data Protection Officer who will then follow TfL's Privacy and Data Protection complaints handling procedure. TfL's Data Protection Officer's contact details are available on the TfL website here: <https://tfl.gov.uk/corporate/privacy-and-cookies/your-information-rights> and on the Wi-Fi connection's dedicated webpage as part of the privacy notice.
Risk -

37 *Is there a documented process for monitoring on-going compliance with privacy and data protection requirements as part of this project or initiative?*

Status No Issue
Response Yes
Risk -

37.1 *Please describe (or attach a copy of) the process which will be used to monitor compliance with privacy and data protection requirements.*

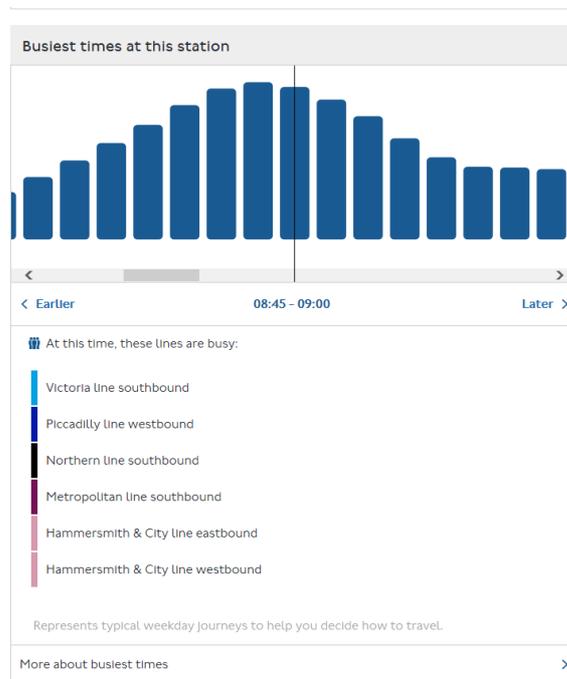
Status No Issue
Response TfL's Information Governance team will work closely alongside key stakeholders across TfL to continuously monitor compliance. Key stakeholders include Data and Analytics, the Cyber Security and Incident Response Team and Audit and Assurance. This will include, but not be limited to monitoring the draft ePrivacy Regulations and any new guidance issued by the Information Commissioners' Office or the European Data Protection Board.
Risk -

Questions and Answers: *Open data*

38 *Do you intend to make any data or analysis derived from the processing of personal data as a result of the project or initiative, available in connection with TfL's transparency and/or open data obligations?*

Status No Issue

Response Yes - Using aggregated analysed Wi-Fi connection data, we will convert this into a customer-friendly scale, showing the relative 'busy-ness' of LU stations and make this available on our website. We currently provide this for each of our LU station pages, but today we can only share historic patterns. Wi-Fi data will enhance this significantly, allowing near real time reporting. The information currently published for King's Cross St. Pancras is below.



We could also provide a version of station 'busy-ness' data via our free open-data API, which may allow app developers, academics and businesses to utilise the 'busy-ness' data for new products and services.

All data used for these purposes will be derived aggregated data and at no point could an individual be identified from the data.

Risk -