F7526 A3 Data Protection Impact Assessment (DPIA) Checklist

Any initiative, project or proposal to change processes that involves the processing of personal information (or the use of privacy intrusive technologies) is likely to give rise to various privacy and data protection concerns. Undertaking a DPIA helps to ensure that data protection risks are identified as soon as possible. A DPIA should continue to be maintained and updated throughout the project lifecycle. The GDPR makes a Data Protection Impact Assessment (DPIA) mandatory for certain types of processing, or any other processing that is likely to result in a high risk to individual's interests.

This assessment tool is designed to examine a new project / initiative, or a significant change to an existing process at an early stage. It will result in an initial assessment of privacy risk and determine which level of further assessment is necessary. The Privacy and Data Protection team will assess the completed DPIA and may request further information to assist in the identification and mitigation of privacy risks.

Your details			
Name:	Lena Ciric	Date DPIA completed	04/11/2020
Job title:	Associate Professor	Proposed launch date	11/11/2020

Name and description of the project:	COVID-19: Reducing the Risk of Transmission on London's transport vehicles To understand the Covid-19 infection risk on London Underground systems. The pathogen causing COVID-19, SARS-CoV-2, can be transmitted from one person to another through the air, by direct contact or via inanimate objects. The risk of transmission on public transport vehicles is the subject of this study.

	In collaboration with Transport for London (TfL), the proposed interdisciplinary project aims to investigate a number of parameters which influence transmission by gathering evidence which will allow for the quantification of transmission risk. The work packages include • WP1: Microbiological studies will be performed in the laboratory and on site to identify contamination hotspots, • WP2: High-resolution computer simulations will be carried out predicting airflows and mucus droplet dispersion to allow predictions of how the virus may transmit through the air or how and where droplets settle, • WP3: Air quality on vehicles will be investigated to estimate of virus concentration in the air, and • WP4: in-vehicle CCTV images will be analysed to identify the surfaces people touch in vehicles (and touching frequencies), as well as the passenger positions (orientation of passengers with respect to breathing airflow and proximity to each other and parts of the vehicle). This DPIA concerns WP4. The collected data in WP4 will be used to inform microbiological sampling of WP1, passenger locations in the air flow simulations in WP 2, and passenger occupancy levels for air quality analysis in WP3. Note that our primary focus is risk of in-vehicle virus transmission.				
Personal Information Custodian (PIC)	David Kelly	Is PIC aware of this DPIA?	Y	Project Sponsor	Samantha Phillips

A DPIA is **mandatory** in certain circumstances. Please tick each box where it likely that the proposal will meet the criteria:

Use profiling or automated decision-making to make decisions that will have a significant effect on people. Significant effects can include financial or legal outcomes, intrusions into private life or restrictions on access to services, opportunities or benefits.		Process special category data (relating to: racial or ethnic origin; political opinions; religious or philosophical beliefs; trade union membership; genetic or biometric data; health; sex life or sexual orientation) or criminal offence data on a large scale.		Make changes to processes and systems that are likely to result in significantly more employees having access to other peoples' personal data, or keeping personal data for longer than the agreed period.	
Use data concerning children or vulnerable people. A person with vulnerability is usually described as someone who is at a higher risk of harm than others.		Process personal data which could result in a risk of physical harm or psychological distress in the event of a data breach.		Process children's personal data for profiling or automated decision-making or for marketing purposes, or offer online services directly to them.	
Systematically monitor a publicly accessible place on a large scale – e.g. through the use of CCTV or Wi-Fi tracking.	X	Process personal data in a way which involves tracking individuals' online or offline location or behaviour.		Match, compare or combine datasets, or have the potential to deny anonymity or re-identify people.	
Use new technologies or make novel use of existing technologies.		Process personal data on a large scale or as part of a major project.		Process personal data without providing a privacy notice directly to the individual.	
Use personal data in a way likely to result in objections from the individuals concerned.		Apply evaluation or scoring to personal data, or profile individuals on a large scale.		Use innovative technological or organisational solutions.	
Process biometric or genetic data in a new way.		Undertake systematic monitoring of individuals.		Prevent individuals from exercising a right or using a service or contract.	

Step 1 – Identify the need for a DPIA	
Explain broadly what your project aims to achieve and what type of data and processing it involves. You may find it helpful to refer or link to other documents, such as a project proposal. Summarise why you identified the need for a DPIA.	WP4 of the project aims to investigate the surfaces people touch in vehicles (and their touching frequencies), as well as the positions of passengers and the directions they face in vehicles (we call them passenger positions). This data can then be correlated with surface swab results and airflow/ventilation modelling to create a picture of the risks of covid19 spread on public transport vehicles. The project as a whole will also involve loadweigh, NetMIS and speed profile to correlate the results with passenger density, passenger travelling time, and vehicle accelerations as well as data about the vehicle configuration such as vents, air filters, and vehicle volume. The results of the surface touch analysis will be linked to the results of virus sampling (WP1), so that eventually we can estimate the amount of virus each passenger may have on his/her hands according to his/her in-vehicle travelling time, passenger density and other factors. The results of the passenger position analysis will be fed into the air flow simulation (WP2), so that we can estimate, in the case of presence of an infected passenger, the possibility of other passengers inhaling virus aerosols according to their positions in the vehicle. Passenger position and number data collected in WP4 will also be used to model air quality within vehicles (WP3). Since we would need to observe passenger movements in vehicles that may be crowded, use of CCTV already installed in the trains would be the best way to avoid disturbance to passengers. The data processing involves manual observation of CCTV images, to code (anonymously) the passenger positions without reference to any personal characteristic of any passenger. This processing is detailed in the attachment 'Analysis Plan' This DPIA is required because of use of CCTV, which contains personal data relating to passengers travelling on the carriages identified as part of this study.

Step 2: Describe the nature of the processing	
How will you collect, use, and delete data? What is the source of the data?	The source is CCTV collected by Transport for London (TfL). The on-train CCTV footage will be from the train carriages selected for microbiological study (surface swabs). Five different carriages on three different train lines have been selected to be part of this study. The vehicle specification for these trains' carriages will be collected to identify the cameras required. This involves roughly total 60 hours of CCTV footage across cameras identified on the carriages selected. TfL will keep a record of all footage provided. The CCTV data will be provided on TfL approved encrypted USB drives following TfL cyber security policies. The CCTV data will be stored in UCL's Data Safe Haven (DSH), a digital environment specifically designed to support research using sensitive data including personal data of NHS patients, has been certified to ISO27001 and is compliant with NHS Digital's Data Security and Protection Toolkit. Only researchers named in the ACA will process the data. All UCL researchers in this project complete annual mandatory training in Information Governance procedures. Researchers are made aware of their responsibilities under the Data Protection and Freedom of Information Acts. Before conducting manual coding, UCL will run an algorithm on all the CCTV video data to blur faces. Timestamps will also be replaced with our own reference numbers, which will be used for comparison of the results of the CCTV analysis with those of the planned sampling and other datasets (see below). These reference numbers are essentially pseudonymised numbers and do not mean anything to the project researchers who will code the videos. We will also split the videos into small pieces. We call the pieces 'pseudonymis video data'. After this video pseudonymisation process, the project researchers will watch the pseudonymis video data and record information (we call this 'coded information') on a spreadsheet (see the attachments for details) all on the DSH. Researchers will only collect the kinds of data specified in the coding framework; no other types of data will be collected.' The coded information will be compared to other data (see below) and will be used in models with other data about the trains and surfaces to create a picture of covid19 transmission. Upon the completion of the project, all the CCTV footage will be deleted by the UCL's Head of Windows Infrastructure Services Group who is responsible for the DSH and who will also provide a data deletion certificate.
Will you be sharing data with anyone?	UCL

Are you working with external partners or suppliers?	UCL
Is there an agreement/contract in place with the third parties? (If so, please provide a copy with the assessment.)	Academic Confidentiality Agreement with UCL
Will the data be combined with, or analysed alongside, other datasets held by TfL? If so, which ones?	The coded information will be combined with other data. In particular, we will compare the coded information to the following datasets: NETMIS (service operation records), Loadweigh (passenger loading) in order to relate the analysis results with passenger density and vehicle travelling time. We would like to access to track topology, vehicle acceleration and braking graphs in order to understand acceleration (in the directions both in parallel with and at a right angle to the travelling direction).
How and where will the data be stored?	UCL Data Safe Haven
Will any data be processed overseas?	No
You might find it useful to refer to a flow diagram or other way of describing data flows.	Data flow is as follows: • Request specific on-train CCTVs footage for train carriages identified as part of this study. Approx. total of 60 hours of CCTV footage. • The CCTV data will be provided on TfL approved encrypted USB drives following TfL cyber security policies. • Transfer the data to UCL's Data Storage (Data Safe Haven Server), complying with TfL's specifications. • Pre-coding data pseudonymisation: The videos will be processed by an algorithm that blurs people's faces. Timestamps will be replaced with our reference numbers. The videos will be divided into small sections as well. Data quality check will be carried out to make sure the quality of the pseudonymised video data is good enough for the video coding tasks proposed. • Use within the Data Safe Haven environment: Video coding [using MS Office, IrfanView - Version 4.44, both of which are available on UCL DSH portal and have been security checked). We will record proximities (where passengers stand/sit, which directions they face) and surfaces touched by passengers, as well as the service details (date, Service ID, carriage No, etc). Note that DSH is in compliance with 10 Steps of Cyber Security.

	Note that all data transfers require approval and are carried out through secure portals which are fully audited. Access to the UCL Data Safe Haven is via a remote desktop and this requires multi-factor authentication. The researchers will follow UCL's Offsite Working Policy and Procedure (which is attached to this form) set out for DSH users and security staff in charge of UCL's building entry control systems and CCTVs. Note that DSH has a 15 minutes inactivity timeout which automatically logs out the user. In addition to a strong password each user has to use a six digit number generated by a smartphone app or physical token at each login. Passwords must be changed at regular intervals, and unused accounts are automatically disabled after a fixed period. Anonymized data set to be taken out of DSH for use in further analysis • Deletion: following the TfL confidential data protocols, the data will be deleted by the UCL's Head of Windows Infrastructure Services Group who will also provide a data deletion certificate. The secure deletion process uses a multi-pass overwriting procedure to ensure the data is not recoverable after deletion

Step 3: Describe the scope of the processing	
	The scope of the processing is shown in the attached 'Analysis Plan'. The proposed data coding sheets are also attached.
Who does the data relate to?	Travelling passengers who are traveling in the carriage selected as party of this study
How many individuals are affected?	Not known
Does it involve children or vulnerable groups?	It may involve these groups, but their characteristics will not be recorded in the passenger position data set.
If children's data is collected and used, are they aged under 13?	It may involve children under 13, but their characteristics will not be recorded in the passenger position data set.
What is the nature of the data? (Specify data fields if possible; For *example, name, address, telephone number, device ID, location, journey history, etc.*)	CCTV footage of passengers on the train's carriages selected as part of this study.

Specify which special category data or criminal offence data are to be processed?	The raw CCTV footage may allow inference of the following special category data. - racial or ethnic origin; - religious or philosophical beliefs; - health Although some of the observed passengers may be identifiable as having characteristics that fall into these special categories, we will not code any information about special categories of personal data.
Can the objectives be achieved with less personal data, or by using anonymised or pseudonymised data?	Although we will be processing personal data, we will not use personal data as a part of our analysis. Since we need information about passenger position and surface touches, there is no way to develop an algorithm or other means to predict surface touches without a source of actual touches/positions to serve as a baseline. It would be prohibitive, as well as lessen the accuracy of the result if we can not base position and touches on actual observation through CCTV.
How long will you keep the data? Will the data be deleted after this period? Who is responsible for this deletion process?	The raw data (CCTV images) will be deleted upon the completion of the blurring algorithm. TfL will retain the raw CCTV until 22 January 2022. The applicant and the UCL DSH manager are responsible for this deletion process. The DSH manager will provide a certificate upon deletion of the data.
Is the data limited to a specific location, group of individuals or geographical area?	We will use CCTV of three London Underground lines. Currently we are discussing this selection in the Project Liaison Group.

Step 4: Describe the context of the processing	
Is there a statutory basis or requirement for this activity?	TfL public task to deliver a safe and reliable network that promotes healthy living and the outcome of this research will support this. The Mayor develops and implements policies for the promotion and encouragement of safe, integrated, efficient and economic transport facilities and services to, from and within Greater London (GLA Act 1999, Clause 141(1)), and this function is also exercisable by Transport for London (Clause 38(2)).

What is the nature of TfL's relationship with the individuals? *(For example, the individual has an oyster card and an online contactless and oyster account.)*	Customers within the CCTV footage
How much control will individuals have over the use of their data?	None, although the only data processed (position and touch) will not involve personally identifiable information
Would they expect you to use their data in this way?	Passengers are likely to be aware of extensive CCTV in and around public transport networks. Signage is prominent and the privacy page (https://tfl.gov.uk/corporate/privacy-and-cookies/cctv) informs passengers how TfL, including its operating subsidiaries, use personal data collected. On a case by case basis we may use and share CCTV images for research and analysis purposes.
Are there prior concerns over this type of processing or security flaws?	CCTV is a well-established technology. Facial recognition is increasingly controversial. We will be blurring faces using on an algorithm and the coded information with be only capturing position and touch. There will be no use of facial recognition to identify or single out individuals.
Is it novel in any way, or are there examples of other organisations taking similar steps?	We understand that Newcastle University has been given access to CCTV for similar purposes.
What is the current state of technology in this area?	There is no technology to automate the process Although algorithms can estimate the distance between people and some other large objects, this is not sufficient to determine if an individual is touching a surface with their hands nor to determine where their hands are located. Both data points are needed in this research to generate accurate results about the transmission of Covid 19. A number of methods of blurring faces in CCTV footage have been identified.
Are there any security risks?	See Section 8 for risks we identified.

Are there any current issues of public concern that you should factor in?	No, CCTV monitoring of public areas is common and on public transport is common and accepted. Passengers are likely to be aware of extensive CCTV in and around public transport networks. Signage is prominent and the privacy page (https://tfl.gov.uk/corporate/privacy-and-cookies/cctv) informs passengers how TfL, including its operating subsidiaries, use personal data collected. Indeed, this research falls within TfL's scope for its use of CCTV set out at https://tfl.gov.uk/corporate/privacy-and-cookies/cctv Our use is within one of the purposes: '*Protecting the health and safety of employees, customers and members of the public*'. It also states: '*On a case by case basis we may use and share CCTV images for research and analysis purposes*', and these clauses cover its data sharing of CCTV with UCL for this project.
Are you or your delivery partner signed up to any code of conduct or certification scheme?	Yes. DSH has been certified to ISO27001 and is compliant with NHS Digital's Data Security and Protection Toolkit. There is a user training regime for DSH users (i.e. any new user needs to receive NHS Digital's Data Security Awareness training even though the data is not of NHS) and we will follow this. Further, the DSH complies with the National Cyber Security Centre's 10 Steps to Cyber Security.

Step 5: Describe the purposes of the processing	
What do you want to achieve?	To extract the following information • the surfaces people touch in vehicles (and touching frequencies) • the positions of passengers in vehicles and the directions they face (we call them passenger positions). Combining this data with other findings should enable us to create better understanding of how Covid19 is transmitted on public transport vehicles and to recommend measures to reduce transmission. Our hypothesis is that both surface touches and passenger positions are a function of passenger density inside the vehicle, train travelling time between adjacent stations, etc, and hence we would like to compare the results with such potential factors.
What is the intended effect on individuals?	No adverse effect.

What are the benefits of the processing – for TfL, for other external stakeholders, for the individuals concerned and for society in general?	For TfL and other stakeholders The research will provide a better and quantified knowledge of COVID-19 infection risk within London Underground systems. It also provides basic evidence for any future pandemic. The quantified knowledge can be used for, for example, setting a threshold for the crowding level acceptable on the London Underground network and for improving the vehicle cleaning regime. For individuals concerned and for society in general London Underground passengers and the general public will have a better understanding of virus transmission risks. In addition, the scientific evidence this research will provide on virus transmission risks on London Underground, which is one of the main transport modes in London, would allow the Government and GLA to develop better policies in control of viruses transmission in the public, which will benefit the public as well.

Step 6: Consultation process	
Consider how to consult with relevant stakeholders: Describe when and how you will seek views from the individuals whose data you will be collecting – or justify why it's not appropriate to do so.	Because we are not coding personal data and the data we use will be anonymised, we think consultation with passenger stakeholders is not required. However, we will consult with other stakeholders through two key mechanisms. First, with relevant TfL members, we have formed Project Liaison Group (PLG) which is led by Marian Kelly (Health & Safety for London Underground) and includes people from relevant sections across TfL. PLG will have monthly meetings, which will be the main channel between TfL and the project team. Secondly, the project will jointly participate in a Project Advisory Group (PAG) with DfT-led TRACK project, which aims to assess virus transmission risks in use of public transport services and uses CCTV data of TfL, Network Rail and other Train Operating Companies. The PAG will include academics and practitioners which subject expertise and have regular meetings with the project team. The PAG has not been confirmed yet at the stage of this application. We expect that once the project has results, we may publish the results to TfL, and in some cases to the general public, but the timings and the contents will be discussed and agreed by the PLG beforehand.
Who else do you need to involve within TfL?	• Marian Kelly (Heath and Safety for LU) • David Kelly (CCTV manager) • Peter Harries (Service Planning) • Andrew Hyman, Lisa Almond (Academic research) • Rita Scollan, Simon Guild (Data protection)

Have you discussed information security requirements with CSIRT? computer security **incident response** teams (**CSIRTs**)	Yes, we have discussed the project with managers of the DSH who maintain a documented incident response here: https://www.ucl.ac.uk/information-security/technical-advice/incidentresponse CSIRT have reviewed UCL Data Safe Haven
Do you plan to consult with external stakeholders? If so, who?	We will liaise with the DfT-led TRACK project. Our contact is [redacted], Assistant Private Secretary to the Chief Scientific Adviser, DfT Office for Science, Department for Transport
Who will undertake the consultation?	Dr Lena Ciric (UCL Department of Civil, Environmental and Geomatic Engineering) who is the overall lead of the project
What views have been expressed by stakeholders?	The UCL has had two meetings with PLG. As far as we are aware, there is no major issue or obstacle.

Step 7: Assess necessity and proportionality	
Describe compliance and proportionality measures, in particular: Does the processing actually achieve your purpose?	Data Processing is necessary as we believe CCTV is the only data which can answer our research question. The processing of CCTV data into anonymous passenger position and touch information is fundamental to achieving the purpose of being able to show how these factors, as well as crowding, affect the transmission of covid19.
Is there another way to achieve the same outcome?	The behaviour of passengers can be physically observed, but this would mean that observers would need to travel on the trains for long periods and their observation may be challenged by passengers.
How will you prevent function creep?	The scope of the CCTV coding has been well defined and function creep is unlikely. The defined research project also makes it easier for the research team itself to identify changes in purpose that might lead to function creep. We will set up a review group that will continuously monitor our data use and our strong project definition makes it easier for the group to identify function creep if they see it. The Personal Information Custodian of the project will lead this group.

	Project Liaison Group will function as the supervisor of the project for Data Protection compliance. We will report our data usage in each meeting. In cases where a change in purpose is identified by either the review group or the researchers we will review and update this DPIA or undertake a new DPIA to reflect changes in the purpose or the means by which we process personal data. The process will involve consultation with TfL Privacy Team and the PLG. The terms of the processing will be tightly described in the ACA
How will you ensure data quality and data minimisation?	Data minimisation: - The videos will be processed by an algorithm that blurs people's faces to minimise the personal data viewed - Timestamps will be replaced with reference numbers, which will be used for comparison of the results of the CCTV analysis with those of the planned sampling and other datasets (see below). These reference numbers are essentially pseudonymised numbers and do not mean anything to the project researchers who will code the videos. - The videos will be split into small pieces, call the pieces 'pseudonymis video data'. Data quality: - Data quality check will be carried out to make sure the quality of the pseudonymis video data is good enough for the video coding tasks proposed. Clips without good quality will be removed.
What information will you give individuals about how their data is used?	TfL's privacy page (https://tfl.gov.uk/corporate/privacy-and-cookies/cctv) informs passengers how TfL, including its operating subsidiaries, use personal data collected via CCTV across London's transport network. On a case by case basis we may use and share CCTV images for research and analysis purposes. This DPIA will published on privacy pages and TfL to consider adding a statement to cover COVID research.
What measures do you take to ensure suppliers processing personal data on our behalf provide adequate assurances about their ability to process this data safely and lawfully?	Academic Confidentiality Agreement Project Liaison Group will function as the supervisor of the project for Data Protection compliance. UCL will report our data usage in each meeting
To be completed by Privacy & Data Protection team	

What is the lawful basis for processing?	TfL public task to deliver a safe and reliable network that promotes healthy living and the outcome of this research will support this.
How will data subjects exercise their rights?	TfL's Access your data and your information rights webpages covers how data subject can exercise their rights for London Underground CCTV. Requests relating to this research with be coordinated with UCL.
How do we safeguard any international transfers?	n/a (data will stay within the UK)
Could data minimisation or pseudonymisation be applied?	Yes, the following measures will be taken: - Blur the faces in the footage to minimise the personal data viewed - Split the footage into small sections to minimise the personal data viewed by one researcher. - Pseudonymise the timestamp by replacing it with a reference number. The timestamp (Date, Time, Station and Camera location) will be held separately against the reference number in a secure location and will not available to the researcher carrying out the coding process.
Are data sharing arrangements adequate?	Academic Confidentiality Agreement with UCL

Step 8: Identify and assess risks

Describe source of risk and nature of potential impact on individuals. Include risks of damage or distress as well as associated compliance and corporate risks as necessary.	**Likelihood of harm** Remote, possible or probable	**Severity of harm** Minimal, significant or severe	**Overall risk** Low, medium or high
Data will be intercepted during the file transfer	Possible	Significant	Medium

Password is stolen by a third party, and the data is accessed by them on UCL DSH	Remote	Significant	low
The video analysis process is intentionally or unintentionally seen by a third party	Possible	Significant	Medium
Researcher sees someone they know in the footage	Possible	Significant	Medium
Researcher sees an incident such as a crime or injury	Possible	Significant	Medium
Project scope or function creep	Possible	Significant	Medium
Objections to use of data for this purpose	Possible	Minimal	Low

Step 9: Identify measures to reduce risk				
Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 8				
Risk	**Options to reduce or eliminate risk**	**Effect on risk** Eliminated, reduced or accepted	**Residual risk** Low, medium or high	**Measure approved** Yes/no
Data will be intercepted during the file transfer	• The CCTV data will be provided on TfL approved encrypted USB drives following TfL cyber security policies Then	Reduced	Low	Yes

	connect it to the DSH via their portal system, to transfer all the data over to the DSH and return the TfL approved encrypted USB drives to TfL. • Use DSH, data transfer protocol: All data transfers require approval and are carried out through secure portals which are fully audited. Access to the UCL Data Safe Haven is via a remote desktop and this requires multi-factor authentication. In addition to a strong password each user needs to use a six-digit number generated by a smartphone app or physical token at each login. Passwords must be changed at regular intervals, and unused accounts are automatically disabled after a fixed period.			
Password is stolen by a third party, and the data is accessed by them	• Strict password management (no to write on any or record on the computer; regular change) 2 factor authentication (you have to enter a 6-	Reduced	Low	Yes

	digit number sent to your smart phone)			
The video coding process is intentionally or unintentionally seen by a third party	• Only researchers named in the ACA will process the data. • All UCL researchers in this project complete annual mandatory training in Information Governance procedures. • Researchers are made aware of their responsibilities under the Data Protection and Freedom of Information Acts. • The researchers will follow UCL's Offsite Working Policy and Procedure set out for DSH users and security staff in charge of UCL's building entry control systems and CCTVs. Note that DSH has a 15 minutes inactivity timeout which automatically logs out the user. • Use of spyware detection software. The UCL Data Safe Haven implements	Reduced	Low	Yes

	[redacted] which provides firewall, IDS/IPS and malware protection to the environment. All end points within the environment are protected by [redacted] Data Safe Haven ISO27001 SoA reference: 12.2.1 Controls against malware			
Researcher sees someone they know in the footage	• Use of an effective blurring algorithm. Researcher will flag footage where the blurring algorithm has failed to blur any of a specific person's face to a significant degree. The algorithm parameters would then be tweaked, and the blurring algorithm reapplied. • Researcher will stop watching and flag the footage. • Training to make sure to keep confidentiality and to not add any personal attributes to results • The researcher carrying out the coding will not	Reduced	Low	Yes

	have access to when the footage was recorded.			
Researcher sees an incident such as a crime or injury	• TfL LU CCTV manager will review if any incidents were reported during the time period if so, this footage will not be provided. • Use of an effective blurring algorithm. Researcher will flag footage where the blurring algorithm has failed to blur any of a specific person's face to a significant degree. The algorithm parameters would then be tweaked, and the blurring algorithm reapplied. • The researcher should stop watching and any incidents should be flagged to TfL to ensure appropriate action is taken. • The researcher carrying out the coding will not have access to when the footage was recorded.	Reduced	Low	Yes
Project scope or function creep	• We will set up a review group that will continuously monitor our data use and our strong	Reduced	Low	Yes

	project definition makes it easier for the group to identify function creep if they see it. The Personal Information Custodian of the project will lead this group. • Project Liaison Group will function as the supervisor of the project for Data Protection compliance. We will report our data usage in each meeting. • In cases where a change in purpose is identified by either the review group or the researchers we will review and update this DPIA or undertake a new DPIA to reflect changes in the purpose or the means by which we process personal data. The process will involve consultation with TfL Privacy team and the PLG. • The terms of the processing will be tightly described in the ACA			

Objections to use of data for this purpose	• Total volume of recorded footage involved is small and it is unlikely that the same person would appear more than once in footage obtained. • There are number of safeguards in place as described in this DPIA. • This DPIA will be published on TfL privacy page.	Reduced	Low	Yes

Step 10: Sign off and record outcomes		
Item	**Name/date**	**Notes**
Measures approved by Privacy Team:	Rita Scollan 06/10/2020	Integrate actions back into project plan, with date and responsibility for completion.
Residual risks approved by Privacy Team:	Simon Guild 3/11/2020	If accepting any residual high risk, consult the ICO before going ahead.
Privacy & Data Protection team advice provided:	Rita Scollan 06/10/2020 Simon Guild 3/11/2020	Privacy & Data Protection team should advise on compliance, Step 9 measures and whether processing can proceed.
Comments/recommendations from Privacy and Data Protection Team:	Further documentation, completion of data protection training and an Academic Confidentiality Agreement have been provided and approved to ensure appropriate processing and security is applied to the CCTV data. Recommend providing a sample of the redacted images to carry out user acceptance testing to ensure effective blurring algorithm which still enables the researcher to carry out the video coding task.	

DPO Comments:		
PDP Team / DPO advice accepted or overruled by (this should usually be the Project Sponsor):	Accepted Samantha Phillips 22/11/2020	If overruled, you must explain your reasons below.
Comments:		
Consultation responses reviewed by:		If your decision departs from individuals' views, you must explain your reasons.
Comments:		
This DPIA will kept under review by:	Lena Ciric (UCL) Rita Scollan (TfL)	The DPO may also review ongoing compliance with DPIA.

Glossary of terms

Anonymised data	Anonymised data is information held in a form that does not identify and cannot be attributed to individuals. Anonymous information is not subject to the GDPR, and, where possible and appropriate, should be used in place of identifiable or pseudonymised personal data, particularly where sharing information with third parties or contemplating publication of data. Anonymised data will often take the form of statistics. If you are reporting statistics on a small number of individuals, or there is a level of granularity that allows reporting on small groups of individuals within the overall data set, you must exercise caution to avoid inadvertently allowing the information to be linked to an individual. If information can be linked to an identifiable individual the data is not anonymous and you must treat it as personal data.

Automated Decision Making	Automated Decision Making involves making a decision solely by automated means without any meaningful human involvement. Automated Decision Making is restricted and subject to safeguards under the GDPR. You should consult with the Privacy and Data Protection team before rolling out a process involving Automated Decision Making based on personal data.
Biometric data	Biometric data is a general term used to refer to any computer data that is created during a biometric process. This includes test samples, fingerprints, voice recognition profiles, identifiers based on mouse movements or keystroke dynamics and verification or identification data excluding the individual's name and demographics. Biometric data is subject to additional safeguards under the GDPR when it is processed for the purpose of identifying individuals.
Data breaches	A 'personal data breach' means a breach of security leading to the accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, personal data that is transmitted, stored or otherwise processed. Personal data breaches must be reported immediately to DPO@tfl.gov.uk.
Data minimisation	Data minimisation means using the minimum amount of personal data necessary, and asking whether personal data is even required. Data minimisation must be considered at every stage of the information lifecycle: • when designing forms or processes, so that appropriate data are collected and you can explain why each field is necessary; • when deciding what information to record, you must consider what information is required, what is relevant and whether any information is excessive; • when deciding whether to share or make use of information, you must consider whether using all information held about an individual is necessary for the purpose. Disclosing too much information about an individual may be a personal data breach. When deciding how long to keep information, you must consider what records you will need, and whether some personal data can be deleted or anonymised.
Data Protection Rights	The GDPR provides the following rights for individuals: • The right to be informed; • The right of access; • The right to rectification; • The right to erasure; • The right to restrict processing; • The right to data portability;

	- The right to object; - Rights in relation to automated decision making and profiling.
Data quality	The GDPR requires that "*every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay*." This means you must take steps to ensure that the data you use is sufficiently accurate, up to date and comprehensive for your purposes, and that you take steps to effectively mitigate any detriment to individuals that is likely to result from inadequate data.
Function creep	Function creep describes the gradual widening of the use of a technology or system beyond the purpose for which it was originally intended, especially when this leads to potential invasion of privacy. Review and update your DPIA, or undertake a new DPIA to reflect changes in the purpose or the means by which you process personal data.
Genetic data	Genetic data is personal data relating to the inherited or acquired genetic characteristics of a natural person which result from the analysis of a biological sample from the natural person in question, in particular chromosomal, deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) analysis, or from the analysis of another element enabling equivalent information to be obtained.
Marketing	Direct marketing is "the communication (by whatever means) of advertising or marketing material which is directed to particular individuals". This covers all advertising or promotional material directed to particular individuals, including that promoting the aims or ideals of not-for-profit organisations. Genuine market research does not count as direct marketing. However, if a survey includes any promotional material or collects details to use in future marketing campaigns, the survey is for direct marketing purposes and the privacy regulations apply. Routine customer service messages do not count as direct marketing – in other words, correspondence with customers to provide information they need about a current contract or past purchase (e.g. information about service interruptions, delivery arrangements, product safety, changes to terms and conditions, or tariffs). General branding, logos or straplines in these messages do not count as marketing. However, if the message includes any significant promotional material aimed at getting customers to buy extra products or services or to renew contracts that are coming to an end, that message includes marketing material and the privacy regulations apply.
Personal data	Personal data is information, in any format, which relates to an identifiable living individual. Personal data means any information relating to an identified or identifiable person (data subject). An identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

	This definition provides for a wide range of personal identifiers to constitute personal data, including name, identification number, location data or online identifier, reflecting changes in technology and the way organisations collect information about people. The definition can also include pseudonymised data (where we hold data that has had the personal identifiers replaced with codenames); depending on how difficult it would be to re-identify the individual.
Privacy notice	A privacy notice must let people know who we are, what we intend to do with their personal information, for what purpose and who it will be shared with or disclosed to. TfL adopts a layered approach to privacy notices, with clear links to further information about: • Whether the information will be transferred overseas; • How long we intend to keep their personal information: • The names of any other organisations we will share their personal information with; • The consequences of not providing their personal information; • The name and contact details of the Data Protection Officer; • The lawful basis of the processing; • Their rights in respect of the processing; • Their right to complain to the Information Commissioner; • The details of the existence of automated decision-making, including profiling (if applicable).
Processing	Doing almost anything with personal data. The GDPR provides the following definition: 'processing' means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction
Profiling	Profiling is the use of personal data to evaluate certain personal aspects relating to an individual, in particular to analyse or predict aspects concerning that individual's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements.
Pseudonymised data	Pseudonymisation separates data held about an individual from information that identifies the individual. This can be achieved by encrypting (hashing) the individuals name, MAC address or ID code, masking an individual's exact location or changing an image to make an individual unrecognisable.

	TfL can hold the same data in identifiable and anonymous form, provided appropriate controls are in place to prevent re-identification of the pseudonymised data. The advantages of pseudonymisation are that it may allow further processing of the personal data, including for scientific, historical and statistical purposes. Pseudonymised data (if irreversible) is not subject to the individuals rights of rectification, erasure, access or portability. Pseudonymisation is an important security measure and must be considered as part of Privacy by Design and Default approach. If you use pseudonymised data you must ensure that an individual can not be re-identified with reasonable effort. The risk of re-identification is higher when information about the same individual is combined. For example, whilst a post code, a person's gender or a person's date of birth would be very unlikely to identify an individual if considered without other reference data, the combination of these three pieces of information would be likely to enable a motivated individual to re-identify a specific individual in most circumstances. If you use a "key" to encrypt or hide their identity you must ensure it is sufficiently protected to prevent the individual being re-identified. A Data Protection Impact Assessment can help you assess whether pseudonymisation is reversible in a given scenario.
Significant effects	A DPIA will be required for processing relating to an individual, or group of individuals that has an effect on their legal status or legal rights, or will otherwise affect them in a significant way. These effects may relate to a persons: • financial circumstances; • health; • safety; • reputation; • employment opportunities; • behaviour; or • choices

Special Category data	Special category data consists of information about identifiable individuals': • racial or ethnic origin; • political opinions; • religious or philosophical beliefs; • trade union membership; • genetic data; • biometric data (for the purpose of uniquely identifying an individual); • data concerning health; or • data concerning a person's sex life or sexual orientation. Information about criminal convictions and offences are given similar protections to special category data under the Law Enforcement Directive.
Statutory basis for processing	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 Act also states that we have a duty to help the Mayor complete his duties and implement the Mayor's Transport Strategy. In particular, we are required to provide or secure the provision of public passenger transport services, to, from or within Greater London. As a highway and traffic authority for GLA roads, we regulate how the public uses highways and we are responsible for: • Traffic signs • Traffic control systems • Road safety • Traffic reduction We are also the licensing authority for hackney carriages (taxis) and private hire vehicles (minicabs). 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 and to provide or secure the provision of public passenger transport, 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. Further miscellaneous powers are set out in Schedule 11of the Act. Activities may have a statutory basis related to other legislation, for instance the requirements to publish information under the Local Government Transparency Code.

Systematic processing or monitoring	Systematic processing should be interpreted as meaning one or more of the following: • Occurring according to a system • Pre-arranged, organised or methodical • Taking place as part of a general plan for data collection • Carried out as part of a strategy Examples of activities that may constitute a regular and systematic monitoring of data subjects include: • operating a telecommunications network; • providing telecommunications services; • email retargeting; • data-driven marketing activities; • profiling and scoring for purposes of risk assessment (e.g. for purposes of credit scoring, establishment of insurance premiums, fraud prevention, detection of money-laundering); • location tracking, for example, by mobile apps; • loyalty programs; behavioural advertising; • monitoring of wellness, • fitness and health data via wearable devices; • closed circuit television; • connected devices e.g. smart meters, smart cars, home automation, etc.
Vulnerable people	A person is vulnerable if, as a result of their situation or circumstances, they are unable to take care of or protect themselves or others from harm or exploitation. All children are considered vulnerable by virtue of their age and immaturity.