

Moving Forward with Data

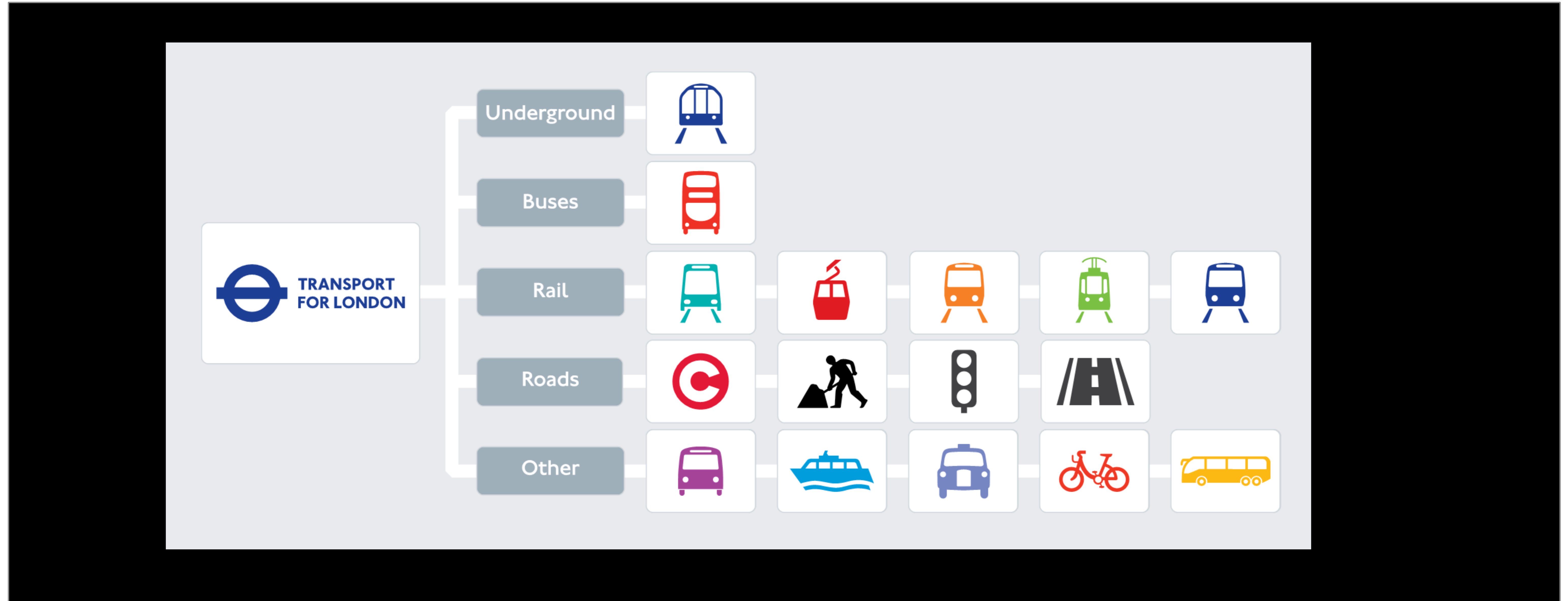
Understanding London's Travel Whist Protecting Privacy

September 2021

Lauren Sager Weinstein
Chief Data Officer



Our responsibilities



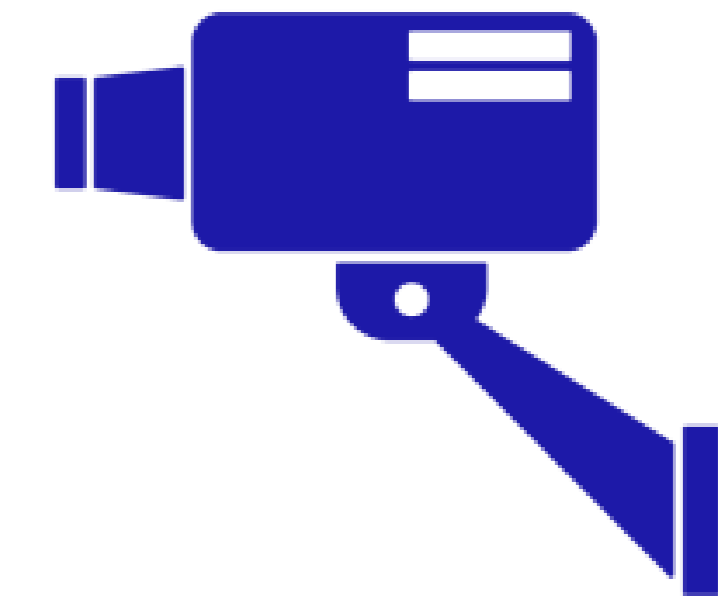
We are data rich...

On a typical (pre-COVID) day



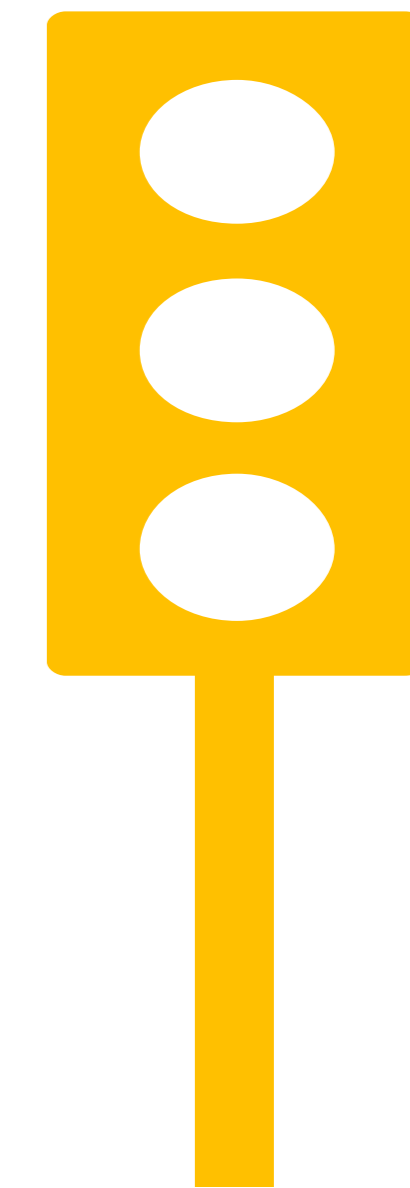
19 million smartcard
ticketing
transactions a day
from **12 million**
active cards

12 million ANPR
registration plates
from the **1600**
cameras across our
road network



500,000 rows of
train diagnostic
data on the Central
Line alone

15,000 SCOOT
detectors creating
5.2bn records



650 million iBus events
daily including **80**
million GPS
records



250,000 daily train location and event
data from NETMIS

DATA ITSELF IS NOT ENOUGH. We must make it useful.



Exploring new opportunities for data with Depersonalised WiFi Data Analytics

Privacy and transparency are key

By transforming depersonalised WiFi connection data into movements could this help us

- Provide better information to **customers** for journey planning and avoiding congestion?
- **Operate and manage** our stations better?
- **Plan** timetables, upgrades etc. more efficiently?
- By measuring footfall, could we generate additional **income to reinvest** in our services?



www.tfl.gov.uk/privacy

Wi-Fi data collection

We collect Wi-Fi connection data at this station to better understand journey patterns and improve our services

We will not identify individuals

You can opt out by turning off your device's Wi-Fi

For more information visit tfl.gov.uk/wifi-data-collection

TRANSPORT FOR LONDON

Plan a journey Status updates Maps Fares Help & contacts

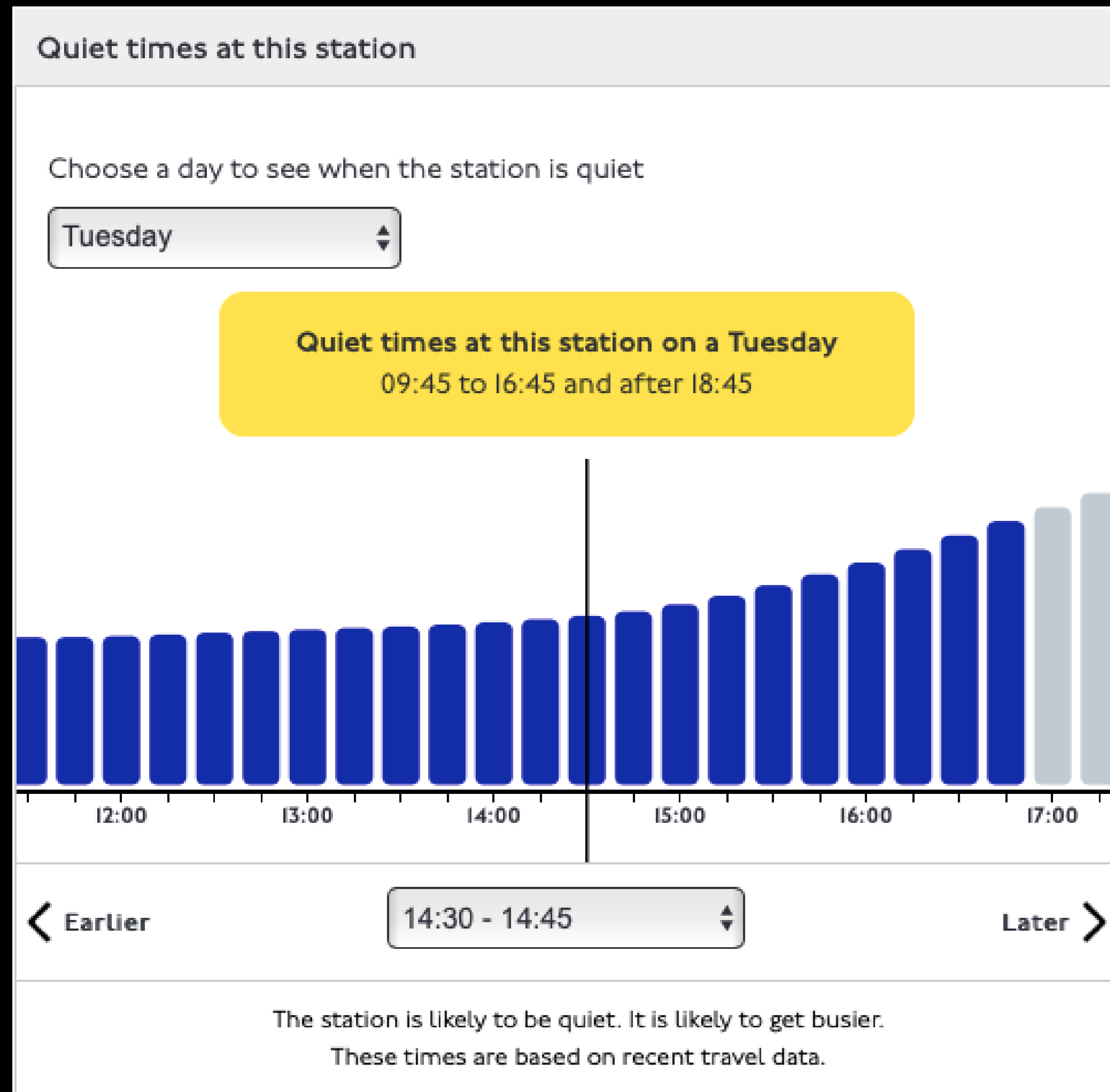
Privacy & cookies Wi-Fi data collection

Wi-Fi data collection

This page explains how and why TfL collects Wi-Fi connection data from mobile devices at London Underground stations.

- ▼ What Wi-Fi connection data is
- ▼ Data protection impact assessments
- ▼ How we collect it
- ▼ Length of time we keep Wi-Fi connection data
- ▼ How we make sure we can't identify people
- ▼ Keeping information secure
- ▼ How we process it and how to prevent processing
- ▼ Sharing information
- ▼ Why TfL is doing this
- ▼ Your information rights
- ▼ Legal basis for using this information
- ▼ Changes to this page

Wi-Fi connection data collection began on 8 July 2019.



TfL Website

TfL Go App

Vauxhall

Live Quiet now

Quieter times on Mondays
Usually 09:15 - 16:30 and after 18:45

Victoria

Northbound Platform 1 2 min then 6, 12

Southbound Platform 2 4 min then 8, 10

Location 28 min

Vauxhall

Live Busy now

Quieter times on Mondays
Usually 09:15 - 16:30 and after 18:45

Victoria

Northbound Platform 1 2 min then 6, 12

Southbound Platform 2 4 min then 8, 10

Location 28 min

Vauxhall

Live Very busy now

Quieter times on Mondays
Usually 09:15 - 16:30 and after 18:45

Victoria

Northbound Platform 1 2 min then 6, 12

Southbound Platform 2 4 min then 8, 10

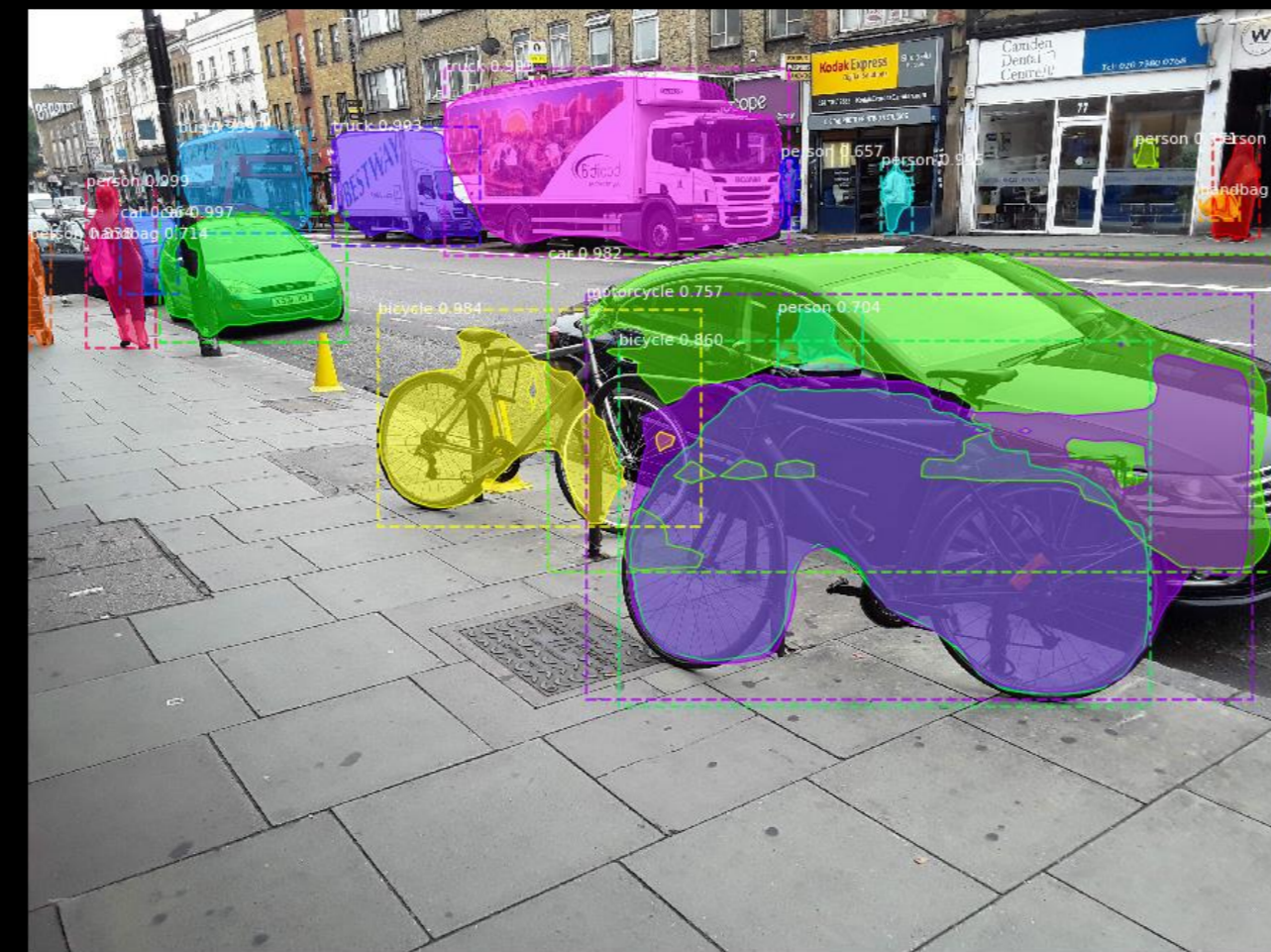
Location 28 min

Aggregated & depersonalised data powers our Customer Information



Image redaction using AI

- Cycling infrastructure database (CID) contains ~500k images of TfL cycling assets (e.g. signs, road markings)
- Images contain **sensitive information** (faces, vehicle registration plates) that needs to be redacted before they can be shared widely within and outside the business
- We used an **object detection and segmentation model** to redact the images

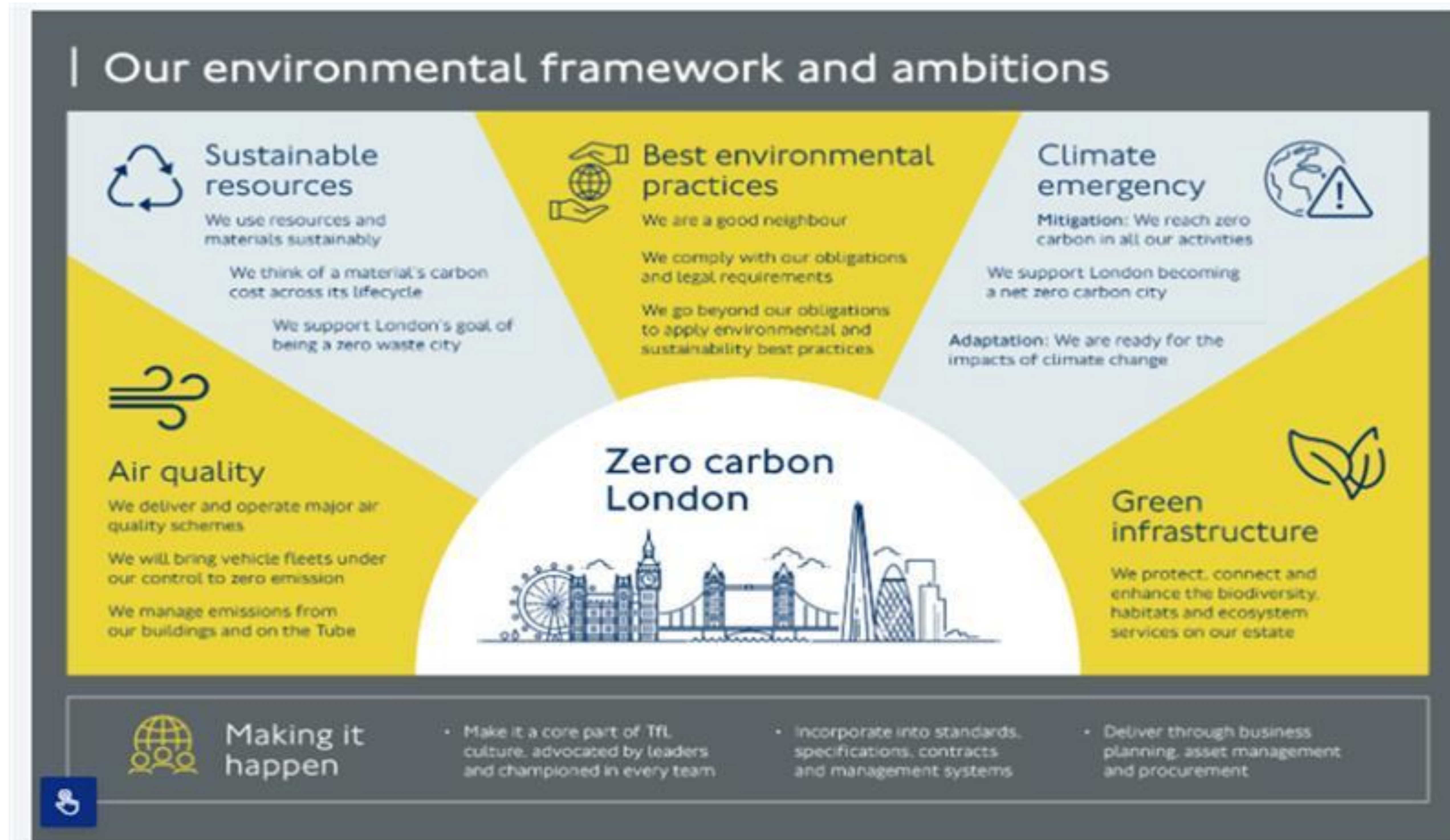


Opportunities

Harnessing data from our sensors to operate the transport network more efficiently

Data driven decision-making to adopt best environmental practices and cleaner infrastructure

Insight we can offer to our customers to encourage green travel



Our upcoming challenge: Use Data to address the Climate Emergency

