Connected Systems for Connected journeys

SOCIAL DISTANCE MANAGEMENT IN PUBLIC TRANSPORT

Version 1.1

www.journeo.com | info@journeo.com

Table of Contents

Table of figures

Document control:

Version	Date	Author	Checked	Signed
1.1	16/05/2020	Darren Maher		

www.journeo.com | info@journeo.com

Executive Summary

In this paper, Journeo seek to demonstrate how a series of disparate technologies can be brought together to provide a robust methodology to provide Social Distancing Management in public transport applications on-board buses, trains, trams, at stops and stations throughout the UK.

The UK economy relies on moving large numbers of people to centres of economic benefit, enabling them to fulfil their roles as key workers in our hospitals, construction sites, offices and more. The most effective and environmentally friendly way to do this was, and remains, mass public transit.

COVID-19 and the fear of contracting a virus in contained and often crowded public spaces has resulted in a requirement to rebuild confidence in the safety of public transport. This document will focus on how operators and authorities can demonstrate occupancy through the accurate collection of data and the processing of that data so it can be displayed to potential passengers, **to show that public transport is open for business**. This document primarily looks at on-vehicle data collection and at-stop data collection.

Operators and authorities will then be able to display that there is safe capacity on their network and within their vehicles.

In turn, it is hoped that this will contribute to an increasing level of confidence within public transport.

For data capture on-vehicle, Journeo has found over-door APC sensors to provide unrivalled accuracy when held in comparison against other methodologies, such as MAC harvesting and ticket factoring. When simultaneously ensuring that a vehicle remains safely socially distanced yet open to provide an essential service for key workers, accuracy is paramount.

For data capture off-vehicle, accuracy is more challenging as members of the public will not necessarily be passing through a focal point such as a doorway meaning technologies that rely on line of sight (such as APC sensors and video analytics) become less useful. In these circumstances, the challenge becomes to gauge the best level of approximation available and, in this case, the most suitable technology would be MAC harvesting. Not only will it provide a good level of estimation, the price and rapid deployment of the solution makes it an attractive solution.

The communication of this occupancy data to the right person at the right time presents another challenge. It is, however, a challenge that can be met by leveraging existing data standards prevalent in the UK and deployment through existing infrastructure. Occupancy data of a vehicle in isolation will not, by itself, provide enough information for a passenger to make an informed decision, so there is a requirement to associate vehicle occupancy with trip and service information. This document will also highlight how this can be achieved, both on a vehicle/at a stop and in the back office. Data consumers will then be able to easily ingest this data and interpret it for display within their platforms.

About Journeo

We are a UK-based transport technology systems business, listed on the AIM section of the London Stock Exchange (JNEO).

We provide "critical cost of failure" services to the transport community which capture, process and display essential information to improve journeys.

We serve towns and cities with passenger information, fleet operators with safety and efficiency systems and bring these together in fully integrated solutions for transit hubs such as airports. Our systems contribute to wider smart city initiatives as transport becomes more intelligent and connected.

We are customer-led with over 20 years' experience solving complex operational requirements based on deep technical and market knowledge. We are recognised leaders in the systems engineering and services needed to partner our customers and to install and support their solutions in the factory, on the street and on the vehicle. Our technical resources focus on increasing customer value by developing products and software that meet the transport communities' unique requirements.

Our customers

We are fortunate to work with Passenger Transport Executives (PTEs), local authorities and transport operating companies of all sizes in the UK, Scandinavia and beyond, enabling us to gain a deep understanding of their, and the wider transport community needs'.

Delivered through

Bus occupancy levels

Prior to considering which technology should be used to capture occupancy level data, one must consider how that data is to be used. The goal, in this instance, is to collect occupancy data and publish it to several different endpoints that could include (and may not be limited to):

Consumer	Purpose	Requirement
Real Time Information displays (Typically located in bus shelters and at transport interchanges)	To disseminate information to the public	Real time
Websites and journey planning/transit applications	To disseminate information to the public	Real time
Bus operator back office systems	To monitor vehicle capacity and employee safety	Real time and historic reporting
Local authorities/transport executives	To monitor utilisation of the network	Real time and historic reporting
Drivers	To monitor the number of persons on the vehicle	Real time

These data consumers (with the notable exception of driver interfaces) already ingest data in standardised transport formats, with TXC (TransXchange) forming the basis for most static timetable data and SIRI (Service Interval for Real time Information) in its various guises, forming the real time updates that can be applied to scheduled data. This data is then used to fulfil a variety of requirements, the most recognisable of these being published arrival/departure predictions on information displays and on websites and mobile applications.

How this information should be displayed is covered in the **Displaying the data** section of this document.

For predictions to be generated, a matching process must take place, to associate any scheduled departure with a real time positional update (often referred to as Automatic Vehicle Location or 'AVL' data and often delivered to third-party consumers as SIRI-VM data). Thus, matching service-specific information with trip-specific information.

As occupancy data can be considered trip-specific data, there are two possible routes for matching live occupancy levels with other trip-specific data; matching at source and outputting through existing means, or matching in the back-office. Both methodologies will be explored later within the **Processing the data** section of this document and a preferred route is highlighted to indicate the preferred methodology for resolving this challenge.

The key, however, to any bus occupancy system, is accuracy. Being able to accurately gauge the number of persons on a vehicle is a challenge that has been faced before and, overcome. This is explored in the **Capturing the data** section.

Capturing the data

Bus occupancy data is not something that, historically, has been particularly prevalent in the UK market and, consequently, there has not been a great deal of work undertaken on it by many suppliers. Fortunately, Journeo has worked throughout Europe and, in particular, Scandinavia, where the franchised model of services has resulted in a greater interest in occupancy data and passenger counting equipment has been fitted as standard. Journeo has been working in Scandinavia installing and caring for systems for over a decade.

When we talk of passenger counting, there are many technologies that have been trialled, but only one adopted as the de-facto technology – for one simple reason: it has demonstrated time and again to be a robust and reliable method for counting passengers on and off vehicles

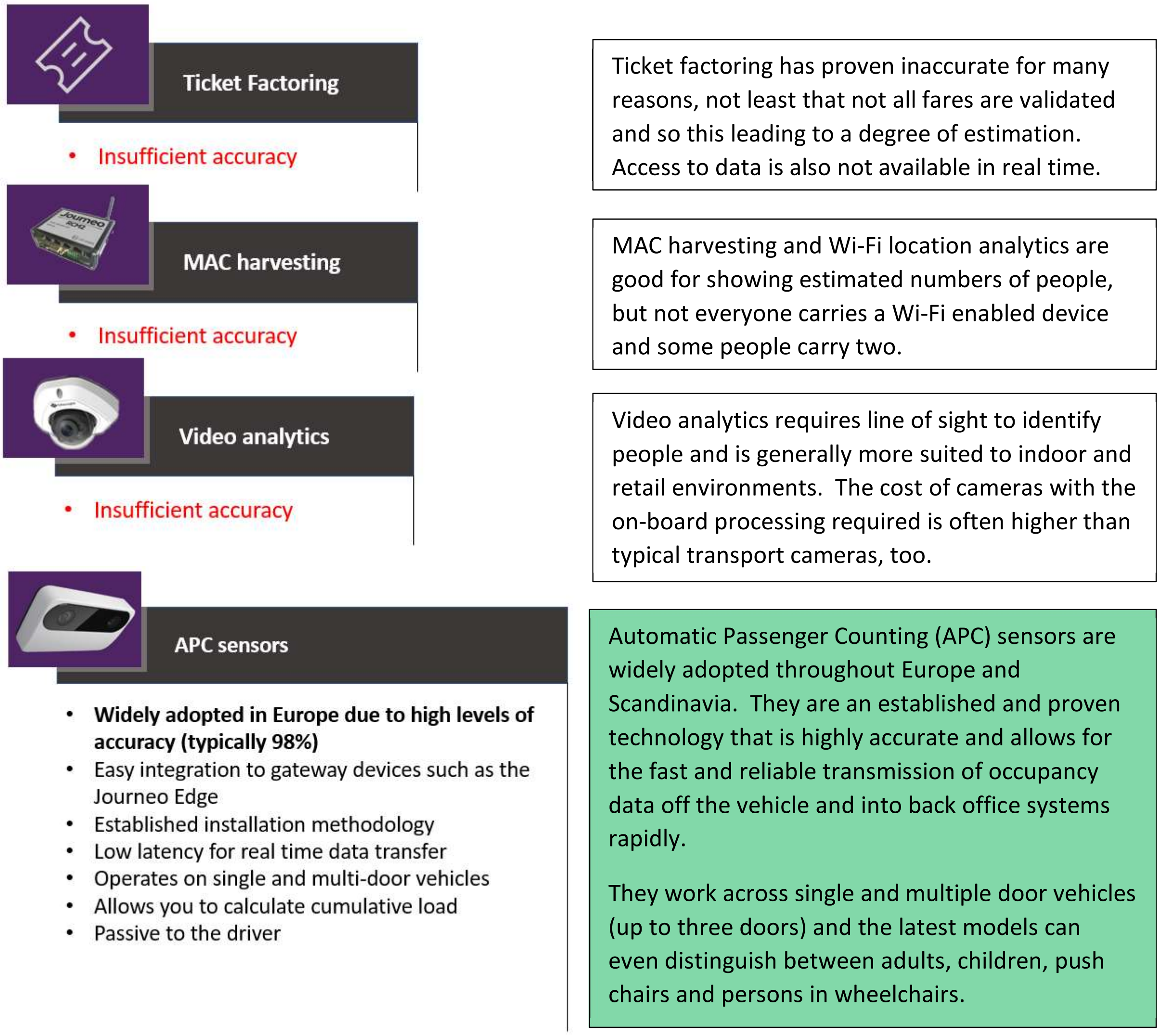

Figure 1 - APC sensors are the industry-proven de-facto standard

Our experience gained in projects across various bus fleets, from those operating in small geographical areas such as airports, to covering entire cities has helped us identify the range of technical solutions and which are best suited to the main use cases, these being to:

- Calculate a real-time bus occupancy load, and/or
- Store boarding and alighting count data for historical analysis

It is possible that both are required yet can use very different technical solutions to achieve them. In the case of real-time loads, the accuracy of the sensor is of primary concern as it has to calculate from boarding and alighting numbers a cumulative calculation without the benefit of adjustment, weighting or post-processing. Whether the resulting number is converted to a broader percentage load, for example, 'Bus is >75% full' is a matter of data presentation, covered in a later section of this proposal.

Should historic analysis be the primary objective, then less accurate sensors may be selected, and information extrapolated within a post-processing environment to observe trends of passenger numbers across the same route or to filter-out any inaccuracies.

With the aim of this project being quite specific, we can focus more on the higher quality sensors that offer better accuracy so that:

- A vehicle occupancy load can be calculated each time the doors close within a matter of seconds

Furthermore, because sensors are almost always connected to a door signal (to know when to start counting and stop), you can further derive

- The vehicle dwell time of each stop

This may be useful to capture because the number of people boarding and dwell time might provide a boarding rate to help derive passengers left behind at the stop, i.e., a faster boarding rate indicates people are ready to board and are doing so in a shorter period than you might have at a less busy stop where vehicles doors are left open for longer to collect passengers.

Fixed-sensor technology

These discreet sensors are typically mounted in housings above the doorway and are observing the space from a top-down perspective. Their image of the bus is therefore limited to passenger heads, shoulders and very rarely, faces. Even if they are able to see a face, they are not storing images in order to operate normally and so cannot identify an individual. For that reason, we call them sensors and not cameras even though some may have what looks to be a camera lens (or lenses) as part of their design.

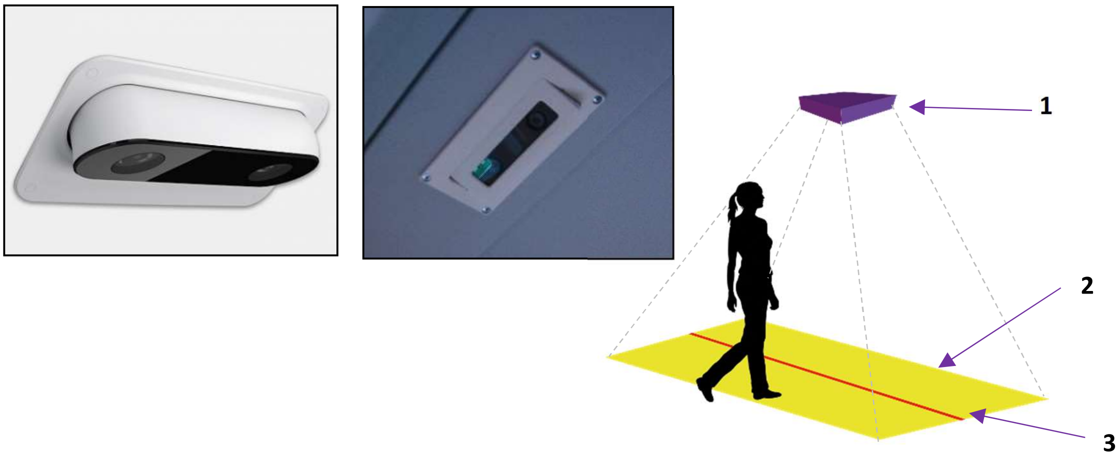

Figure 2 - Fixed sensor and mode of operation

1 Sensor mounted overhead
2 Configured counting area
3 Counting line

Key facts are:

- One sensor needed per door – whether a single or double width
- Can count bi-directional passenger exchanges
- Works in all lighting conditions
- Allows for data validation using remote video access

Other types of sensor are available, for example:

- Pressure sensors
- Beam sensors; or
- Image analytics

However, Journeo and our customer's experience with this sensor type has indicated varying levels of counting success with these sensors, with accuracy generally sitting between the 70 and 90% mark. Overhead fixed sensors, in comparison, can be up to 98% accurate.

Count and occupancy information

Our sensors provide accurate boarding and alighting data at each bus stop or when the vehicle doors close, and our onboard Journeo EDGE device (an on-board gateway used to transmit data off the vehicle) collects this information for onward transmission over 4G to our servers, the Journeo Transit back office.

The EDGE device supports ethernet and serial interfaces to other on-board systems, allowing count data to be transmitted alongside service, route and trip information. This is explored further in the section entitled **Processing the data**.

The EDGE has its own GPS antenna and so the point where counts have been captured will always include the Latitude and Longitude co-ordinates together with the Vehicle Fleet Number that has been programmed into the device at point of installation.

If the on-board systems support the ITxPT information exchange principles, then our EDGE device can act as a data broker to collect this information via a simple Data Subscription or provide the count data as a Data Publisher.

Data capture and transmission

In cases where the EDGE device is transmitting data off-board via the GSM network, it has intelligence allowing information to be stored when a 4G signal is unavailable and can continue doing so for many hours irrespective of the number of stops or passenger exchanges. Once a signal is available, it will forward its cache of data to the back office.

- Information could be sent within 2 seconds of the bus door closing

Where multiple doors are being monitored, the information transmitted can include the aggregate of boarding and alighting and also for each door.

Furthermore, the sensor can be configured to include height-calculated and object-identified attributes within the count data, for example:

- The number of Adult, Children, Buggies/Pram or Wheelchair boarding/alighting

This functionality has yet to be extended to the reporting platform but is currently under review for integration.

The Journeo EDGE device connects to our Journeo Transit back office over a VPN connection using an available GSM network and SIM monitored by us using M2M network tools.

This ensures a secure connection to our Azure hosted Journeo Transit modules to provide a number of ways to view data.

- Secure, scalable and available

The Azure environment we use is built to scale and provide high availability to end-users or third-party systems.

System accuracy

An important question often asked is how accurate a system using sensors is. You might look instead at what level of inaccuracy would be acceptable before the system becomes unusable and so, as we have explored above, it depends on the application of information.

For example, if a real time occupancy number is broadcast to the bus stops ahead on the route, you will want a high degree of confidence in what is displayed on those displays or those waiting could be led to making an incorrect decision. If drivers are being asked to manage occupancy levels based on the sensor data, then they may lose confidence if they can see more or less people on the bus than the system is stating

- Accuracy is dependent on the quality of installation, configuration, and system validation

In our experience, each door should provide an accuracy level in excess of 97% using the sensors that we supply or a tolerance of +/- 3%.

In cases where there are high levels of passenger exchanges, some additional tools may be used to provide periodic 'resets' in order to take the occupancy to a zero level in those cases where very busy routes might naturally accumulate some error.

- Driver switching journey on the ETM
- Driver resets using a discreet button mounted next to the Occupancy Panel (a small read-out showing the number of passengers on the bus).
- A GPS point reset (this might be a specific location where buses are always empty); or
- A Stop reset point (where the bus arrives at the end stop and all passengers disembark)

Deployment

A typical deployment of passenger counting technology comprises the following main parts whether installing onto a new bus in the factory or as part of a retrofit program:

- A sensor per door and in some cases, a surface mount container if it cannot be flush-mounted
- Cabling between the sensor and the Journeo EDGE and between the sensor and vehicle power
- A door signal typically from the door to the Journeo EDGE
- The Journeo EDGE for data collection and transmission
- GSM/GPS antenna
- Optionally, we could provide a small driver panel designed to display the current occupancy to the driver

Each Journeo EDGE will, in turn, be connected to the Journeo Transit servers to collect, store and make available information as required, for example:

- Realtime map with occupancy;
- Historical reports summarising data; or
- Via APIs to consumer Apps, as part of a SIRI feed or other APIs to BODDS

Processing the data

As highlighted earlier in the document, once you have captured occupancy data, you must then match that with other trip specific information in order for it to become usable to systems that the data is delivered in to. There are two ways to do this:

1. Match the data at source – i.e. with other trip specific data-generating devices. For example, providing the occupancy data to the ETM in order for the ETM to publish the information within its AVL (Automatic Vehicle Location) data.
2. Match the data on a gateway – i.e. take the data you need from multiple systems and publish cumulative data from this.

Matching the data at source

The immediate temptation would be to fulfil the brief by extending the scope of current SIRI formats to include occupancy data within real time updates and trip-specific information derived at source. However, there are several obstacles that would present with this methodology:

1. **Established standards take time to change** – The process for agreeing standards changes and suppliers adopting these changes can be time consuming and result in further cost. The current SIRI standard does allow for occupancy data, but only on a vehicle full/not full basis – not the live count. With no way to determine what is considered "full" and no methodology to dynamically adjust that as social distancing rules change, this would become problematic.
2. **Trip-specific information is published by the Electronic Ticket Machine (ETM)** – Whilst not exclusively true, the vast majority of AVL update data is sourced from ETMs, meaning that any data captured by on-board sensors would have to be published back to the ETM for dissemination. Whilst it is likely that this is technically surmountable, this has repercussions, not least that would require the modification of all ETMs and their back-office architecture, to ingest the new data, communicate it to their back-office systems and publish the data.

Match data on a gateway

Through the installation of passenger counting sensors and a gateway device, data can be collected on-board from the sensors alongside route progression data consumed from the ETM. This would then be transmitted off-board, as seen in Figure 3, below.

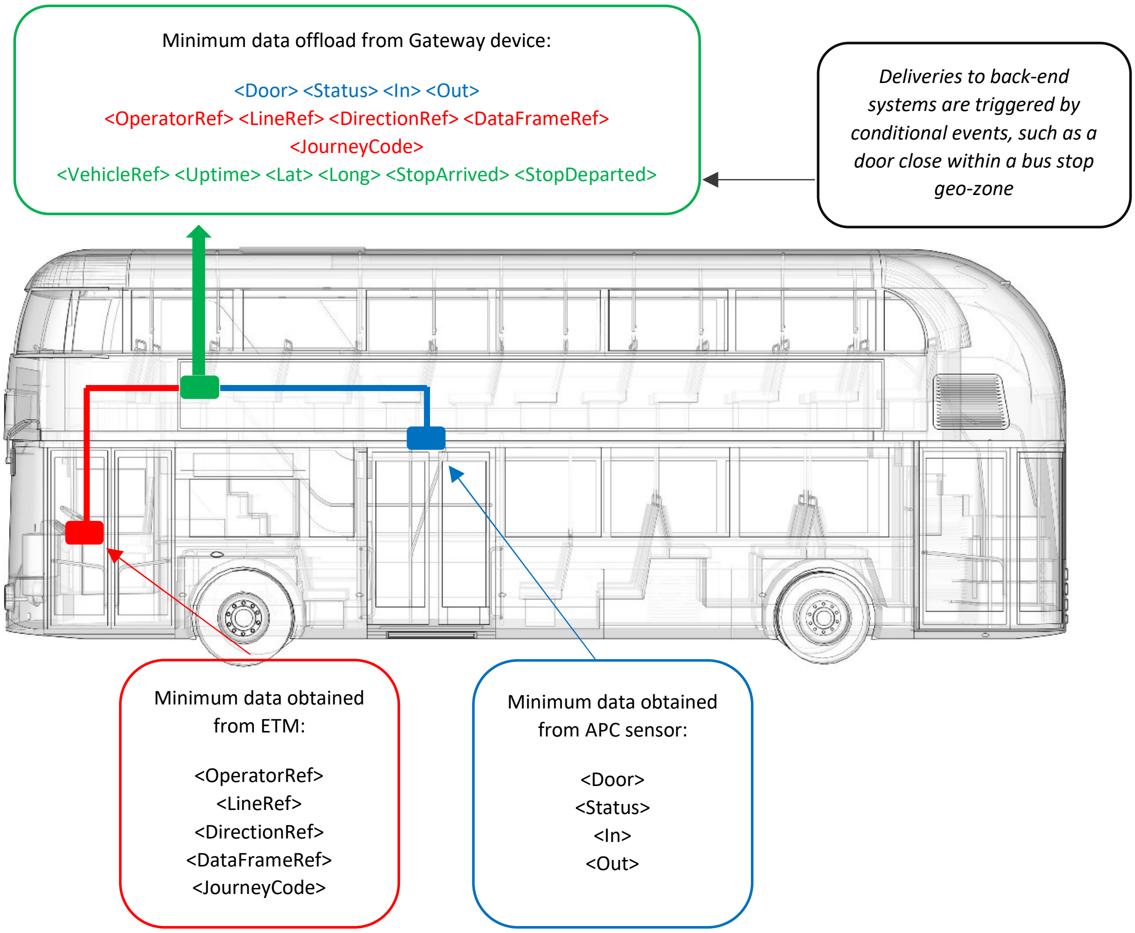

Figure 3 - on-board data matching via gateway device

- **Red data** – is sourced from the ETM and is typically owned by the operator. The data gathered, however, is no different to data that operators will be required to publish to BODDS (Bus Open Data Digital Service) when vehicle tracking information is also required for the service.
- **Blue data –** is sourced from the APC sensors. There is potential that if the project for installation of sensor technology is funded by the authority, the data would be owned by them. Operators, however, would probably view the data as their own commercially sensitive data.
- **Green data –** is sourced from the on-board vehicle Gateway (in this case the Journeo Edge device) and is likely to be subject to the same commercial discussion as blue data.

The second option, to consolidate data via a gateway, presents advantages over the first option:

1. ITxPT methodologies for publishing and subscription of messaging between on-board systems already exists and can easily be adopted if current onboard systems are compliant to the standard.
2. All hardware involved in the installation currently exists, cutting down on development time as only interfaces to ETMs would need to be reviewed on a per-manufacturer basis. Journeo can already interface to many major ETM variants from work completed on previous projects.
3. Interfaces could be rapidly designed and deployed to align to match with current transport data standards, including the publish/subscribe methodology of SIRI or OpenAPI interfaces.
4. There is scope to publish occupancy levels to other on-board systems, such as the vehicle blinds.

Processing the data for display

There are established standards for providing real time updates on bus journeys. The most common being GTFS-RT and SIRI. SIRI is the most prominent standard used within the UK. When processing data for multiple vehicles operating multiple trips, it makes sense to leverage this standard (although not utilise the existing standard that does not allow for specific occupancy data) to publish data to third party systems. This will allow other consumers of the data to similarly consume and display the data with very little modification/development until the SIRI standard allows for such granularity.

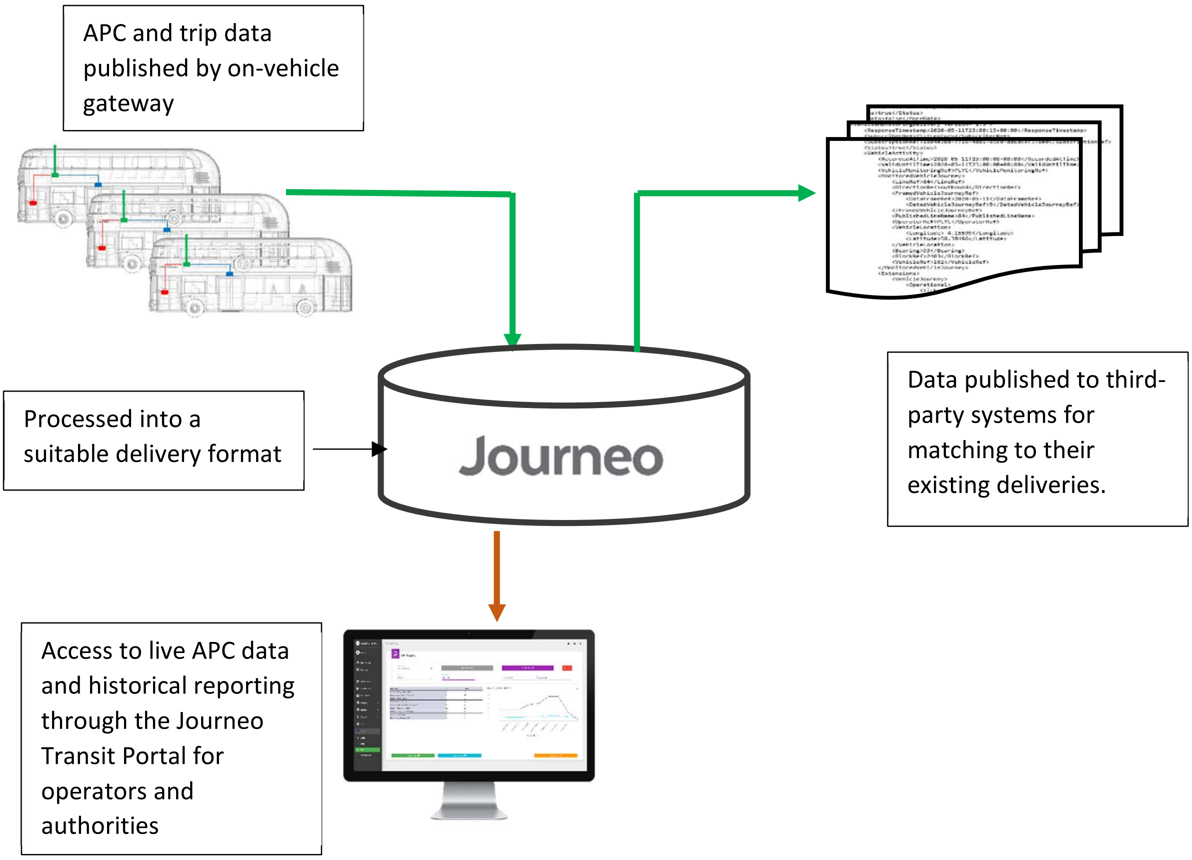

Figure 4 - High level architecture of the solution

To provide suitable information to data consumers, two processes need to take place, matching event information with the trip to form a cumulative count and creating an output that can be matched to existing transport data standards.

1. **Matching event information with the trip to form a cumulative count**: This matching process will associate on/off counts provided by the APC sensors (from single or multiple doors on a vehicle) with data previously stored from a trip, allowing a cumulative live occupancy level to be published. When the ETM signals that a new trip, identified by a new <JourneyCode> (this will usually be generated by the driver logging on to a new trip), this will be taken as an instruction that one trip has finished and another began (inferred from the TXC) and use this as a trigger to implement a zero count after the last count of the previous trip (as seen in Figure 7, below). The Cumulative count would be output as <CumulativeCount>. This process can either take place on the EDGE gateway device, or in the Journeo back office, the only significant difference being the former option providing a greater speed of processing and the latter a smaller payload of data reducing SIM costs.

   The methodology can be seen from Figure 5, below:

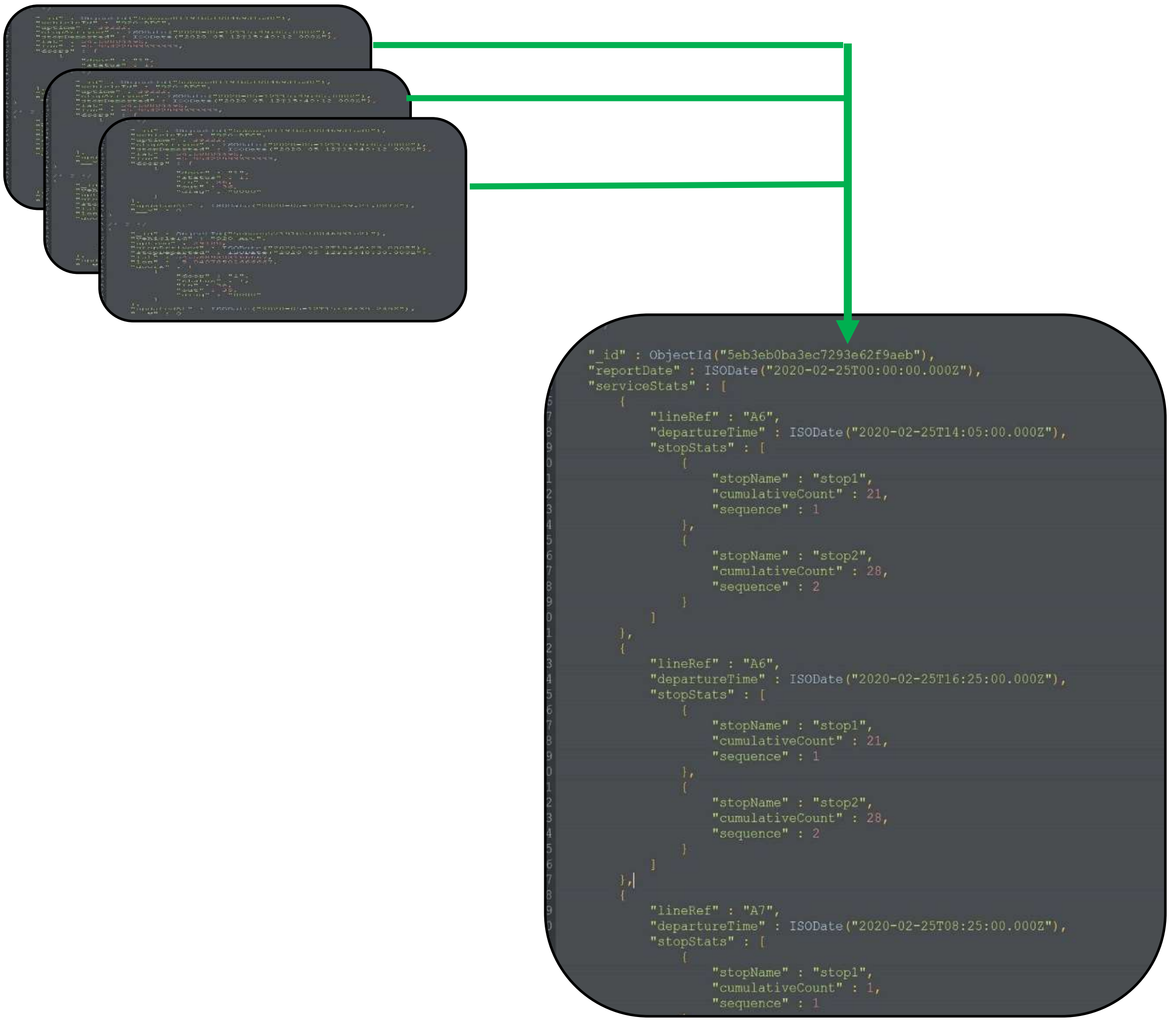

Figure 5 - Successive trip deliveries allow a trip-specific cumulative total to be calculated

2. **Creating an output that can be matched to existing transport data standards:** In order for the solution to work, it is essential that the system generates output data that can be "matched" to existing transport standards, that may be reporting data on the same trip, but for a different purpose, such as SIRI VM/SM/ET data.

   By capturing the requisite fields from the ETM, coupled with the cumulative count by Journeo (On the EDGE or centrally), data can be packaged for delivery to other Journeo systems that display data (such as EPI, the RTI display Content Management System) or that will allow third-party consumers to match against both TXC and SIRI deliveries from other transport systems. Figure 6, below, demonstrates the ability to match the proposed captured data with a SIRI SM delivery. In practice, there is suitable information generated from the Gateway to match against any SIRI format. This will allow the data consumer to overlay the cumulative count data with the other trip specific information.

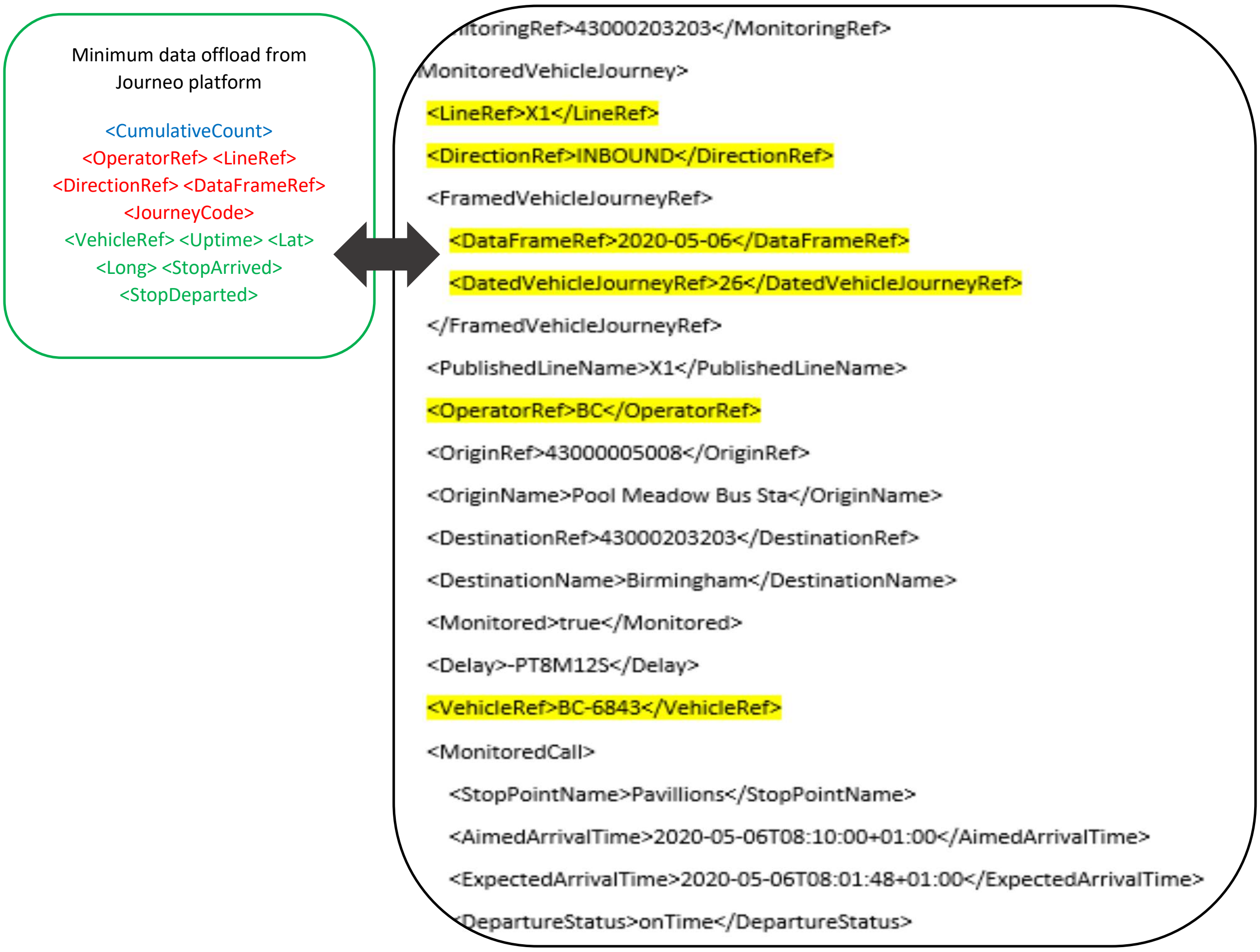

Figure 6 - Matching against other transport data formats

Displaying the data

To infer occupancy levels and display in a meaningful format, an additional data source is required, that could potentially be sourced from data held by the DVLA/O licence/bus registration submissions. Ideally, this data would have to indicate the potential capacity of the vehicle and consumer systems would have to regularly scrape/consume this dataset (in the same manner one would scrape the NaPTAN database) to update vehicle capacity as social distancing rules change.

However, as this data is unlikely to exist, Journeo would be able to collect this information during the installation of each vehicle and cumulatively build a database of this information. An interface could be easily designed to allow authorised users to alter the capacity of each vehicle, as social distancing measures change.

This would allow the consumer to not simply show occupancy numbers, but a more useful occupancy state, such as a RAG status indicating

- **Green** - <75% capacity
- **Amber** - >75% capacity
- **Red** – 100% capacity

RTI displays

Once a consumer has matched the data against existing SIRI updates for services, they will be able to display it in a format suitable to their use case. For example, should the data consumer wish to display the information on TFT in-shelter displays (As Journeo themselves would wish to do through EPI, already in use throughout your RTI estate) they will easily be able to do so, as seen in the mock-up provided in Figure 7, below.

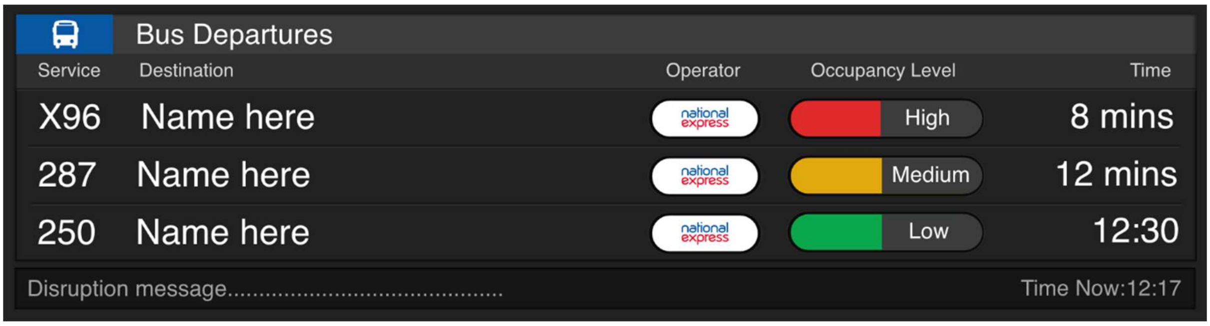

Figure 7 - Occupancy levels can be displayed on TFT RTI displays with ease

The information remains relatively easy to display on high-density colour LED displays, but becomes more problematic and less recognisable on traditional 3-line orange LED displays, such as those seen in figures 8 and 9 respectively, below.

Figure 8 - High pixel density, full colour LEDs will lead to easily recognisable data

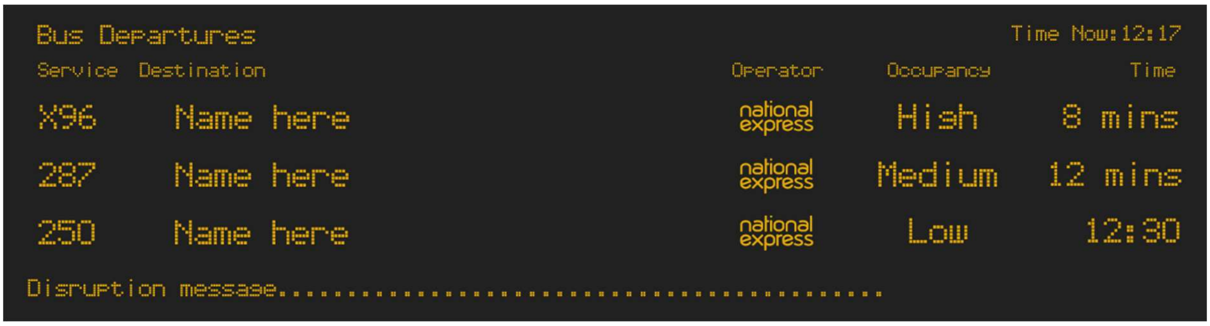

Figure 9 - the information is less recognisable on traditional 3-line LEDs

Website/journey planning mobile applications

By providing access to the data in a format that can be easily matched to existing transport standards, large scale consumers and bedroom developers alike would be able to access the data and match it to their existing datasets, allowing them to present the data as best suited to their individual applications.

Bus operators' back office systems

Bus operators would, through the Journeo portal, be able to see live occupancy information and obtain historical reporting data on stop and service utilisation, furthermore, they would be able to receive alerts (the Journeo portal currently utilises toaster notifications and alert mails to notify users of a metric nearing or exceeding configurable parameters).

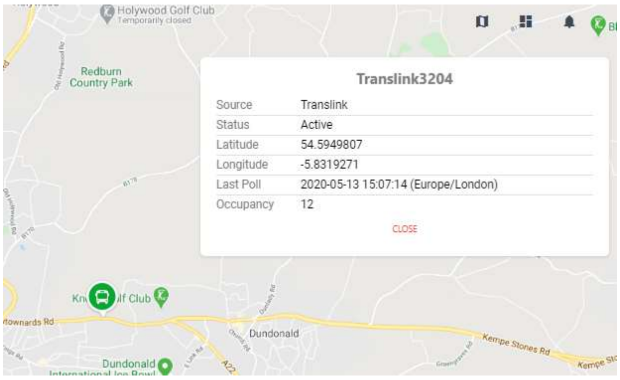

Figure 10 - Live occupancy on the Journeo Portal

Real time occupancy rates (as seen in Figure 10) would allow them to ensure that their vehicles are being operated in a manner that allows safe social distancing.

Local authorities and transport executives

To monitor the utilisation of the network, including high volume bus stops where additional social distancing measures may need to be implemented, local authorities would be able to access historical reporting on network usage via the Journeo Transit platform. This will assist them in making informed decisions about network capacity.

Figure 11 - Stop and service utilisation reports are available on the Journeo platform.

Drivers

It is important that drivers are given the information they need to manage their vehicle safely. Whilst not yet developed, a simple driver interface could be developed to display information on current vehicle occupancy levels.

Bus stop crowdedness

The issue of Bus Stop crowdedness and the presentation of data to the public viewing remotely through mobile applications presents more of a challenge. APC sensors cannot be used as members of the public do not pass a specific “gate” before arriving at the stop.

The more accurate technologies, such as video analytics, would be expensive and time consuming to implement (issues such as line of sight may result in many cameras monitoring a single stop). The natural ebb and flow of people passing bus stops and the constant change in levels of people queuing or congregating around the stop lends itself to delivering an approximate crowdedness level to end users to help them make informed decisions about using public transport.

This could be achieved by again combining multiple datasets; MAC address harvesting at stop, augmented by historic stop usage information collected from the over-door sensors collected from the on-vehicle sensors.

Journeo have previously completed proof of concept trials on MAC address harvesting to estimate groups of persons gathered at bus stops utilising the same Edge gateway device that is proposed for installation on-vehicle.

In the trial, the Edge device scanned available Wi-Fi channels and reported captured data to the Journeo back office in data blobs at 5-minute intervals – this was configurable and the 5-minute interval selected due to the use levels of the bus stop. It is anticipated that high volume bus stops would need a greater level of updates.

MAC addresses reported to back-end systems were hashed to ensure identities were obscured but that relationships could be drawn between collected blobs.

Figure 12 - MAC harvesting/Wi-Fi location analytics would be most suitable for an "at-stop" solution

The proof of concept trial was successful and showed that MAC address harvesting was able to give good indicative data on how many people were in the locality of the stop.

However, in order to differentiate between the level of people in the vicinity of the stop and the amount of people crowded at the stop waiting to board a service, further work would be required to assess dwell time.

As seen in the earlier section, 'Journeo Transit', The Journeo Transit portal collates data on historical bus stop usage, based upon boarding and alighting data from vehicles that have frequented the stop.

Figure 13 - The Journeo portal collects historic APC information

Given further testing, statistical analysis could be performed on this data to determine normal levels of bus stop usage for any given time of day/day of the week.

An algorithm would then be developed to compare known statistical norms for stop usage, aligned to historical statistical norm MAC address quantity captures, against current MAC address quantity capture in order to determine an approximation of the current level of bus stop crowdedness. This could then be delivered to BODDS in the same fashion.

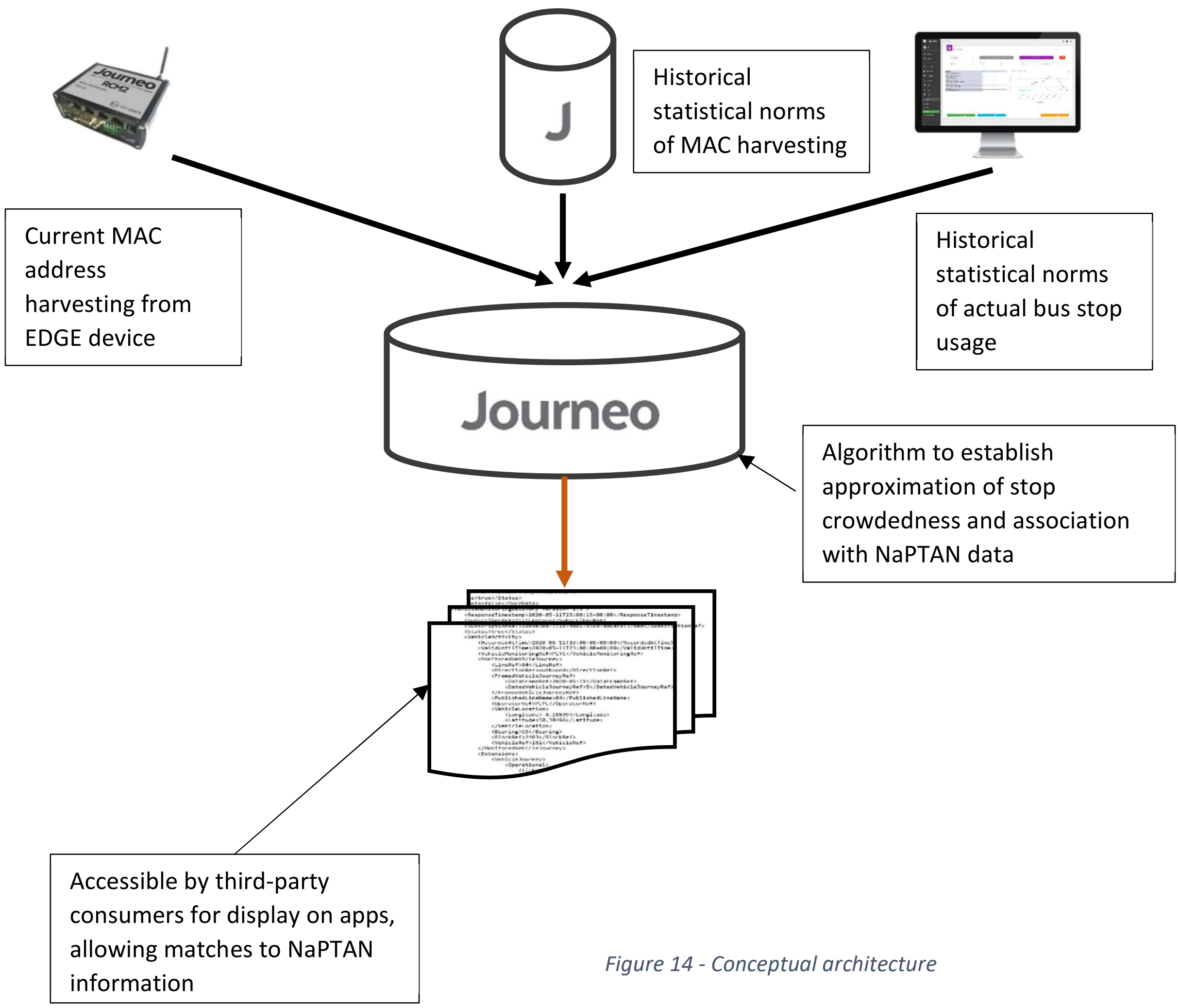

Figure 14 - Conceptual architecture