Task 7.1 Collecting, Maintaining, and Utilizing Transportation Data

Part 1: Future Data Needs

VAUGHAN

HDR

Memo

Date:	Monday, May 25, 2020
Project:	City of Vaughan – Transportation Master Plan Update
To:	Christopher Tam – City of Vaughan
From:	Jonathan Chai, Azadeh Heydari, Jason Zhou, Elli Papaioannou, Dan Lu - HDR
Subject:	Collecting, Maintaining and Utilizing Transportation Data – Task 1

Introduction

HDR has been retained by the City of Vaughan (the City) to undertake the City-wide Transportation Master Plan Update, also known as the Vaughan Transportation Plan (VTP). As part of the VTP, the City would like to explore available, potential data sources in order to more effectively support the City's transportation planning.

Advancements in transportation data collection, including application and GPS-based collection of travel behavior, and in transportation modelling, such as activity-based modelling will change how infrastructure and policies are developed and evaluated in the future. These advancements should be leveraged to maximize the benefit of investments in transportation infrastructure.

In addition, there is an immediate need to understand the significant changes in travel behaviour which are occurring due to the on-going COVID-19 Pandemic, and which may continue to occur after the Pandemic is resolved.

This memo outlines the City's current data sources and their application, emerging data sources, best practices in application of emerging data sources in North America, and a summary of constraints and opportunities.

City of Vaughan's Current Data Sources

The City currently relies on traditional transportation data sources for its transportation planning and modelling efforts. This includes:

1. A periodically updated GIS database for existing transportation infrastructure,
2. The Transportation Tomorrow Survey (TTS) for travel demand (i.e. mode choice, origin-destination),
3. Ground counts which are undertaken typically on a project-specific basis, and this may include road surface traffic sensors, floating car surveys and traffic cameras, and license plate surveys.
4. Collision data which is supplied by York Region Police.

As a lower-tier municipality, the City is reliant on York Region for GIS information for the Regional network and for traffic counts at all Regional intersections. Efforts to consolidate the

City's transportation infrastructure asset data in a GIS database is on-going as part of efforts to meet provincial legislation requirements (O. Reg. 588/17 and related legislation).

For its traffic counts on its own City streets, project-specific counts are typically conducted through third-party survey companies with no formal database maintained by the City. The City is currently reviewing options for data management software to organize this data.

The TTS is conducted every 5 years with a 5% sample size and has well-known limitations due to the survey approach[1]. Respondents are asked for their travel behaviour on a typical day and as such recreational trips, which often include active transportation trips, are known to be under-reported.

The City documents collisions in individual pdf reports but does not currently maintain its own database for collisions on City streets. As other municipalities do manage this information internally (including York Region), the City is currently looking into a data management software for collision data.

Emerging Data Sources

Traditional data collection methods, particularly for travel demand and ground counts noted previously, have been well used among transportation professionals, but with limitations such as the additional cost of every survey, data storage and transfer issues, accuracy issues associated with human errors, small sample size, etc.

Emerging data sources have the potential to be more efficiently collected and readily available to the City, either through third-party data sources or through implementation as part of the City's infrastructure. Three main areas where emerging data sources may be utilized to the City's benefit include:

- Big Data Providers
- Smart Video Detection
- Predictive Data Analytics

These areas and their benefit to the City are discussed in the following sections.

OD Information - Big Data Providers

Current Origin-Destination surveys with known limitations can be supplemented by "Big Data". The proliferation of cellphone usage which can anonymously track location has resulted in the ability to supplement traditional data collection methods by emerging "Big Data" sources which can be utilized efficiently with new collection techniques and data processing methods.

There is currently no consensus among transportation professionals as to the definition of Big Data. However, four characteristics have been identified in Big Data. Note that Big Data

[1]https://www.proquest.com/docview/2322201758/previewPDF/95FBA8B4EC4A4D50PQ/1?accountid=5603

should be distinguished from large datasets which do not necessarily have these characteristics, also known as four "V" as listed below:

- **Volume** refers to the massive amount of data collected from various sources such as mobile apps, Bluetooth, and GPS.
- **Variety** refers to the various industry standards, sampling rates, and data types such as video, pictures, and text.
- **Velocity** refers to the quick arrival rate and access to data.
- **Veracity** refers to the potential for missing or erroneous data due to various reasons such as unreliable sources and equipment failures.

Some Big Data providers are listed below. Note there some excerpts from the report: *Leveraging Big Data for Managing Transport Operations (2018)*[2].

HERE (https://www.here.com/): HERE provides mobile navigation services, including 3D maps, live traffic, and public transport information. According to HERE, it collects data from more than 80,000 sources over 196 countries and one billion mobile devices are using the service.

TomTom (https://www.tomtom.com/en_us/): TomTom is a Dutch manufacturer of automotive navigation systems. The data underpinning TomTom's traffic data services come from a range of sources including TomTom navigation device GPS data, Global System for Mobile Communication (GSM) data, and Traditional road agency data. It also regularly produces regional urban traffic congestion indices for cities in Australia and New Zealand.

Google Maps (https://www.google.com/maps) including **Waze** (https://www.waze.com/): Google provides colour-coded maps of traffic speeds across roads. The data underlying the maps are sourced from third-party data sources and crowd-sourced data. Waze collects crowd-sourced traffic information from users and provides real-time traffic information back to users. Waze also allows users to report accidents, police stakeouts or any other hazards. Google purchased Waze in June 2013 when it had over 50 million users worldwide.

INRIX (https://inrix.com/): INRIX offers live traffic information, direction and driver services and flexible developer apps and tools. The Real-time Traffic Flow services are based on road sensor data and mobile devices. Based on the past traffic flows, INRIX predicts future traffic flows per day, season, and holidays. It also forecasts weather, accidents and road construction, as well as other events such as school schedules, sports games, and concerts.

StreetLight Data (https://www.streetlightdata.com/): StreetLight Data is a data analytics company, providing readily usable information about transportation and mobility. It uses multiple data sources for its products including Archival anonymized cellular data, Traffic

[2] Leveraging Big Data for Managing Transport Operations

flow data, Archival anonymized GPS data, US census and other demographic data, Geospatial yellow pages, and Open source maps.

TERALYTICS (https://www.teralytics.net/): TERALYTICS data source is also from mobile device but different from other companies which rely on Apps locational feeds. It partners with telecom operators to analyze aggregated and anonymized mobile signal data.

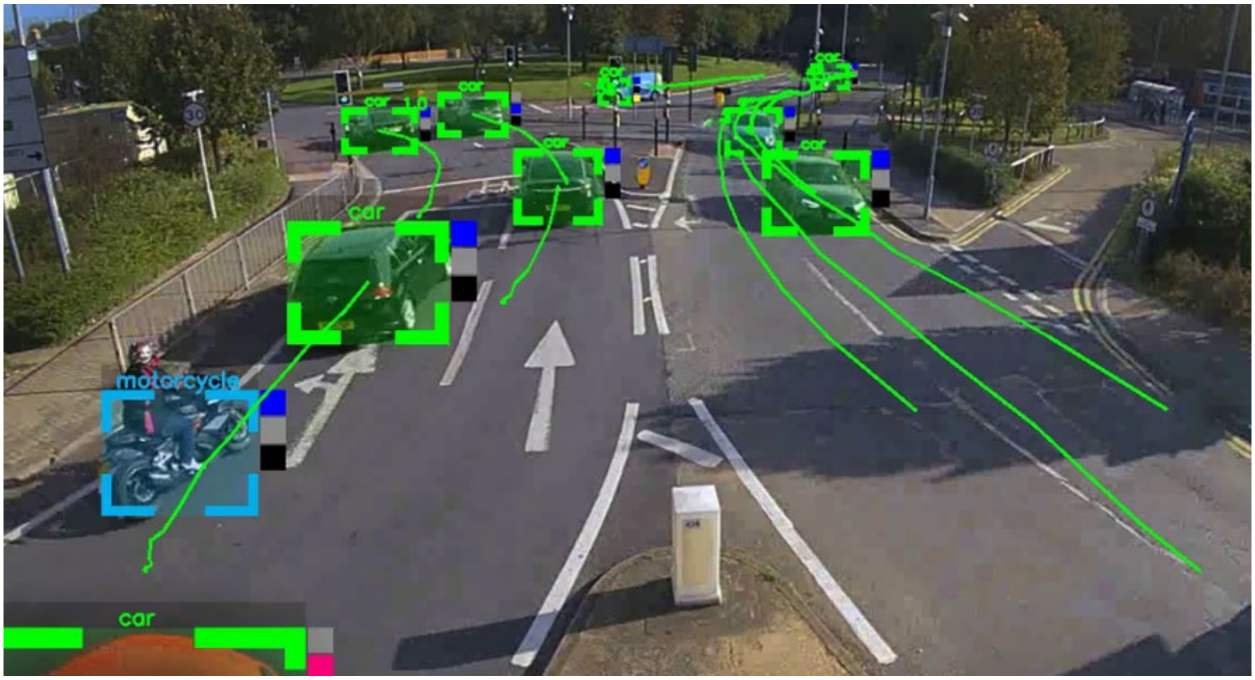
AirSage (https://www.airsage.com/): AirSage collects mobile device signal data, which then transforms into anonymized information. It contains data about when and where people are moving. It offers the following services: OD trip matrices for specified geographic areas, OD commuting trip matrices for specified geographic areas, and route-based OD matrices for specified geographic areas. It has important relationships with two major national US wireless carriers.

Ground Counts - Traditional Survey + AI

Artificial Intelligence (AI) has been one of the hottest research topics in the past 10 years. The recent development in AI (Image Processing) can distinguish various moving objects from a video clip as shown in **Figure 1**. This technology has been quickly applied in traffic data collection by many AI start-ups. Vehicle, pedestrian and cyclists can be automatically detected and counts from video clips. This emerging data collection and processing method can save time and money which was needed for traditional data collections, such as manually counting vehicles, pedestrian, and cyclists on site. Also, AI programs can significantly decrease human errors in manual counting and/or in data processing[3]. With reliable image and data processing techniques, video clips/images can be applied to collect traffic turning movement volumes, speed data, curbside activities, bike/pedestrian counts, parking supply and demand. Cameras can be mounted on roadside streetlight or traffic signal and can provide 24 hour monitoring on hotspots and provide real-time analysis. In addition to videos and images from cameras, roadside sensors such as radar, Lidar, active infrared sensors, and Bluetooth detectors can be used to collect large amount of data, which can be stored on the Cloud and can be used for real time analysis and processing using AI or other applicable algorithms.

[3] https://hellofuture.orange.com/en/ai-reduce-human-error-rate/

Figure 1: Vehicle Detection Using Artificial Intelligence



Below are some data providers that have been working on these techniques:

MIOVISION (https://miovision.com/) is a start-up founded in Waterloo, Canada. It has developed Multi-modal Detection solutions that use a single camera mounted on traffic light or streetlight pole to monitor and detect multi-modal movements at the intersection. MIOVISON has applied both AI imaging processing algorithms and cloud-based platform.

GOODVISION (https://goodvisionlive.com/#usecases): GoodVision helps traffic engineers with traffic data collection from their own cameras and provides deep traffic data analytics with its ultimate one-stop analytical platform. GOODVISION does not provide camera mounting/installation service. The user will need to upload suitable footage and use its platform for storage and data processing.

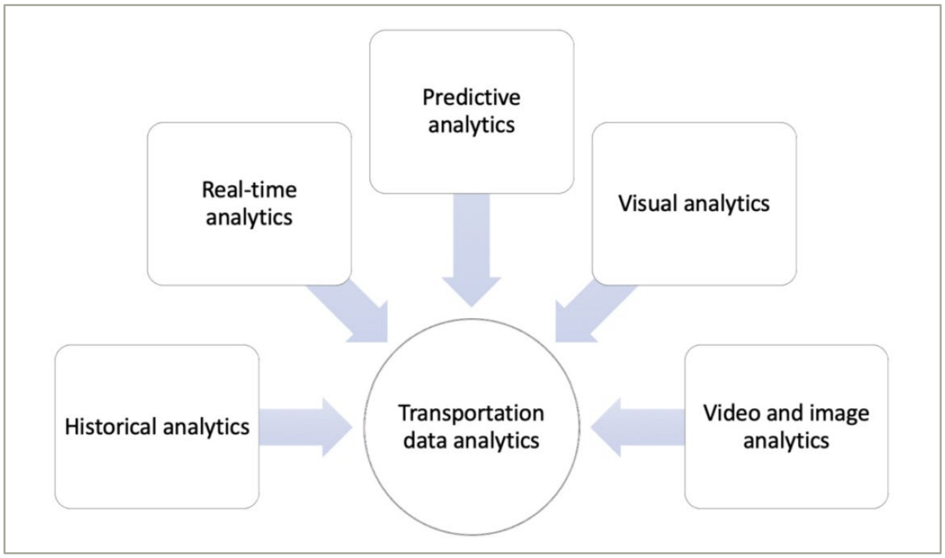
UNLEASH (https://unleashlive.com/). UNLEASH apply AI algorithms to live video stream of traffic to detect movements. Users can use their phones, Go-Pros, drones, or existing IP/CCTV cameras to connect and capture footage.

HIVE (https://thehive.ai/): HIVE is a full-stack AI company, offering solutions ranging from data labeling to model & application development.

Data Analytics

Historical data analytics can be used to understand trends and patterns in order to make long term decisions while real-time data analytics allows for quicker decisions such as coordination and real time optimization of signal timing plans. Predictive analytics allows for prediction of traffic events such as traffic congestion and collisions using methods such as near-miss identifications. Visual analytics allows for mapping of those hotspot collision areas. Finally video and image analytics have the potential of distinguishing vehicle classes while collecting traffic volumes at intersections. **Figure 2** illustrates various kinds of transportation data analytics.

Figure 2: Types of Data Analytics



Source: Neilson, Daniel, and Tjandra - Systematic Review of the Literature on Big Data in the Transportation Domain (2019)

Best Practices and Examples in North America

City of Toronto – Traffic Adaptive Pilot Project (Smart Signal) and Traffic Cameras

City of Toronto (CoT) is currently running a pilot program[4] which includes the installation of two new technologies:

- **InSync** is a technology that determines and implements signal timing plan in real-time based on video-analysis camera detection. It measures vehicle delays and queue lengths and is being tested at 10 locations on Yonge Street; and,
- **Sydney Coordinated Adaptive Traffic System (SCATS)**, which also determines and implements signal timing plan in real-time using radar detection to measure traffic flow upstream and downstream. It is being tested at 12 locations along Sheppard Avenue.

CoT analyzes the performance and effectiveness of InSync and SCATS by comparing before-and-after conditions related to traffic volumes, travel times, pedestrian delays, stops and side street delays.

CoT also utilizes smart traffic cameras (e.g. Miovision Cameras) at various intersections. These 360-degree cameras, such as the one used to monitor the King Street pilot project, allow for monitoring of all modes of transportation including pedestrian and cyclists.

[4] https://www.toronto.ca/services-payments/streets-parking-transportation/traffic-management/traffic-signals-street-signs/traffic-signals-in-toronto/types-of-traffic-signal-systems/

The Regional Municipality of York – Traffic Camera Installation

York Region has traffic cameras installed at many of its major arterial intersections which update every three minutes. The data from these cameras can be utilized to monitor traffic conditions and to analyze queuing and speeds at these intersections.

Dallas, Texas – Downtown Dallas 360 Transportation Analysis

City of Dallas undertook a detailed transportation analysis of all local roadway networks within the City Center in 2016 to evaluate the effects of potential vehicular capacity reduction[5]. The study included use of StreetLight cellphone, and GPS data to capture travel patterns throughout the study area, providing a micro-analysis of motorists' behaviors.

Maryland Department of Transportation (MDOT) – Annual State Highway Mobility Report

MDOT has been publishing annual State Highway Mobility Reports[6] beginning in 2012, which summarize the state's performance measurement efforts. The reports leverage Ritis big data and focus on mobility trends such as transport system management and multi-modalism. MDOT uses raw data to illustrate trends such as speed and travel time. Other data illustrated include bottleneck ranking along major corridors.

Texas Department of Transportation – Crash Records Information System (CRIS)

Crash Records Information System (CRIS) is the repository used by Texas DOT to compile and track crash data statewide. CRIS contains interpreted data fields such as roadway attributes and location-specific data for crashes occurring on the state highway system, crash reports submitted by law enforcement and drivers, and classification of traffic crashes in accordance with national standards.

New York City - Open Data

New York City (NYC) launched its Open Data[7] program in 2011, which provides a portal to a variety of traffic data from different sources. Below are some examples:

As part of NYC's VISION Zero action plan, an open-data map[8] illustrates crashes, including the number of fatalities and nightly crash data at each location. The map also includes various infrastructure and safety-related layers, such as locations of Leading Pedestrian Intervals Signals, Safe Streets for Seniors, and Bike Priority Areas etc., which could be correlated and compared with the collision rates.

Other large datasets such as the Origin-Destination patterns utilized from Taxi GPS allows the City to analyze trip lengths and travel times[9]. GPS devices on buses also collect over 6 million records per day, including bus activity, position and speed at 30-second intervals[10].

---

[5] http://www.downtowndallas360.com/wp-content/uploads/2017/12/8d.-Appendix-Fehr-and-Peers-Transportation-Analysis.pdf
[6] https://ritis.org/usecases/corridorperformance
[7] https://opendata.cityofnewyork.us/
[8] https://vzv.nyc/
[9] https://data.cityofnewyork.us/Transportation/2018-Yellow-Taxi-Trip-Data/t29m-gskq
[10] https://data.cityofnewyork.us/Transportation/MTA-Data/mmu8-8w8b

The NYC Open Data program collects data related to bike sharing programs, such as Citibike, and includes trip durations and travel routes[11].

City of Chicago - CityTech Collaborative

A pilot was initiated by the City Tech Collaborative which aims to collect curb activity[12] data. The project will includes private sector partners like Bosch and HERE Technologies to provide various levels of support and expertise. HERE Technologies will analyze traffic movement, congestion and other data points to identify bottlenecks and areas of friction. The data gathered will enable researchers to take a closer look at what's causing some of the curb space management problems. Phase 1 of the study will largely be a mapping exercise to digitize the curb.

Opportunities and Challenges

Records of digitized activities have created vast amounts of transportation data which can be leveraged to provide functional and efficient transportation infrastructure. As with any emerging technology, however, the real issue is not the acquisition of large amount of data, but it is the application and usage of the data that determines its value.

Large datasets provide benefits and opportunities in the present day, such as sharing real-time traffic information, which can guide users to alternate routes and/ or allow for real time signal coordination and optimization. Large datasets can help urban planners and other professionals create safer streets by using near misses and collision data. Some longer term opportunities lie in the future implementation of connected and autonomous vehicles, which will collect and emit large amounts of data and where it will be of utmost importance to have established data-exchange processes between vehicles (vehicle to vehicle- V2V), infrastructure (vehicle to infrastructure- V2I) and other users (vehicle to everything- V2X), including pedestrians and cyclists.

Currently, due to the Pandemic, traffic patterns have changed drastically, and while this may only be temporary, speculation arises about the potential long-term effects that such a crisis may have on people's travel patterns and behaviour (primarily linked to the prevalence of teleworking, which has proven very successful during these trying times). While the answer to this question is still uncertain, what becomes evident is that there will be an urgent need to observe and understand these changing trends and demands in real time. Big data and new emerging technologies can play an important role to help City planners understand these changes and to respond in a timely fashion.

Leveraging large datasets can also create challenges, in terms of storage needs and processing capacity and expertise. Other challenges relate to data acquisition timing and data quality concerns. Legal privacy barriers and lack of skilled professional to analyze the data have also been identified as challenges that need to be overcome.

---

[11] https://www.citibikenyc.com/system-data
[12] https://www.govtech.com/analytics/Data-Analytics-Will-Reshape-Chicago-Curbside-Management.html

Task 7.1 Collecting, Maintaining, and Utilizing Transportation Data

Part 2: Data Collection Questionnaire

VAUGHAN

HDR

Memo

Date:	Tuesday, April 06, 2021
Project:	Vaughan Transportation Plan
To:	Chris Tam
From:	HDR
Subject:	Data Collection Questionnaire Feedback

HDR has been retained by the City of Vaughan (the City) to undertake the City-wide Transportation Master Plan Update, also known as the Vaughan Transportation Plan (VTP). Currently, the City does not have a formal data collection and management program in place. As part of the VTP, the City would like to identify best practices for municipal data collection and management programs, and to identify new and emerging data sources to more effectively support the City's transportation planning and traffic management efforts.

HDR prepared a questionnaire and sent it to various GTA municipalities to collect input on their data collection and management programs. We have received feedback from six (6) municipalities (City of Brampton, City of Markham, City of Mississauga, Region of Peel, City of Toronto (CoT), and Region of York). Key findings from the survey for the City to consider in its own data management program are summarized below.

Data Collection – Traditional and New Sources

- Many municipalities in the GTA have recognized the importance of conducting pedestrian and cyclist counts: 5 of 6 municipalities are collecting dedicated pedestrian and cyclist counts through permanent or mobile counting stations.
- Only Brampton, York Region and Peel Region regularly collect goods movement/truck data as well as make use of MTO's data on this subject**.** The City may consider working with York Region to collect more truck data to better understand truck movements stemming from the large industrial employment base in the City.
- Big data has become an integral component of the data collection plan for many municipalities. The City of Toronto (CofT) has explored quite a few Big Data methods; notably, it has deployed 60 Miovision Smart Sense units for permanent multimodal counting and was mostly pleased with the results. CoT has also been pulling speed and volumes data from traditional "Watch Your Speed" signs, thus providing a good indication of trends and seasonality.

Data Management and Storage:

- GTA municipalities tend to display consistent trends regarding data management and storage. All six municipalities surveyed provide some form of open data program, although the level of detail of data freely available varies widely. All six also sell some form of transportation data to third parties with pricing based on cost recovery, and all six maintain a geocoded database of counts.
- Four of six municipalities use Traffic Engineering Software (TES) as their database and/or data management interface. Five of six currently use local servers to manage and store data, while Peel Region currently uses cloud storage. CoT will soon be moving to cloud storage as well.
- Three of six municipalities use some form of custom data management interface other than TES – Mississauga uses SQL and Tableau, while CoT uses FLOW and MOVE, and Peel Region is currently establishing an advanced data management warehouse.

Organization, Staffing and Budgeting:

- Staffing and budget levels for data management and collection varied widely across the surveyed municipalities. Upper- and single-tier municipalities tended both to have more dedicated staff for data management and collection, as well as higher budgets, with the exception of Mississauga which has a significant budget of $280,000 per year.
- See **Table 1** below for a summary of budget levels across respondents.
- All six municipalities indicated that Transportation or Traffic Operations, Services, or Engineering departments were responsible for data collection and management. In addition, 5 of 6 respondents indicated no special sources of funding for these programs – typically these funds were drawn from operating budgets or a mixture of capital and operating budgets.

Table 1: Data Collection Program Budget Level Responses

Municipality	Data Collection Budget	Data Management Budget	Supplementary Data Costs Budget
City of Brampton	$50,000 per year	N/A	Several hundred per year per streamline counter station
City of Markham	$50,000 per year (excluding staff time)		
City of Mississauga	$286,000 per year (excluding staff time) (City) Operating budget for 3 FTE staff (MiWay)		
Region of Peel	"Significant budget – varies, grant and NGO collaboration"		
City of Toronto	$300,000 per year	Operating budget for 5 FTE staff	$300-500,000 per year (capital costs of installing devices)
Region of York	$200,000 per year	$150,000 per year	$50,000 per year (maintenance and development)

Other:

- Questions regarding cybersecurity and privacy elicited diverse responses. While 5 of 6 municipalities indicated specific plans to address cybersecurity of transportation data (typically with responsibility delegated to municipal IT departments), only 1 of 6 (CoT) highlighted specific data security clauses for video data and the development of a privacy impact assessment for data collection devices.
- All six respondents are interested in some form of collaboration with the City of Vaughan and/or other GTA municipalities on data initiatives, indicating a strong potential desire for Region-wide data collaboration.

A summary of responses for each question can be found in the following section.

Questionnaire

1. Data Collection – Traditional and New Sources:

Does your agency have an on-going transportation data collection program or are you considering new sources of data? **If not please skip to Question 2**

a. What types of data are typically collected (i.e. TMC, ATR, pedestrians and cyclist counts, heavy vehicle counts)? Info on the frequency and location of these counts would be helpful if readily available.
All six respondents collected classified TMC and ATR data with varying degrees of regularity. 5/6 also collected some form of dedicated pedestrian or cyclist counts through permanent or mobile counting stations. The City of Toronto and Peel Region had more significant data collection programs, including radar speed studies, pedestrian delay counts, and the Cordon Count program.
b. Are transportation counts collected in-house or through third party contractors? What are the considerations for choosing one approach over the other?
All six respondents used third-party contractors. 4/6 solely used contractors; Peel Region and the City of Toronto had at least one in-house counting position as well.
c. Do you collect, or have you explored collecting goods movement/truck specific data (i.e. OD survey), or do you rely on other sources (i.e. MTO).
3/6 respondents regularly collect such data, both from third parties such as the MTO and in-house (Brampton, Peel, York). 2/6 do not collect such data (Markham, Mississauga), while 1/6 (Toronto) undertook a one-off freight and goods movement study, but does not collect data regularly.
d. Can you share any experiences and lessons learned in using Big Data or other new methods in data collection/processing, such as: i. Big Data sources - Streetlight OD, GPS data, Google Maps API, etc. ii. Permanent smart video infrastructure; iii. A.I. applications which collect and process multi-modal counts, provide input on near miss collisions, or inform parking or curbside management iv. Any other sources?

Respondents' experience is varied and extensive. 5/6 respondents indicated the use of some form of "Big Data", with the City of Brampton as the only respondent indicating that they did not. Upper- and single-tier municipalities tended to have more experience with such solutions. Some solutions included: • Radar stations • INRIX and Streetlight data • eCommerce surveys • Smart Freight Centre universities • Bluetooth travel-time sensors • Smart video counting units • Use of "Watch Your Speed" signs to collect volume and speed data and measure seasonal trends, including during the COVID-19 pandemic
e. Can you please share any experiences with specific emerging data providers / companies, reports on pilot projects, or plans for data collection using these new methods?
As above, respondents' experiences were varied and extensive. Specific firms used included Houston Radar, INRIX and Streetlight, Smart Freight Centre Universities, icorridor.org, TTS/COVID-19 surveys, Health and Transportation Surveys, Miovision, traffic and curbside data from Uber and Lyft, commercial vehicle data from Geotab, and Bluetooth travel-time sensing systems from Iteris and Get Go.

2. Data Management and Storage:

Does your agency actively manage and maintain a database of transportation data? I**f not please skip to Question 3.**

a. Do you have an open data program, and is transportation count data shared? Why or why not?
All six respondents have some form of open data program. Varying levels of transportation count data are publicly available – most respondents do share ATR/AADT data but do not share TMC data publicly, although the City of Toronto is moving to share all counts publicly within the next six months. York Region makes its Bluetooth travel-time data, lane/road closure data, and still images available to the public.
b. Do you sell transportation data to third parties? If so, what data formats do you typically provide transportation data in when sold to third parties? (i.e. Excel, PDF, GIS maps, etc.) What is the pricing strategy? (i.e. cost recovery, revenue source, etc.)
All six respondents sell transportation data to third parties (typically TMCs). All six provide data in PDF format, with 3/6 offering Excel as an option. All six base pricing on cost recovery (full or partial), with 2/6 tiered by recency of data.
c. Do you have a geocoded database of counts, and what is the effort to manage and maintain that? If so, who is responsible for compiling and preparing mapping – i.e. is there a dedicated GIS team?
All six respondents maintain a geocoded database. 4/6 use Traffic Engineering Software (TES); the City of Toronto uses the legacy FLOW program but is moving to MOVE to integrate collision data. Only 1/6 (Brampton) indicated the use of a dedicated GIS team for mapping.
d. Does your agency utilize dedicated infrastructure to manage and store transportation data (local server or Cloud services)?
5/6 respondents currently use local servers to manage and store data, with 1/6 (Peel Region) using cloud storage. As part of the upgrade from FLOW to MOVE, the City of Toronto will be moving to cloud storage as well.
e. Do you utilize a custom data management interface? (i.e. business intelligence tools, data warehousing).
3/6 respondents indicated the use of TES as their custom data management interface. Mississauga also makes use of SQL and Tableau, while Toronto uses the custom FLOW and MOVE services. Peel Region is currently establishing an advanced data warehouse in partnership with university partners.

3. Organization, Staffing and Budgeting:

Does your agency have dedicated staff or budget for collecting and managing transportation data? I**f not please skip to Question 4.**

a. What departments are responsible for transportation data collection and management?
All six respondents indicated that their Transportation or Traffic Operations, Services, or Engineering departments are responsible. The City of Toronto has a dedicated Transportation Data and Analytics Unit. Mississauga also has Community Services responsible for active transportation data collection outside of road right-of-way (ex. on trails).
b. What are your staffing needs for data collection and management (i.e. dedicated staff or partial FTE)?
Upper- and single-tier municipalities tended to have more dedicated staff for data collection and management. Peel and Toronto indicated multiple FTEs; Mississauga and York indicated at least one FTE with partial support from other departments. Markham's program is delegated entirely to existing FTEs without dedicated staff.
c. What is your overall budget (including staffing and contracting) for: i. collecting data ii. managing data iii. any other related supplementary costs
Operating budgets range from $50,000 (Brampton and Markham) to $400,000 (York), excluding staffing costs. Upper- and single-tier municipalities tended to have higher budgets, although Mississauga has a budget of $286,000 per year. Supplementary costs varied, including $300,000-$500,000 per year in additional capital installation costs for Toronto. Peel Region received partial funding through grants and NGO collaboration.
d. Is your data program funded through any special sources (i.e. development charges)
5/6 respondents indicated no special sources of funding – typically drawn from operating budgets or a mixture of capital and operating budgets. Peel Region indicated project-specific funding sources, which could include either operating budgets or development charges.
e. Do you have any documentation on your program which can be shared related to overall program management, annual budget, dedicated staffing, etc.
3/6 respondents were willing to share additional documentation. Brampton and Markham were willing to share data collection contract documentation, while Toronto was willing to share organizational charts.

4. Other:

a. When considering connected infrastructure – do you have any plans to address cybersecurity?
5/6 respondents indicated specific plans to address cybersecurity, with Markham being the only outlier. All five placed the responsibility for cybersecurity with their own municipality's IT department, cybersecurity team, or technology services division. York conducts periodic audits by Transportation Services, while Peel also partially uses its partnerships with universities.
b. Have privacy issues been considered in your data collection programs? Is it possible to get copies of any relevant policies or agreements?
Only 1/6 of the respondents indicated dedicated policies and agreements: the City of Toronto has data protection clauses for video data, and has begun developing a privacy impact assessment for data collection devices. Other respondents have no specific policies, or otherwise simply avoid collecting personal information or remove it; York notes that its programs are in full accordance with the *Privacy Act*.
c. Would you be interested in collaborating with the City of Vaughan and other GTA municipalities on any data initiatives (i.e. to reduce effort and cost through a shared platform for Big Data)?
All six respondents are interested in some form of collaboration. The City of Toronto's response was contingent upon scope and goals.

Task 7.1 Collecting, Maintaining, and Utilizing Transportation Data

Part 3: Implementation

VAUGHAN

HDR

Memo

Date:	Friday, February 25, 2022
Project:	Vaughan Transportation Plan
To:	Chris Tam
From:	HDR
Subject:	Data Collection Program – Implementation Memo

Introduction

HDR has been retained by the City of Vaughan (the City) to undertake the City-wide Transportation Master Plan Update, also known as the Vaughan Transportation Plan (VTP). Currently, the City does not have a formal data collection and management program in place. As part of the VTP, the City has requested that HDR identify best practices for municipal data collection and management programs and identify new and emerging data sources to more effectively support the City's transportation planning and traffic management efforts.

Through the results of a questionnaire sent to six (6) peer GTA municipalities collecting input on their own data collection and management programs, as well as through independent research, HDR has prepared a data collection implementation plan which addresses both the "needs" and "wants" of the City's transportation department. This plan is summarized below.

Data Collection – "Needs"

Transportation Data Management System

Without a standardized platform for managing the data it intends to collect on an ongoing basis, the usefulness of the City's data collection program will be severely limited. While the City does not presently make use of a transportation data management system, the results of the peer municipality questionnaire indicate that four (4) of the six (6) respondent municipalities use Traffic Engineering Software (TES) as their data management platform of choice; additionally, the City has already engaged in internal discussions around the utility of implementing TES as a transportation data management system. Since the two (2) respondent municipalities which do not currently use TES (Peel Region and City of Toronto) instead use a custom data management system built in partnership with university partners or by an in-house team (respectively), TES is clearly the standard among peer municipalities.

TES (https://www.tes.ca/software/) is a Canadian firm supplying modular transportation data management software in use in over 100 jurisdictions worldwide. The software suite itself features modules to manage the diverse types of data that the City intends to collect, including multi-modal turning movement counts, AADTs, collision data, and integrated GIS mapping functionality. Pricing for TES is based on a subscription model and is tiered by the subscribing municipality's population [1].

Annual Count Program

Currently, the City does not maintain a formal database of traffic counts on City-owned streets. Historically, such counts have only conducted on a project-specific basis by third-party contractors, although since 2019 the City has begun collecting turning movement counts at a select number of City-owned intersections, even without a data storage solution in place to map or store the collected data. The City does, however, require a greater and more comprehensive understanding of traffic flow on streets under its jurisdiction, and is seeking to supplement data it currently receives from York Region's traffic counts of Regional roads with an annual count program of its own.

In order to meet the City's request that all-modes turning movement counts be conducted at every collector/collector and collector/arterial intersection at least once on a frequent basis, HDR recommends that the City conduct its turning movement count program on a bi-annual basis, with approximately half of the intersections in question being surveyed each year. Using the City's existing ward boundaries as divisions, turning movement counts in each year of the bi-annual count program will be conducted as follows:

- Year 1 – Wards 1,3,5 – approx. 135 intersections (130 collector/collector and 5 collector/arterial)
- Year 2 – Ward 2,4 – approx. 137 intersections (132 collector/collector and 5 collector/arterial)

A map of the count locations identified for the bi-annual count program is given in **Figure 1**, with the City's five (5) wards colour-coded, as well as with collector/collector and collector/arterial intersections marked as green and blue dots, respectively.

[1] https://www.tes.ca/wp-content/uploads/2020/04/TES-Monthly-Pricing.pdf

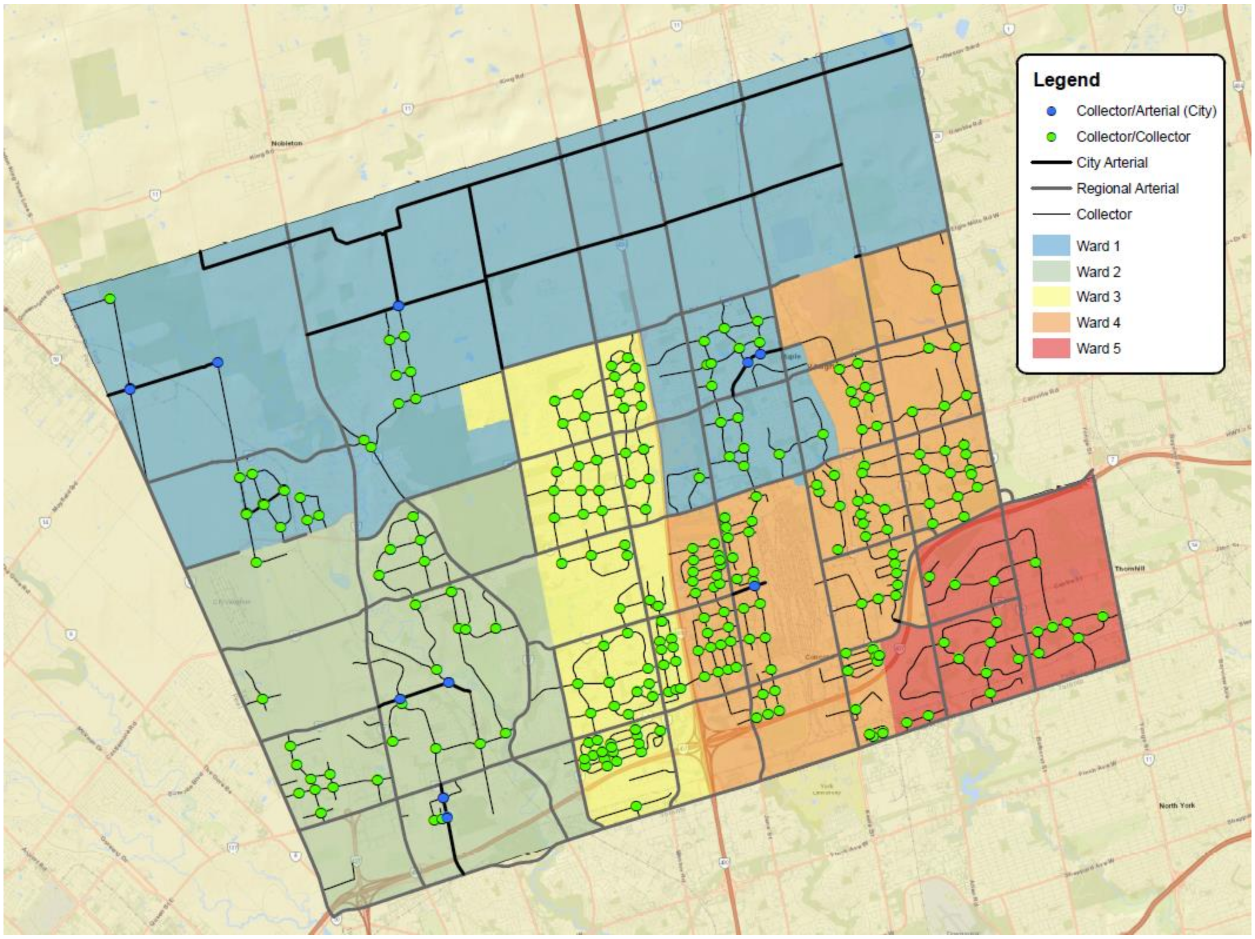


Figure 1: Proposed Bi-Annual Count Program Locations

It is important to note that most arterial roads in the City of Vaughan are under the jurisdiction of York Region. Only approximately ten (10) intersections occur wholly between collectors and arterials under the jurisdiction of the City of Vaughan, as is reflected in the above totals. Collaboration with York Region's turning movement count program is therefore recommended to avoid duplication of efforts where the Region may have already conducted counts of its own.

Vehicle AADT data may also be collected for each collector within each of the City's blocks according to the proposed wards-based bi-annual count program. It is recommended that AADT counts of collectors be conducted at minimum on each segment bounded by arterials or by the boundaries of the City's block system; whichever results in more closely-spaced count locations. Using Vellore Woods Boulevard as an example of a collector, this methodology would require one AADT count be taken on Vellore Woods between Rutherford Road and Major MacKenzie Drive.

Cycling AADT data is to be counted for all separated cycling facilities every year, and must therefore proceed on a City-wide annual basis.

Annual Collision Data Collection

The City does not currently maintain a database of data on collisions which take place on City streets, instead receiving individual PDF reports of collision incidents as supplied by York Region Police. Other municipalities, including York Region, do maintain such a database, and there exists potential for collaboration with the upper-tier municipality for the collection of collision data on City-owned streets within Vaughan's borders.

HDR recommends that the City collect collision data on City-owned streets from York Region Police on an ongoing basis and establish consistent procedures for cleaning and storing this data in a centralized database (for instance, using TES) or a geocoded map as opposed to static and insufficient individual PDF reports. The City of Toronto, as an example, publishes historic and ongoing collision data in a publicly available geocoded map created with ArcGIS Online and the Leaflet Javascript mapping library as a means of tracking progress towards its Vision Zero road safety goals [2].

Arterial Road Travel Time Data Collection

York Region currently operates a network of Bluetooth travel time sensors deployed throughout the Region on arterial roads under its jurisdiction, including 120 sensors located on arterial roads within the City of Vaughan (a map of these sensors' locations is shown below in **Figure 2**). The Region provides live travel time data from its sensors as an XML feed on its Open Data portal [3], with data being refreshed every five (5) minutes.

As the City wishes to collect and summarize travel time data on Regional arterial roads from the Region's Open Data portal, and since no historic database of travel times is accessible via the portal, HDR recommends that the City develop a script to continuously scrape live travel times from the Region's XML feed, whether that be at the feed's refresh rate of every five minutes or at a longer interval. Running the script continuously and storing the data it retrieves from the Open Data portal will provide the City with an ongoing source of arterial travel time data with which to construct its own annual database.

[2] https://www.toronto.ca/services-payments/streets-parking-transportation/road-safety/vision-zero/safety-measures-and-mapping/
[3] https://ww6.yorkmaps.ca/TravelTime/iteris_traveltimes_out.xml

Figure 2: Regional Bluetooth Travel-Time Sensors in Vaughan

Data Collection – "Wants"

Permanent Multi-Modal Count Stations

MIOVISION (https://miovision.com/) is a start-up founded in Waterloo, Ontario. It has developed multi-modal detection solutions ("MIOVISION TrafficLink") which use a single camera mounted on a traffic signal or streetlight pole to monitor and detect multi-modal movements at an intersection. MIOVISION's technology applies an AI imaging processing algorithm and features a cloud-based platform. This 'video analytics' approach means that all movements can be captured and multi-modal counts, as well as collision and near-miss hotspot identification, can be theoretically generated 24 hours a day, 365 days a year.

Following the successful implementation of TrafficLInk cameras by the Town of Milton, the City could purchase and install TrafficLink cameras at key intersections throughout Vaughan to not only establish permanent multi-modal AI-driven count stations, but also to assist in signal optimization projects. The advantage of the MIOVISION system over similar technologies is in the provision of a smart camera which must only be mounted to an existing traffic signal or pole, greatly assisting in the simplicity of setup and use.

Annual Origin-Destination and Trip Generation Data Collection

Existing methods of Origin-Destination and Trip Generation data collection surveys with known limitations can be supplemented by leveraging "Big Data". The proliferation of cellphone usage which can anonymously track location has resulted in the ability to supplement traditional data collection methods by emerging "Big Data" sources which can be utilized efficiently with new collection techniques and data processing methods.

There exist numerous "Big Data" providers which offer opportunities for commercial partnerships on Origin-Destination data. Some providers are listed below; these suggestions are those which are generally most flexible for usage by municipal governments, as opposed to those oriented towards business owners:

- **Google Maps** (https://www.google.com/maps) including **Waze** (https://www.waze.com/): Also features robust platform of APIs for scripting and data collection; HDR has experience in accessing and leveraging these APIs.
- **StreetLight Data** (https://www.streetlightdata.com/)
- **AirSage** (https://www.airsage.com/)

Data aggregators such as StreetLight typically offer packages for either a set number of traffic zones, or an unlimited regional subscription. Transit data development from these sources lags other modes, but is catching up. The 'big data' approach can be very useful, since it can provide valuable data for all time periods, virtually in real-time compared to traditional methods, and covering various vehicle types, freight movement, origin-destination data, and trip generation rates. However, given that it provides a sample, validation and calibration are required, and it is best used in addition to permanent infrastructure in key locations. Depending on specific needs, a regional subscription may provide more value, though its cost needs to be considered.

Annual Truck Origin-Destination and Trip Data Collection

York Region presently conducts its own truck-specific movement data collection program, as indicated in its responses to the peer municipality questionnaire. Given that Regional roads form the backbones of the truck/heavy vehicle movement network within Vaughan, and since the Region has indicated that it already has a form of collaboration with the MTO in its questionnaire responses, collaboration with or use of Regional truck movement data is strongly encouraged.

Deployment of Bluetooth Travel Time Sensors on City Collector Roads

Similarly, York Region has already deployed a network of Bluetooth travel time sensors on Regional arterial roads, including those in the City of Vaughan, as described in the "Needs" section. The Region has chosen Iteris (https://www.iteris.com/products/travel-time) as its vendor for Bluetooth sensors, as can be discerned from examining the XML feed of travel times published to its Open Data portal [4]. Consulting York Region to discuss benefits and/or

[4] https://ww6.yorkmaps.ca/TravelTime/iteris_traveltimes_out.xml

limitations of the specific models purchased and installed, as well as lessons learned for establishing a travel time collection program or for the possibility of collaboration, is also recommended. Bluetooth sensors can also be supplemented by third-party 'big data' sources as long as calibration and validation can be undertaken.

Partnerships with Agencies, Municipalities, and Academics

Advancements in technology have created new transportation data collection and management opportunities, including GPS data, location-based services data, and smart cameras on traffic signal arms and streetlight poles. These new data sources create opportunities as well as challenges for the City to harness the data into meaningful, actionable information. In addition, with a well-connected regional transportation network, the City is not alone in facing these opportunities and challenges to manage the constantly changing travel demand and patterns. The University of Toronto School of Cities recently announced the creation of the Urban Data Centre, which will collect municipal information such as open city data and sensor in a central place to make data more accessible. The City should seek partnerships with other government agencies, municipalities, and the academic world to work on the data collection challenges.

Cost Estimates

The results of the peer municipality questionnaire allow for the estimation of $50,000 per year as a cost estimate for an ongoing data collection program for the City, as this is the approximate cost indicated by the Cities of Brampton and Markham – both of which are of similar size and geographic extent to Vaughan. Capital and one-time costs, such as to purchase smart cameras or travel time sensors, are not included in this estimate – pricing may vary and is generally only available upon contacting individual vendors.

The Regions of York and Peel, as well as the Cities of Toronto and Mississauga, all highlighted in the questionnaire that they employ full-time equivalent staff dedicated to data collection. The City has indicated that funding dedicated staff would be a stretch for its purposes, and that developing a program which can be maintained with minimal full-time staff involvement would be ideal.

Peel Region indicated that the cost burden of its data collection program is reduced through grants and collaboration with non-government organizations (NGOs). The City may benefit from participation in similar partnerships or grant applications and should inquire further with Peel Region for details on these programs.