Today's discussion

1	Achieving Outcomes	Using Flow and Link to achieve the City's goals.
2	Analytics Dashboard	New data sources lead to new possibilities
3	Data Management	Best practices on architecture, privacy and more
4	Scaling Impact	Business Model, Implementation and Project Management

Flow Applications

Kiosk

Flow Analytics

Flow Data Platform

Data Platform

The Smart City Challenge promises a diverse new set of services and capabilities for cities in the future.

A real-time public/private data exchange will be needed to support these new mission-critical services. The Flow data platform provides infrastructure for this exchange with a focus on security, high availability, and open data.
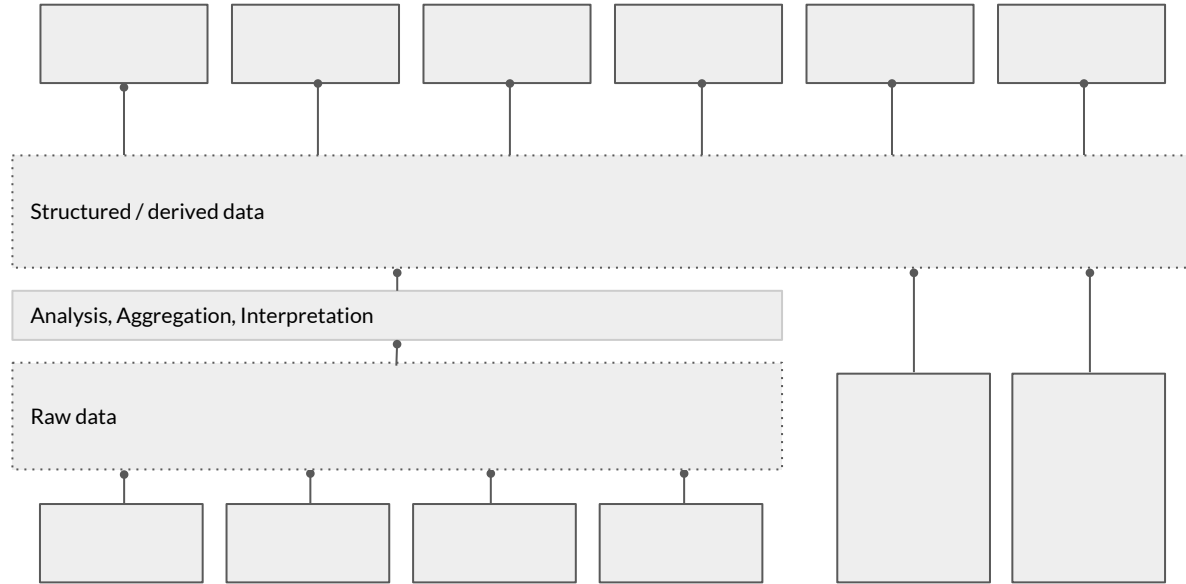
SIDE WALK LABS

Data Platform

Apps and services
City, researchers,
third-party

Data platform

Common interface for
data providers and
services.

Consistent data
schemas for common
city data used by all
participants.

Structured / derived data

Analysis, Aggregation, Interpretation

Raw data

Data sources
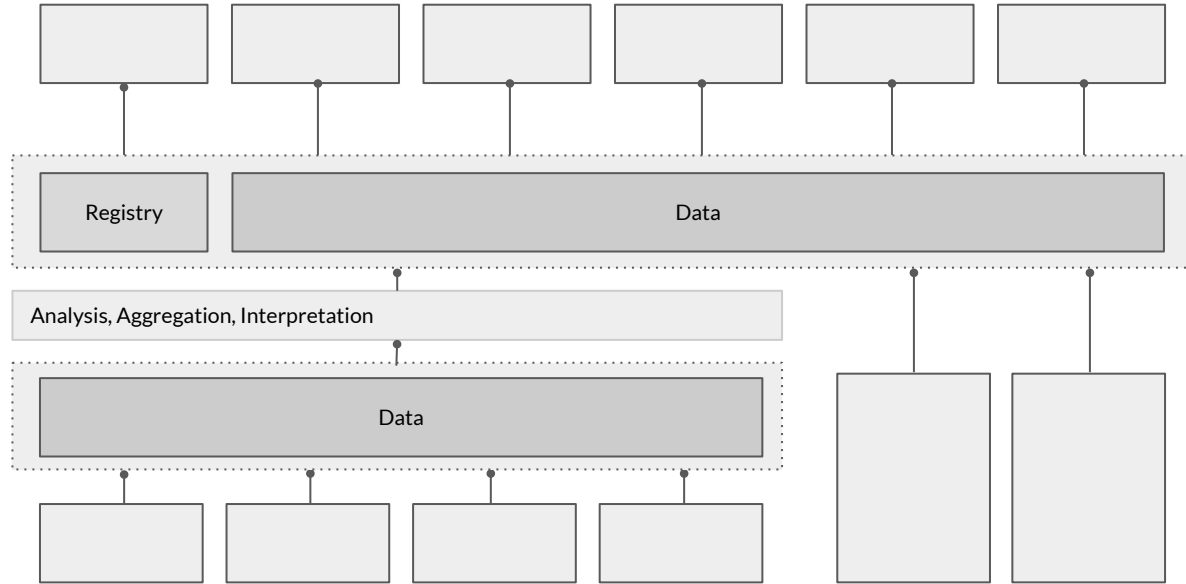Public / private
collaboration for data

SIDE WALK LABS

Registry and Data

Apps and services
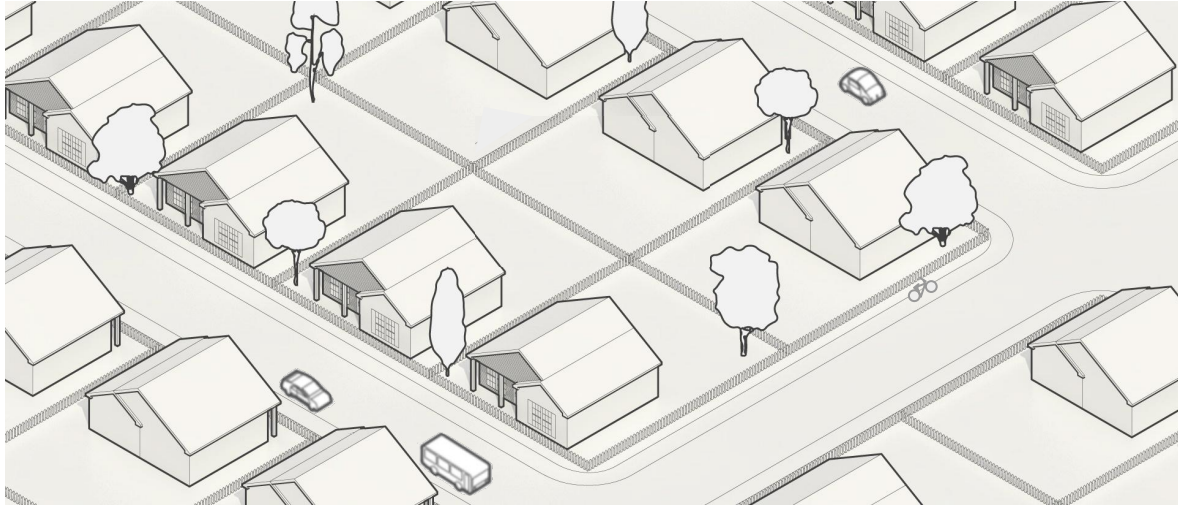City, researchers, third-party

Registry provides single manifest of common entities (parking spaces, streets, etc.)

Data provides structured and raw datastore.

Data sources
Public / private collaboration for data

Registry	Data

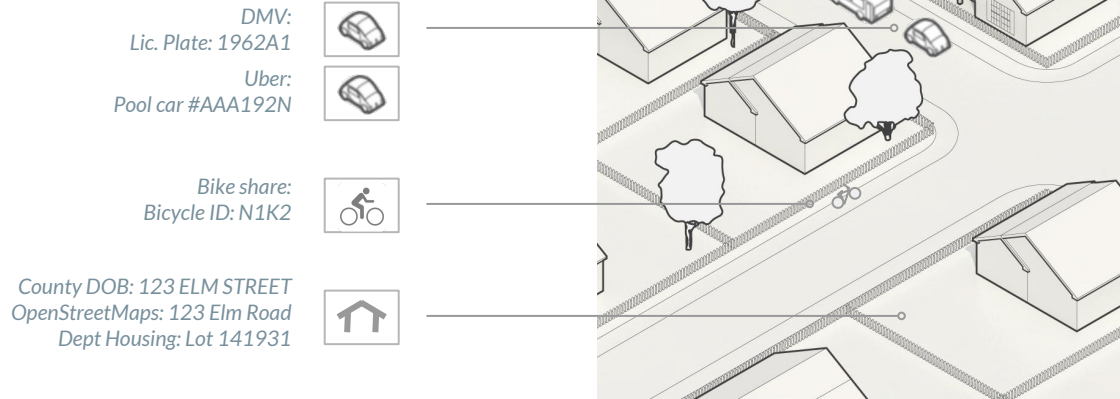Analysis, Aggregation, Interpretation

Data

Registry

In building apps and services for smart cities, many resources need to be referenced by different systems - resources like roads, parking spaces, city vehicles, shared bicycles, and intersections.

Registry

DMV:
Lic. Plate: 1962A1

Uber:
Pool car #AAA192N

Bike share:
Bicycle ID: N1K2

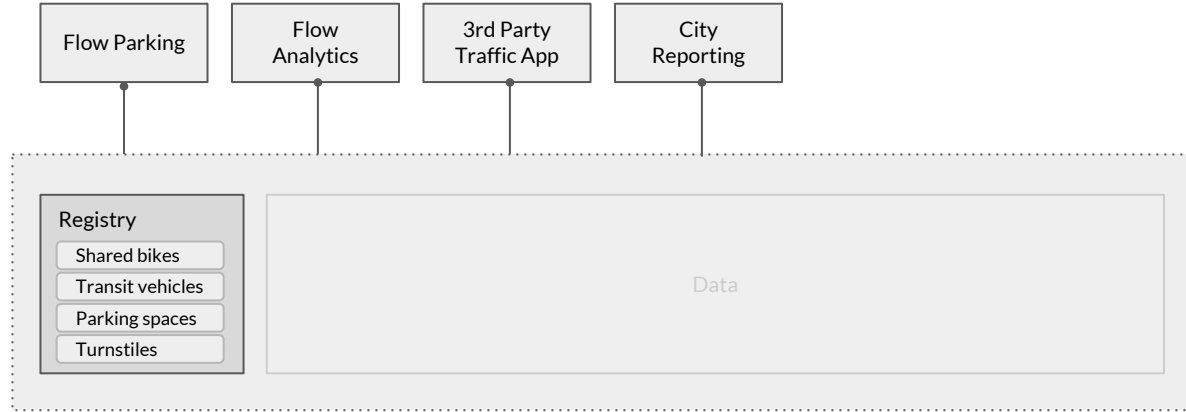County DOB: 123 ELM STREET
OpenStreetMaps: 123 Elm Road
Dept Housing: Lot 141931

The problem today is just that there's no stable, consistent way to represent these - so cities end up with a lot more of them than they actually have, or silos of information that can't be used together.
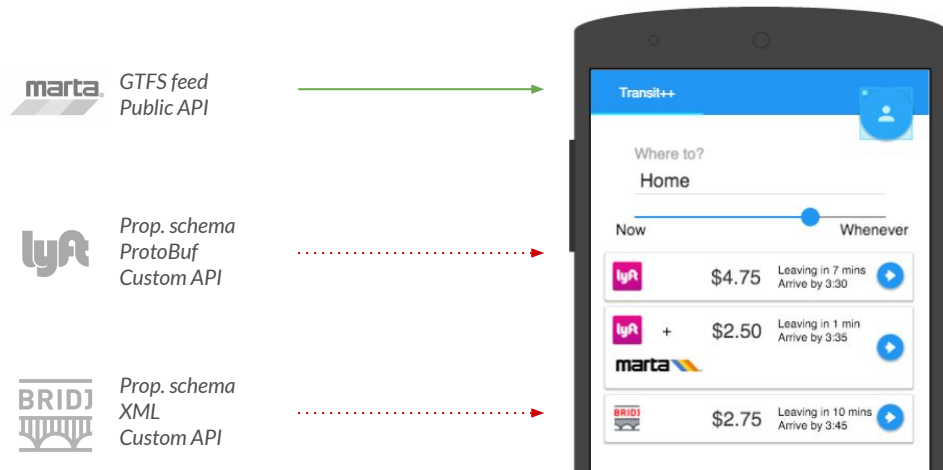
SIDE|WALK|LABS

Registry

For example, Flow Transit can then look for traffic congestion on a street and road closures for it, even if the source of the former was a private company and the latter was the DOT.

| Flow Parking | Flow Analytics | 3rd Party Traffic App | City Reporting |

Registry
- Shared bikes
- Transit vehicles
- Parking spaces
- Turnstiles

Data

The Registry removes the mess and provides a single reference for resources within the city - making it easier to keep track of everything from streets to bus stops across different sources of data.

Data

marta® — *GTFS feed*
Public API

lyft — *Prop. schema*
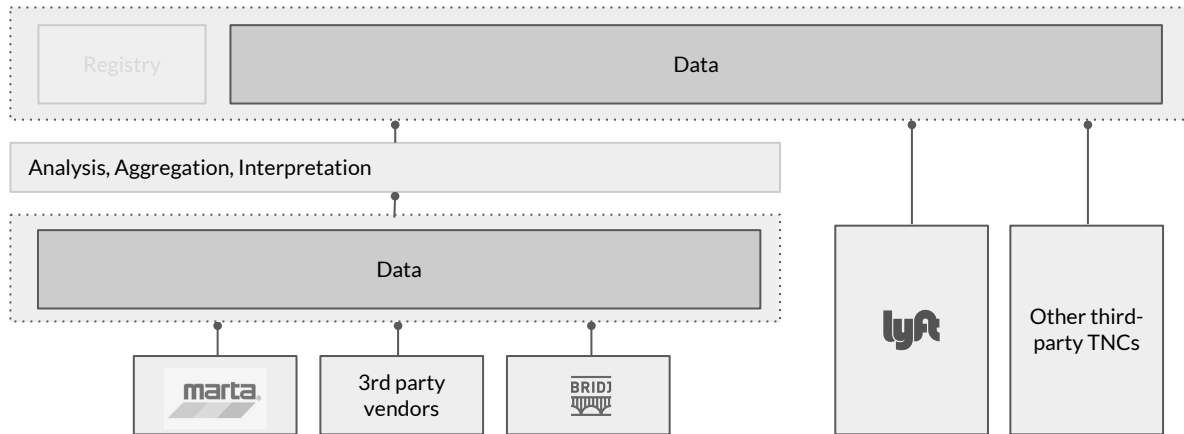*ProtoBuf*
Custom API

BRIDJ — *Prop. schema*
XML
Custom API

Being able to reference city resources consistently is useful, but if the data which references them is stored in a variety of different places, accessible via different APIs, and is exposed in dozens of different formats, it's not usable.

Data

The Data portion is built to store geographic and time-series data using consistent data schemas and formats, while providing unified interfaces for apps and services to make use of that data.

SIDE WALK LABS

Overall data platform

Apps and services
City, researchers, third-party

Flow Analytics	Flow Parking	Transit Pass	Air Quality Reports	Researchers	Waze / Maps

Structured data
Derived from sensor or raw data

Registry	Transit Occupancy	People	Parking Inventory	Road Maintenance	Road Occupancy

Analysis, Aggregation, Interpretation

Raw data
Sensor or inventory system data

Sensor Data	Sensor Data	Meter usage/ pricing	Park pricing/ availability

Data sources
Public / private collaboration for data

Link	MobileEye	Meter data	3rd party parking co.

Third-party traffic / transit apps

Third-party TNCs

Legend:
- City data providers / services
- Third-party data providers / services
- Sidewalk data providers / services
- Sidewalk data platform

- Standard REST + Pub/Sub interfaces
- Open data formats / schema

- All data encrypted at rest and in-flight
- Granular access control list (ACL)

SIDE WALK LABS

Extensibility


```
Gray Line Bus

Bus ID:
Route:
Driver:
Current Location:
A/Q sensor ID:
CO2 level:
PM2.5 level:
...
```

Cities and third-party services can always store new data types - the data platform isn't locked down to the schemas that Sidewalk understands.

This makes it easy to innovate with new products and services, or contract with new data providers, without needing to rebuild the data platform from scratch.
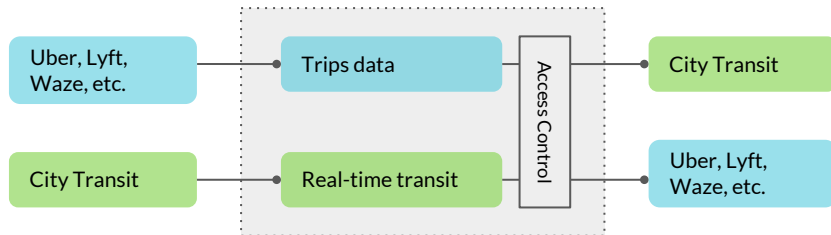
Data types can be extended to fit your needs, too - so if you need to track more information about buses than other cities, you can still keep it in the same storage with the same access controls as the rest of your transit data.

New data providers or more sophisticated use can extend the data model as needed, without breaking interoperability.

SIDE|WALK|LABS

Example data exchange

| Uber, Lyft, Waze, etc. | → | Trips data | Access Control → | City Transit |
| City Transit | → | Real-time transit | → | Uber, Lyft, Waze, etc. |

Uber has trip data which can help the city - the city would like to verify whether underserved communities have sufficient access to Uber and other TNCs.

It would also like to use Uber trip data to judge demand for new public transit lines, and determine net impact of transit outages.

The city has demand response data that can help Uber - Uber would like to know, in real-time, where public transit services are over-capacity and wait times are growing.

This kind of data will allow Uber to route more cars to areas where demand is likely to spike.

The city and Uber agree to data-sharing provisions, using the data platform's access controls - this enables both to grant access to one another for data they need, without giving up data ownership or control.

This model can be extended to other public and private actors.

SIDE|WALK|LABS

Highlights: Technology Stack

Cloud hosted, with publicly documented REST APIs for data providers and services - Standardized interfaces for all app and service integrations. HTTPS required for all requests. **Client libraries provided** to make development easier for both data providers and services.

Open data types and schemas - Data must conform to well-documented, open data schemas, and is never locked into proprietary data types.

Standard data formats -Multiple formats supported depending on the data type and whether it's being uploaded or requested as output.

Publish-subscribe support for realtime data - Several push notification mechanisms for apps and services which need to know about changes in realtime, to enable scenarios like immediate demand response.
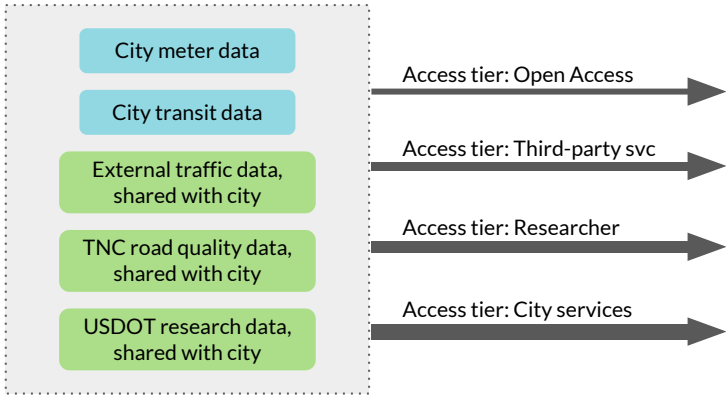
Highlights: Access controls

Data belongs to cities and their data providers - Public and private data providers always maintain control over who can access their data. Fine-grained access controls are provided to make data sharing flexible, but secure.

The city decides who, if anyone, has access to city data and which data providers to trust - it can extend access to city systems, external researchers, or products and services built on top of city data.

City meter data

City transit data

External traffic data, shared with city

TNC road quality data, shared with city

USDOT research data, shared with city

Access tier: Open Access

Access tier: Third-party svc

Access tier: Researcher

Access tier: City services

City data access controls

SIDE WALK LABS

Highlights: Security and Privacy

All data is encrypted in the data platform - Data is encrypted when at rest on-disk, it's encrypted when in-flight between data platform components or apps and services. Data is never transferred in the clear.

Extra safeguards for PII and SPII - Cities can grant access to data likely to include PII, but must always explicitly acknowledge the risks of this access grant before succeeding. Extra safeguards are taken by the platform to ensure PII data does not make its way into long-term logging or system information.

Data access auditability - Any time data is accessed in the data platform, the accessing application or service must present credentials, pass authentication and authorization checks, and generate an audit log entry before being provided with access.

Privacy-sensitive analytics - Flow Analytics and other services from Sidewalk never expose individual or user-identifiable data about end-users to operators.

Highlights: Availability and Test

High uptime and reliability - The service will provide 99.9% uptime during the course of initial pilot program, after which target uptime will be raised further.
Change management - Advanced warning with timeframes for upcoming API changes, new features, and deprecations.

Status dashboard - The data platform and other Flow and Link services can be monitored via a status dashboard, which will indicate when there are known issues. This can help to greatly reduce the investigation time needed for issues relating to city services.
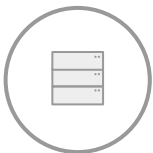
Staging environment for integration testing - The data platform will provide a public staging environment for running integration tests for data providers and services against. This can help services and providers to identify integration problems before they reach production and real users.

Discussion Topic: Data Principles

Strong data management principles are critical for building user trust.

The Sidewalk data platform will be built to support a common set of principles, built collaboratively with cities, for the collection and use of data. We need your input to get these right.

- **Transparency.** Maintain legible, transparent products and services for understanding data collection, storage and use.

- **Individual Control.** Provide individuals with reasonable means of controlling the processing of their personal data.

- **Respect for Context.** Use personal data in a manner consistent with the context in which it was collected and agreed upon at time of collection and provide a reasonable way for individuals to opt out or withdraw their consent.

- **Focused Collection and Responsible Use.** Collect, retain, and use only personal data that is reasonable in light of context and delete, destroy, or de-identify that data within a reasonable time after fulfilling the purposes for which it was collected.

- **Security.** Keep personal data secure from loss, compromise, alteration, unauthorized use, and disclosure. All data handled by the platform will be encrypted when in transit and at rest.

- **Access and Accuracy.** Provide individuals with reasonable access to and ability to correct their personal data.

- **Accountability.** Conduct privacy assessments, adopt privacy-by-design processes, require developers and other recipients of personal data to adhere to these principles, and submit to appropriate regulatory oversight.

Today's discussion

1	Achieving Outcomes	Using Flow and Link to achieve the City's goals.
2	Analytics Dashboard	New data sources lead to new possibilities
3	Data Management	Best practices on architecture, privacy and more
4	Scaling Impact	Business Model, Implementation and Project Management