

**Moving Towards Regional Public Transportation
Systems through Integration and Data Quality
Standards:
Going Beyond the Transit Data Dictionary**

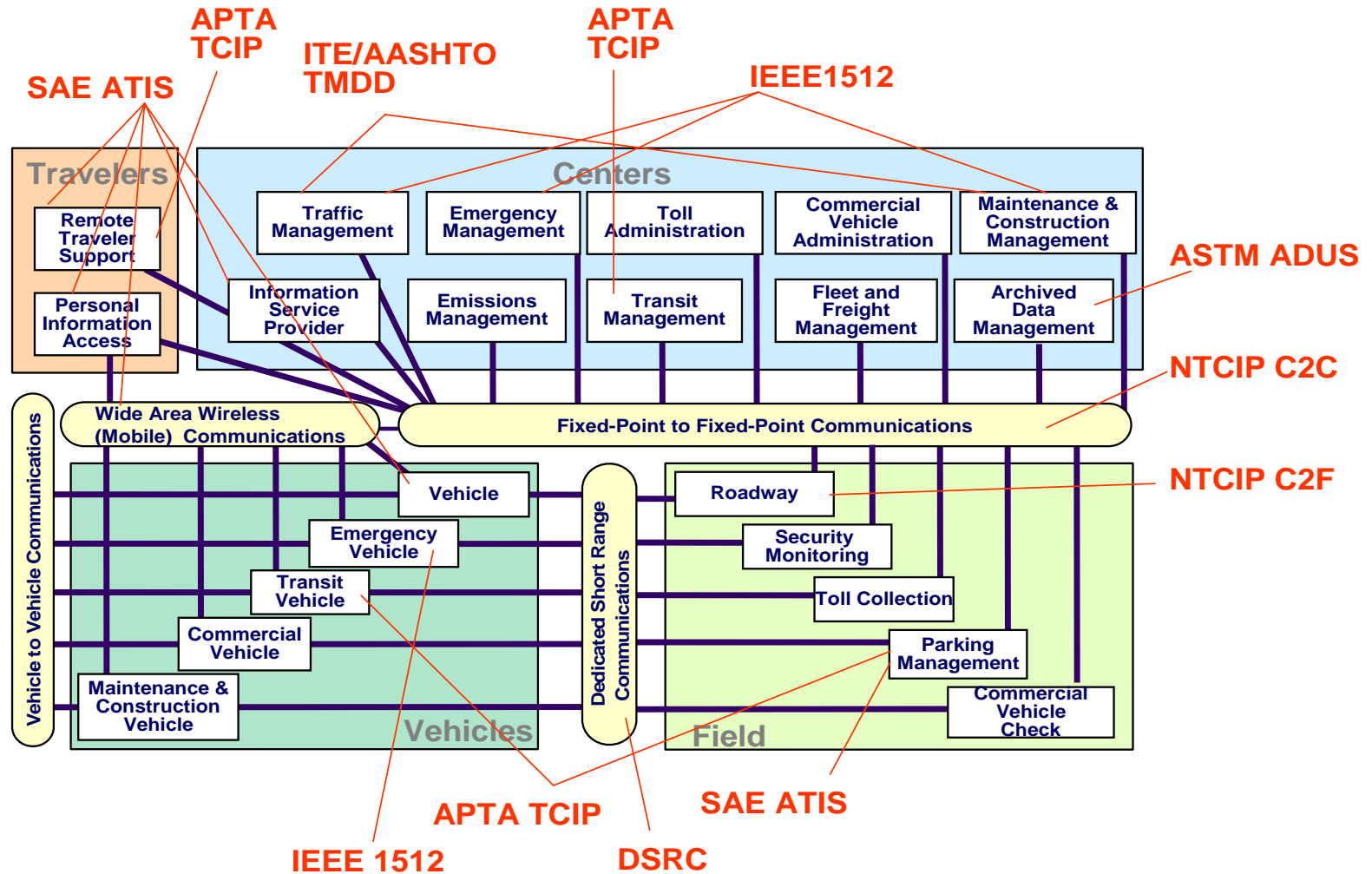
By
Paula Okunieff
Consensus Systems
Technologies Corp.



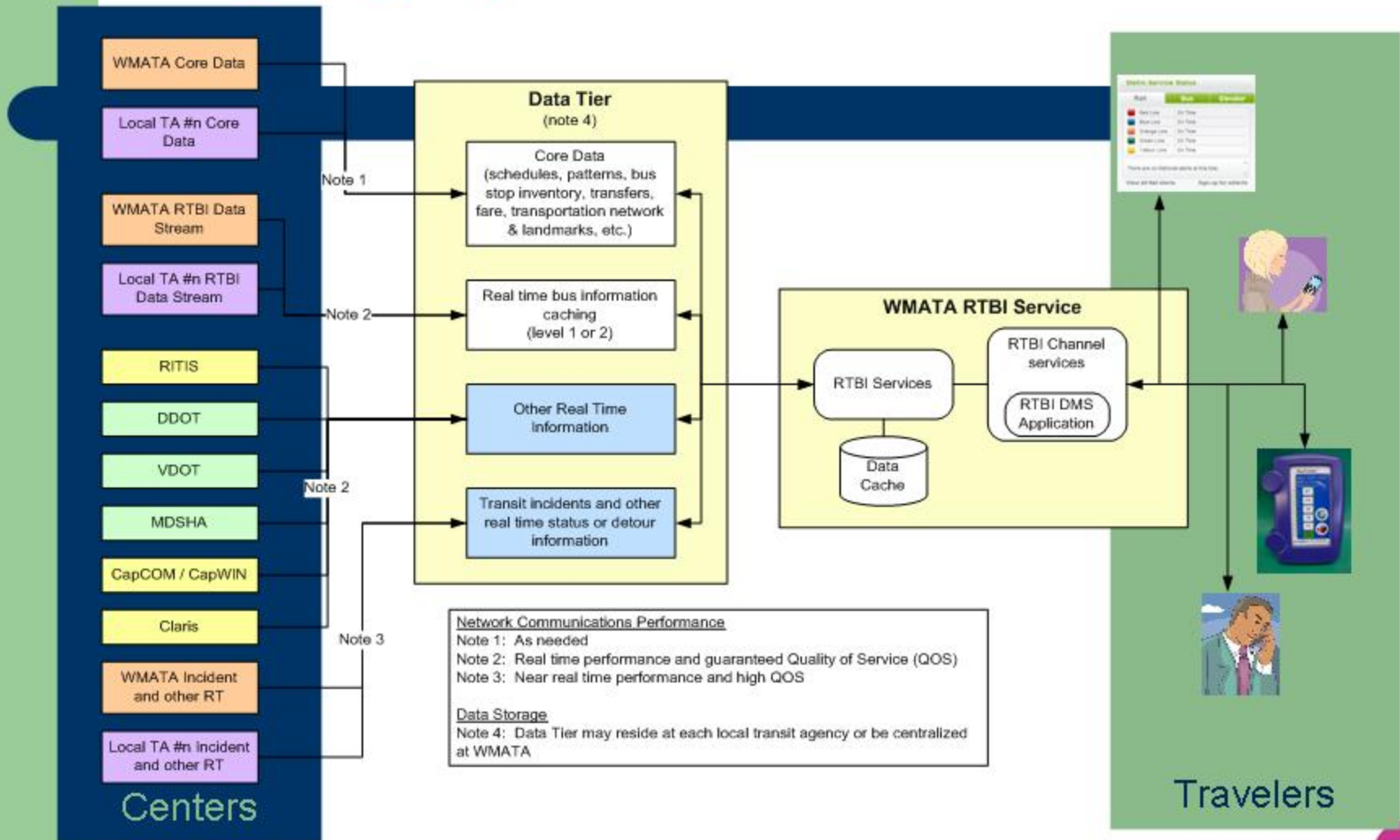
Context for Standards

- Relationship between architecture & standards
- Standards work together
- Interoperable and stackable
 - No one standard can meet all requirements for a system
 - Want standards to be specified in layers (mix and match)

ITS Standards and Architecture



Mapping Systems to Architecture



Web Services Standards Stack

- Adopt a flexible standards stack based on Industry Standards
- Define a Web Services Standard “Profile”



IT, e-Gov and ITS Standards Profile
[Typology adapted from Forrester presentation]

Industry and Business Semantics	TCIP*, Spatial Content Framework Standards Part 7d
Service Presentation / Event Delivery	Portal Standards
Service Orchestration	WS Choreography
Security	WS Security, SSL
Service Discovery	UDDI
Service Description	WSDL, WS-Policy
Message Processing	SOAP, NTCIP 2306 (profile of SOAP 1.1)
Data Extraction	XQuery, SQL/XML
Data Semantics	XML Schema
Data Syntax	XML, XPATH
Transformations, Encoding & Rendering	WAP, WML, Voice XML, XSLT
Transport	HTTP



Partitioning the Standards “Space”

Viewpoint

Levels of Abstraction

	<i>Service Invocation</i>	<i>Information Transfer</i>
<i>Implementation specifications:</i> Recipes, Practice 	<i>Interface:</i> •Web Service specification (SOAP/WSDL) •Datex	<i>Encoding:</i> •Data Structures •XML Schema •ASN.1
<i>Abstract model:</i> Guidance for Design 	<i>Behavior:</i> Exchange Patterns •Query •Publish/Subscribe •Broadcast	<i>Content:</i> •Abstract Reference Model •Message Content, Data Dictionary

US ITS Standards in the Standards Space

	<i>Service Invocation</i>	<i>Information Transfer</i>
<i>Implementation specifications</i> 	<i>Interface:</i> C2C NTCIP 2306 (SOAP/WSDL) TCIP Volume 4	<i>Encoding:</i> TCIP Volume 3 (XML Schema) All other ITS standards GML
<i>Abstract model</i> 	<i>Behavior:</i> TCIP Volume 1, Chpt 7 C2C (exchange patterns)	<i>Content:</i> TCIP Volumes 1* & 2 All other ITS standards IEEE Data Dictionary/ Msg Template

TCIP Dialog Functional Areas

(Business to Business and some Business to Consumer)

ITS Information Standards
(ATIS, IM, LRMS, NTCIP)

Control Center (CC)

Asset & Inventory (CPT)

Fare Collection (FC)

Incident Management (IM)

Passenger Information (PI)

Scheduling (SCH)



Geo Spatial/Location (SP)

Transit Signal Priority (TSP)

Dialog Patterns: Broadcast, Request/Response, Publish/Subscribe, etc.
Conformance: Requirement (PRL)/Implementation (PICS) Specifications Formats

Encoding
Alternatives

Where are the European Standards?

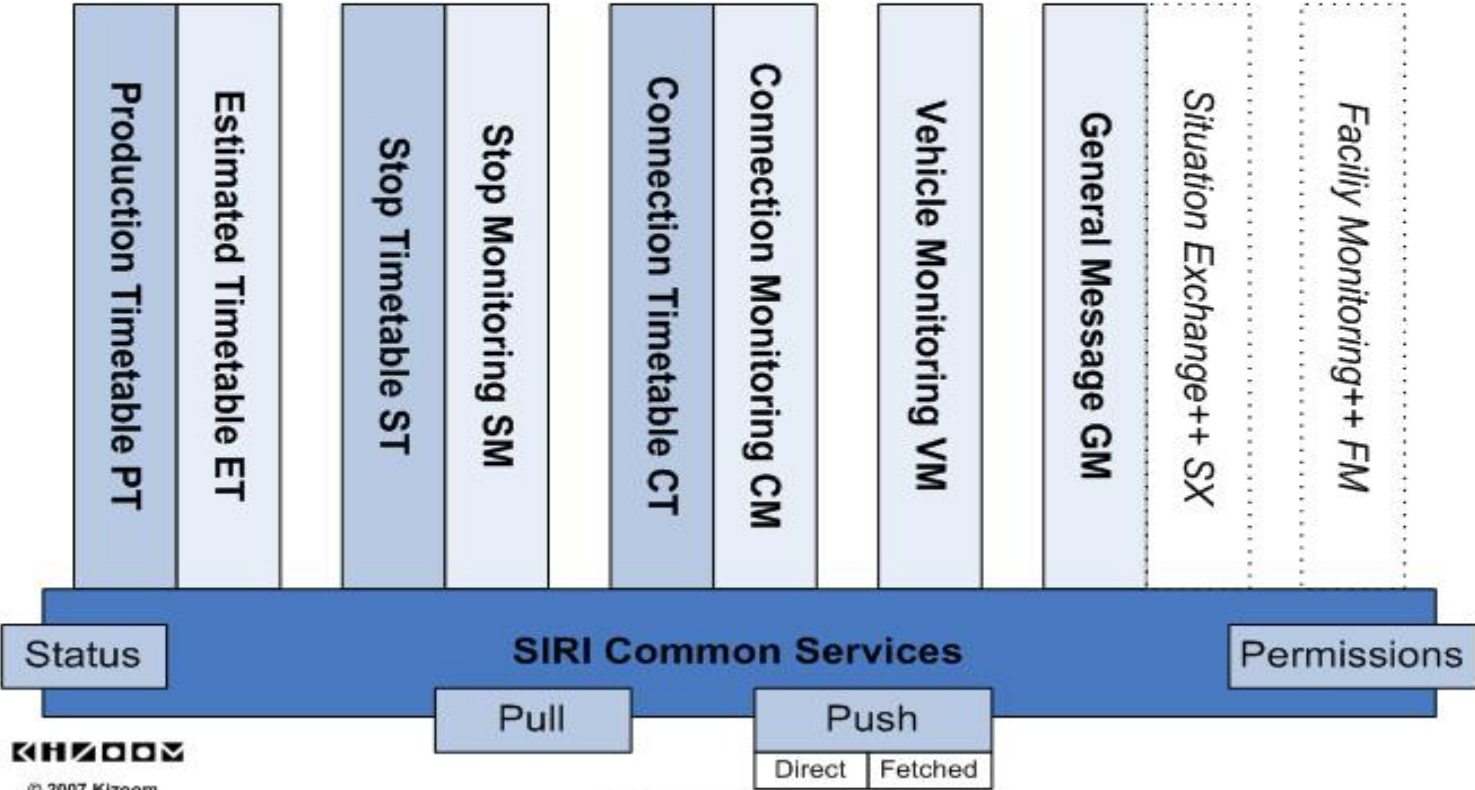
	<i>Service Invocation</i>	<i>Information Transfer</i>
<i>Implementation specifications</i> 	<i>Interface:</i> SIRI (SOAP/WSDL)	<i>Encoding:</i> SIRI (XML Schema)
<i>Abstract model</i> 	<i>Behavior:</i> SIRI (exchange patterns)	<i>Content:</i> TransModel (TPEG, Datex2)

SIRI Functional Services

TPEG2: Situation Model
PT & Road *Situations*

Datex2: Situation Model
Road *Situations*

Transmodel: PT model + Stop Place model (IFOPT)
Stop Points, Vehicle Journeys, Lines, Journey Patterns, Vehicles etc



Web Service: Request/Response, Publish / Subscribe
Topic Filters, Policies, Heartbeat

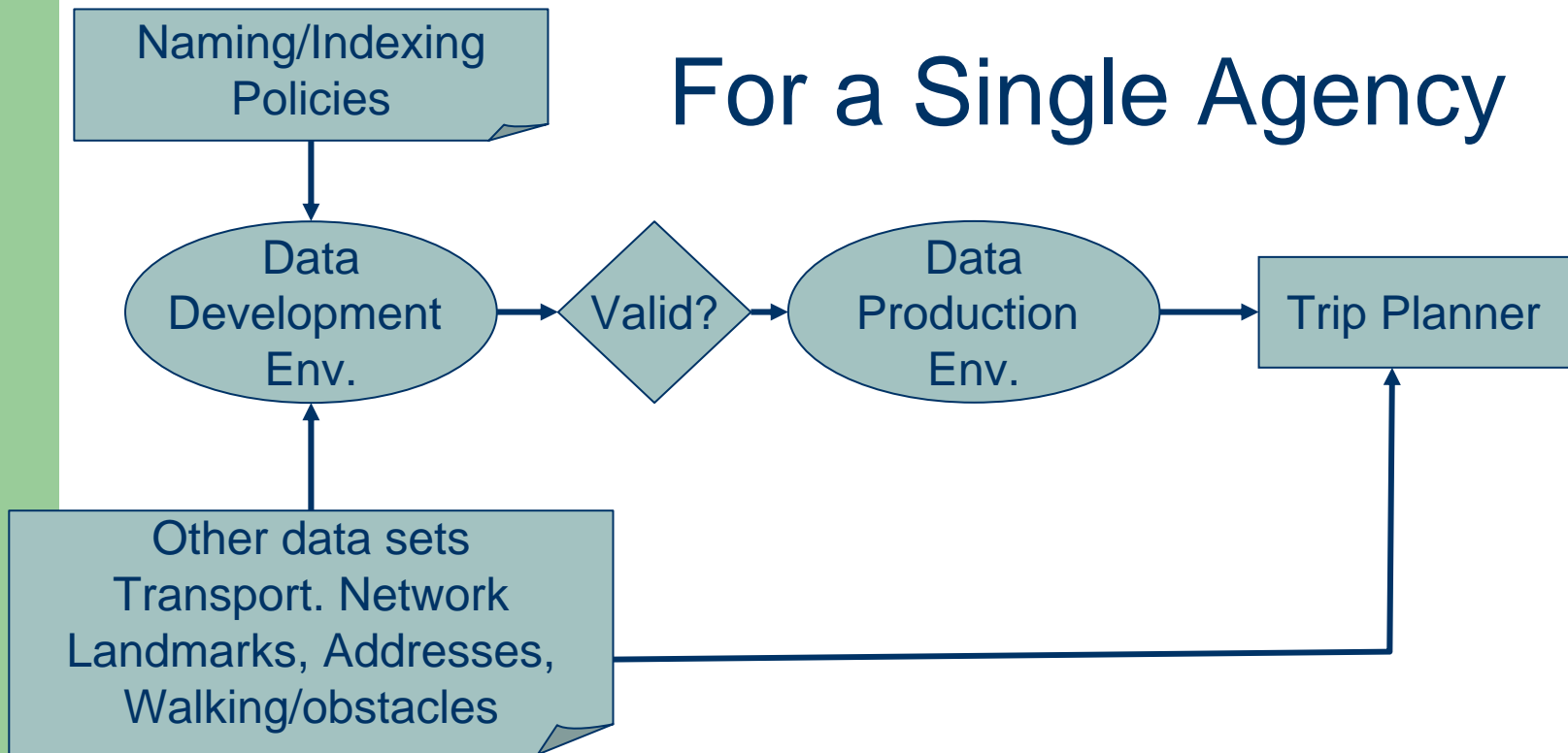


HOW DO STANDARDS APPLY TO MULTIPLE AGENCY DEPLOYMENTS?

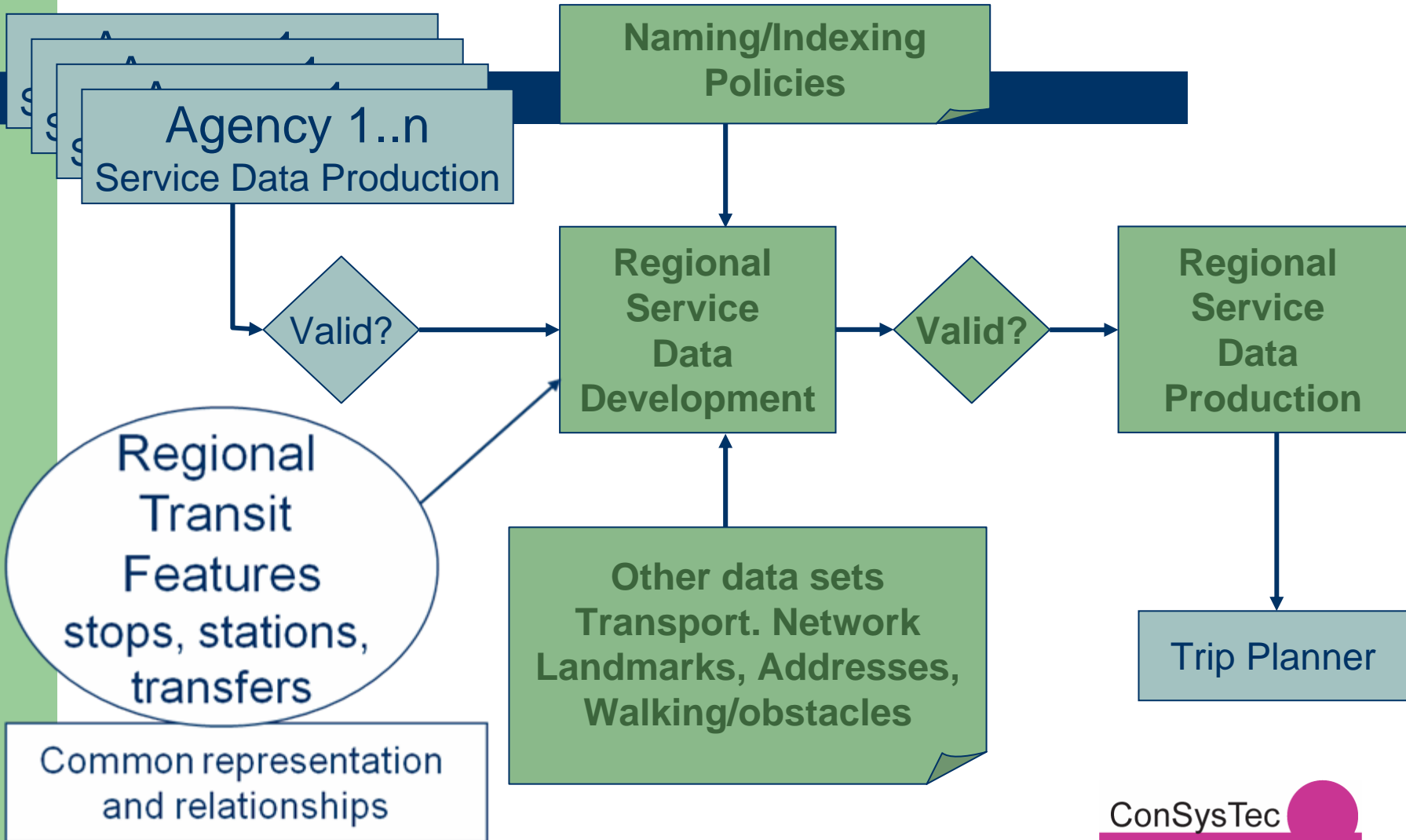


Features and Characteristics

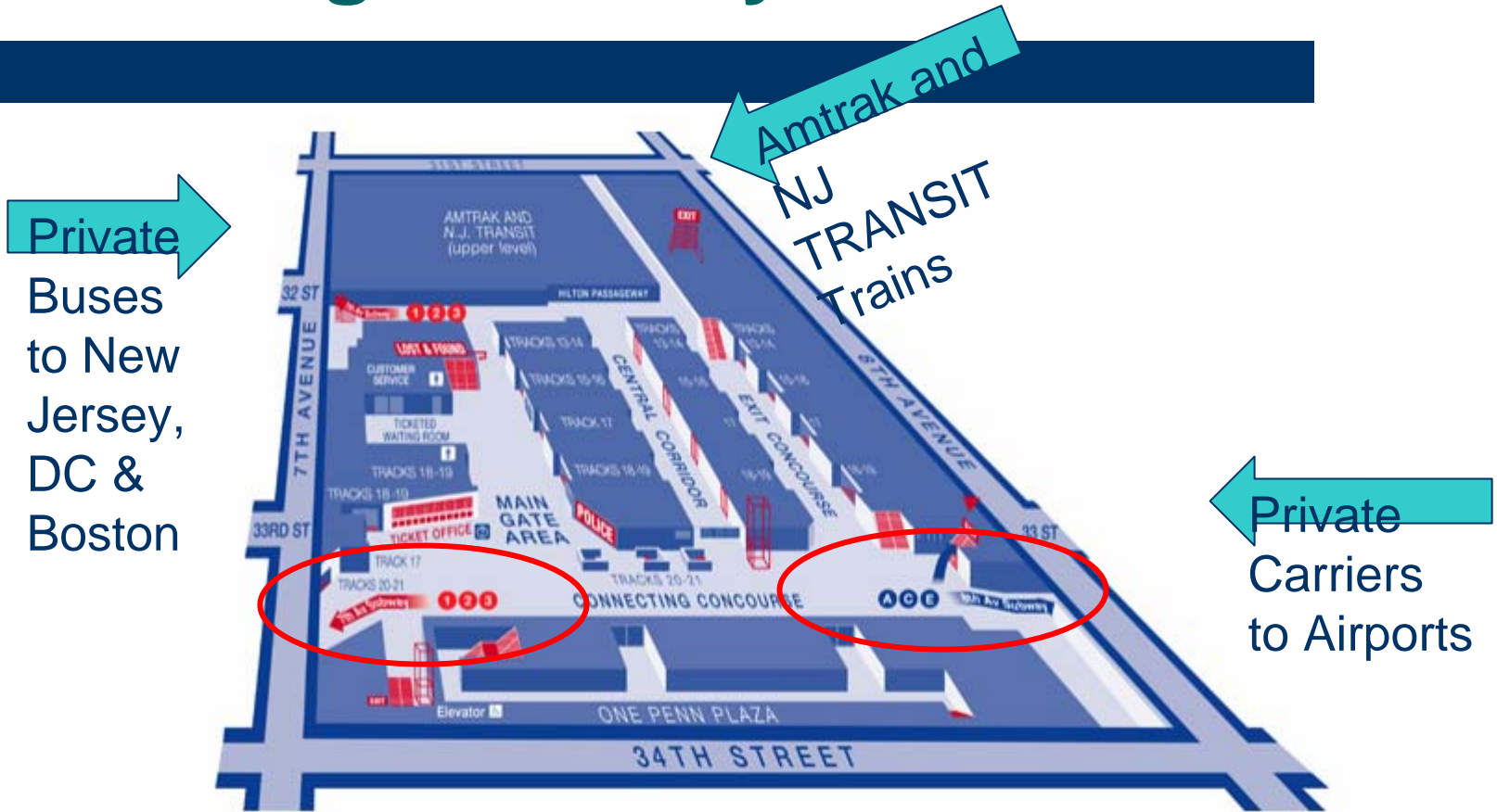
For a Single Agency



Regional Integration Features

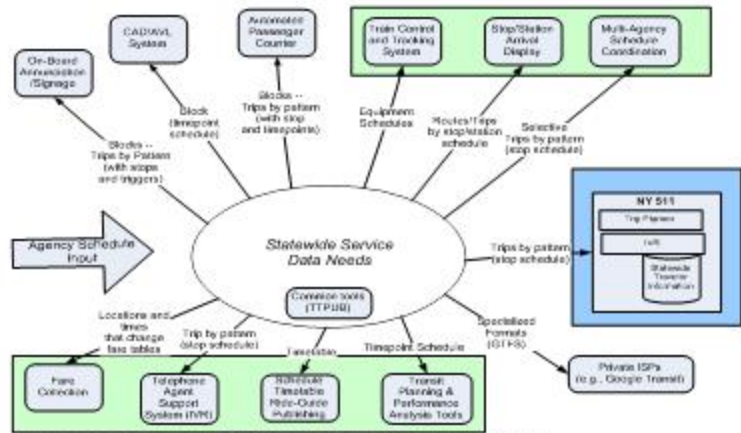
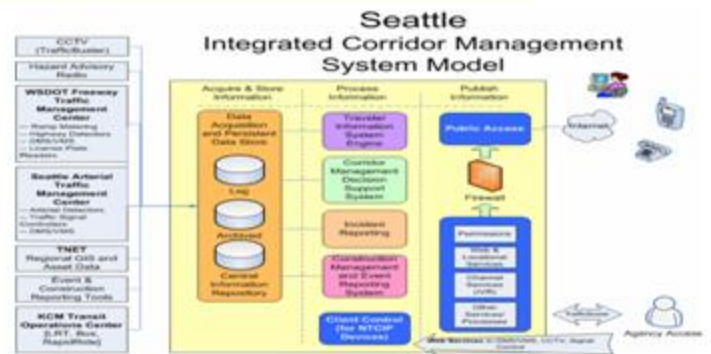
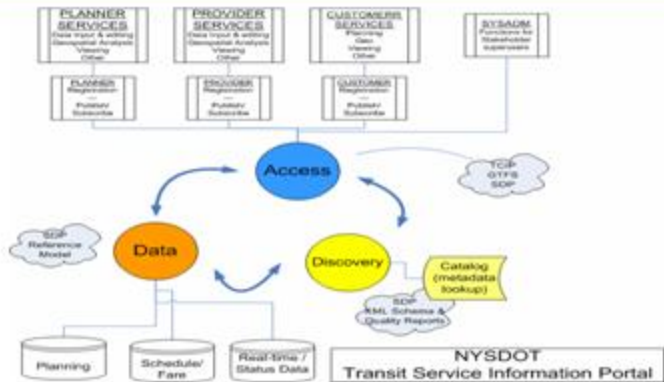


Complex Public Transport Passenger Facility

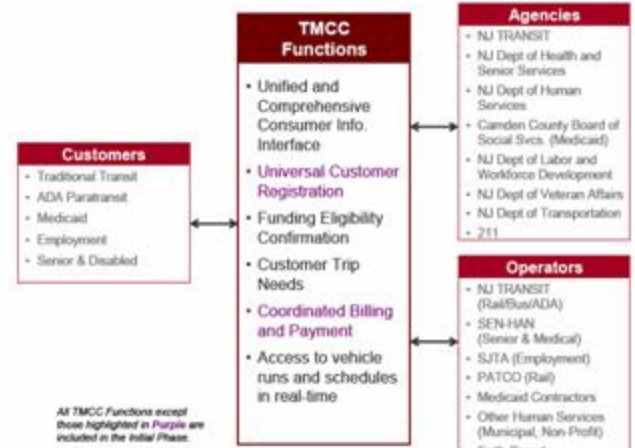


New York Penn Station (only for Long Island Rail Road)

Approach to Transit Data Integration in the US

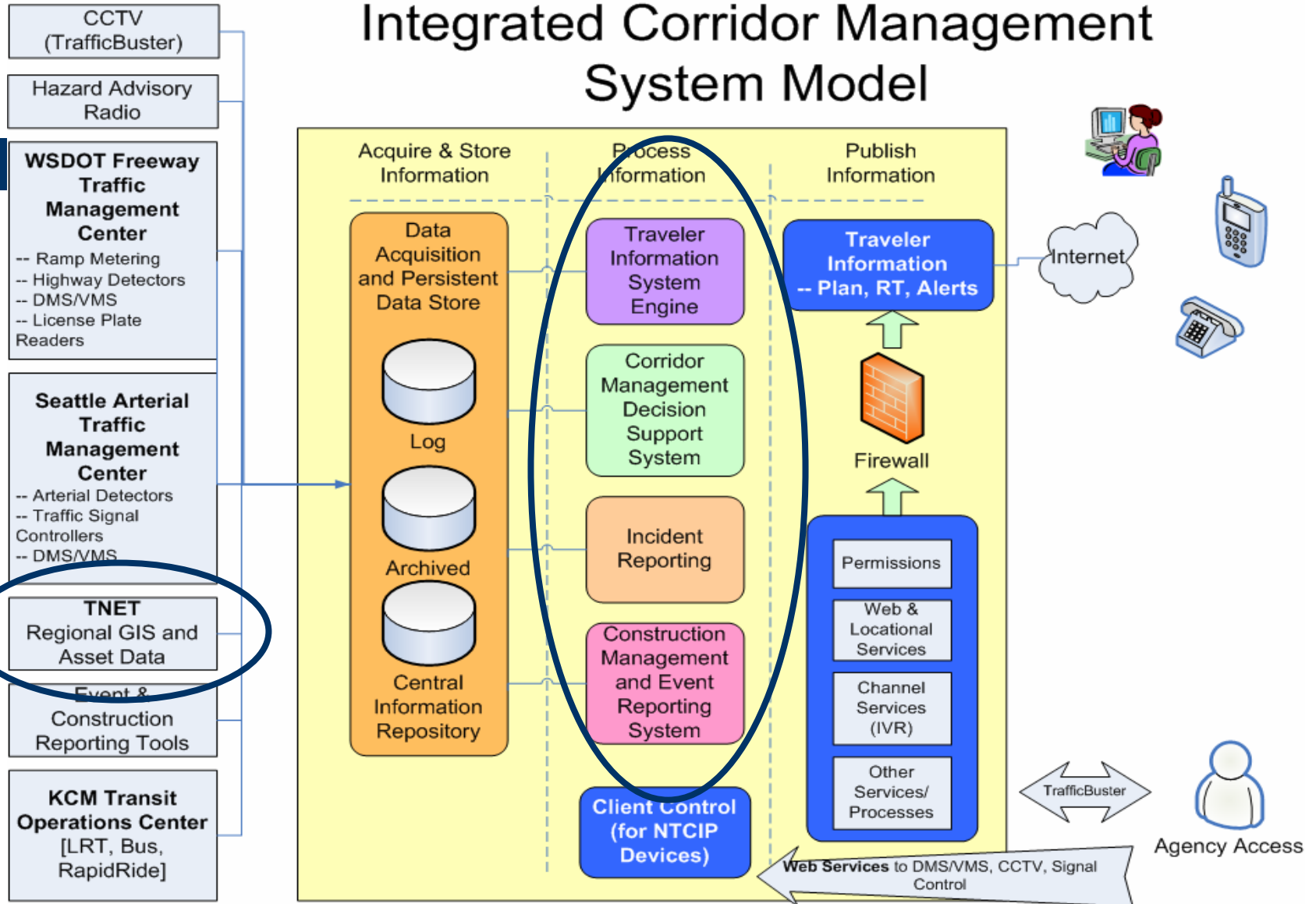


Based on existing processes between NJCT, NJT, and NJT's regional agencies and operational schedule data systems.



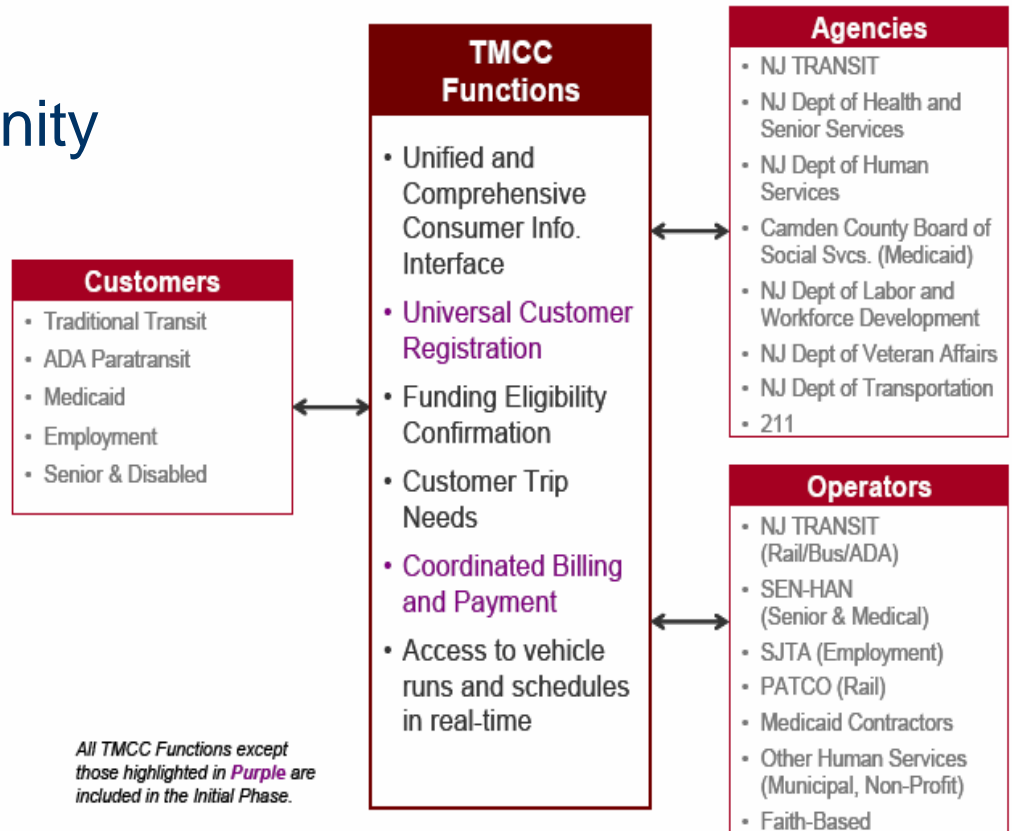
Seattle

Integrated Corridor Management System Model

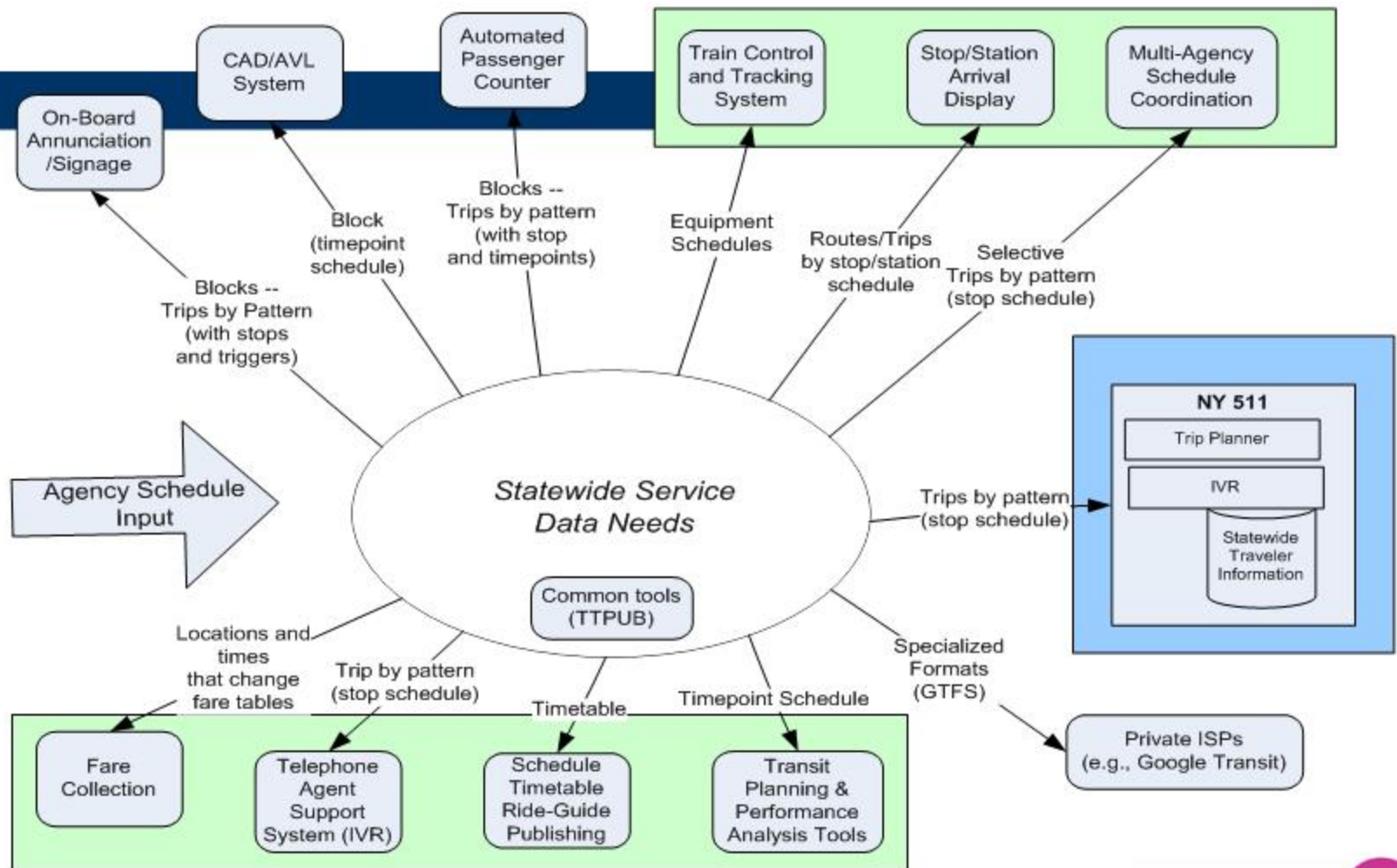


Camden County Transportation Management Coordination Center

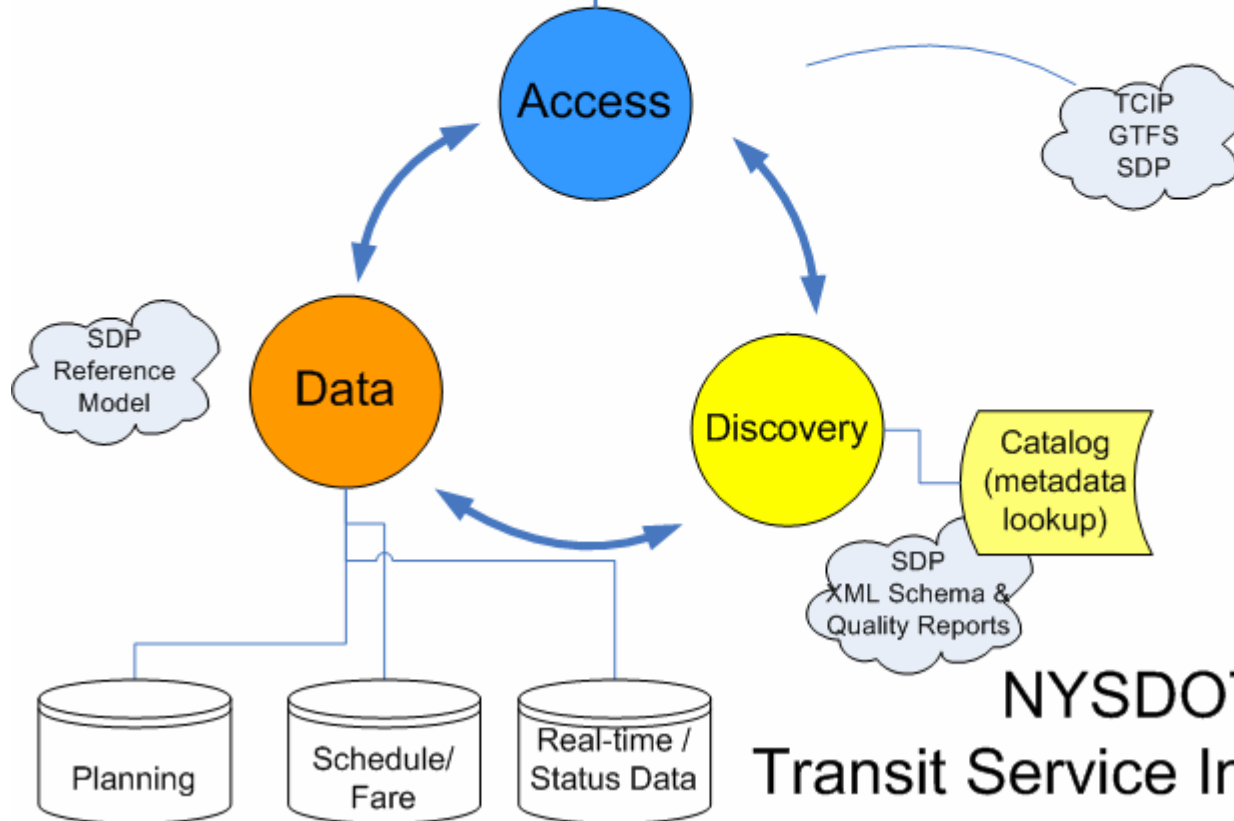
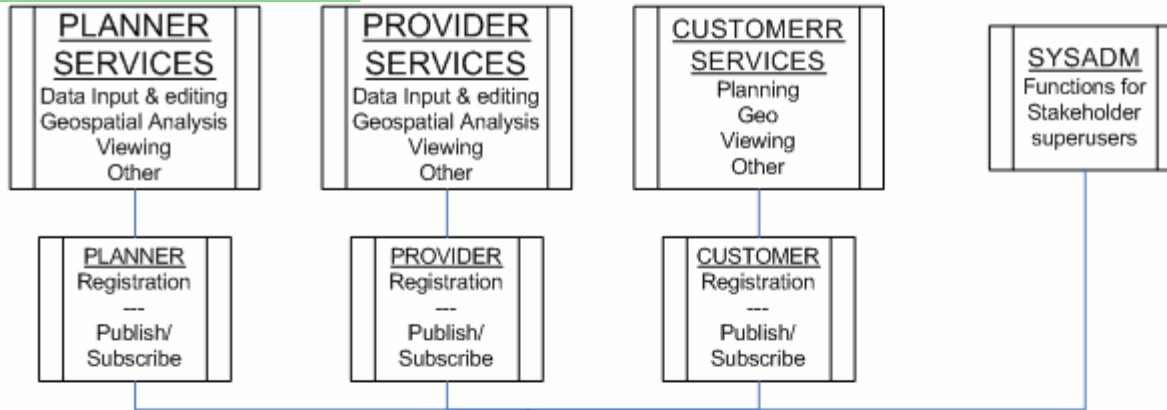
- Maximize funding sources for Community Mobility
- Optimize schedules across providers
- Enhance customer satisfaction



Data Sharing Concept for NY



Boxes surrounding processes indicate current (blue) and future (green) regional applications or coordinated schedule data activities



NYSDOT Transit Service Information Portal



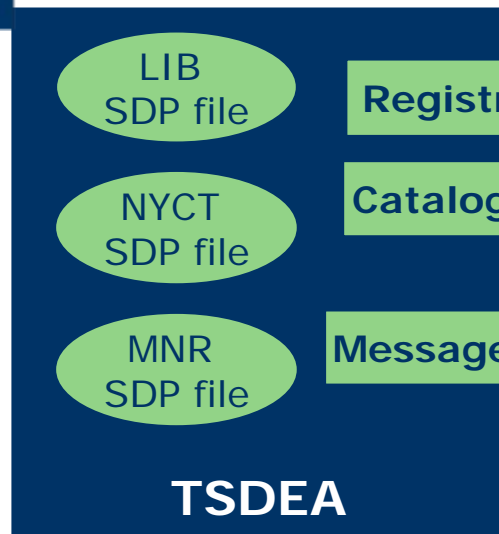
What is the architecture concept?



Schedule Data



Native to SDP Data
Scripts / Validation
and Integrity Checker

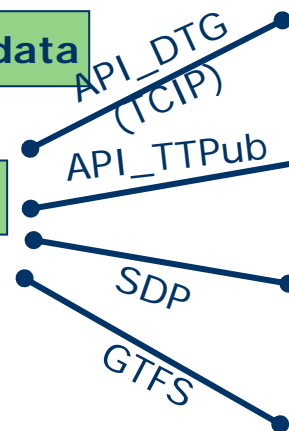


Registry

Catalog/Metadata

Message Editor

TSDEA



**Dynamic
TT Gen.**

**Timetable
Publisher**

TRIPS 123

**Google
Transit**

Case Study on Data Quality

- Example from *NYSDOT Transit Schedule Data Exchange Architecture (TSDEA) Project*
 - Objective: Collect and store multi-agency schedule data that supports multiple downstream applications
 - Developed ***Schedule Data Profile*** (SDP) to ensure consistent transit data quality for use in regional transportation systems.

What is Data Quality?

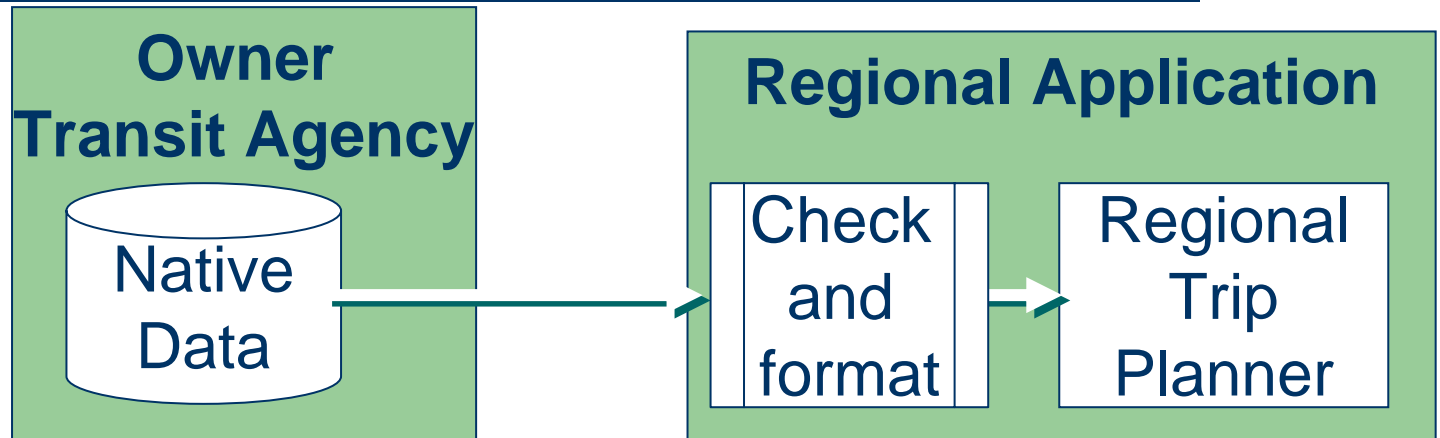
- “*Fitness-for-use* of a particular piece of data for your application”
 - From Transit Location Referencing Guidebook
- What if the data will be used for multiple downstream applications and services?
 - Data requirements “fit for use” for all applications

How is Quality Achieved?

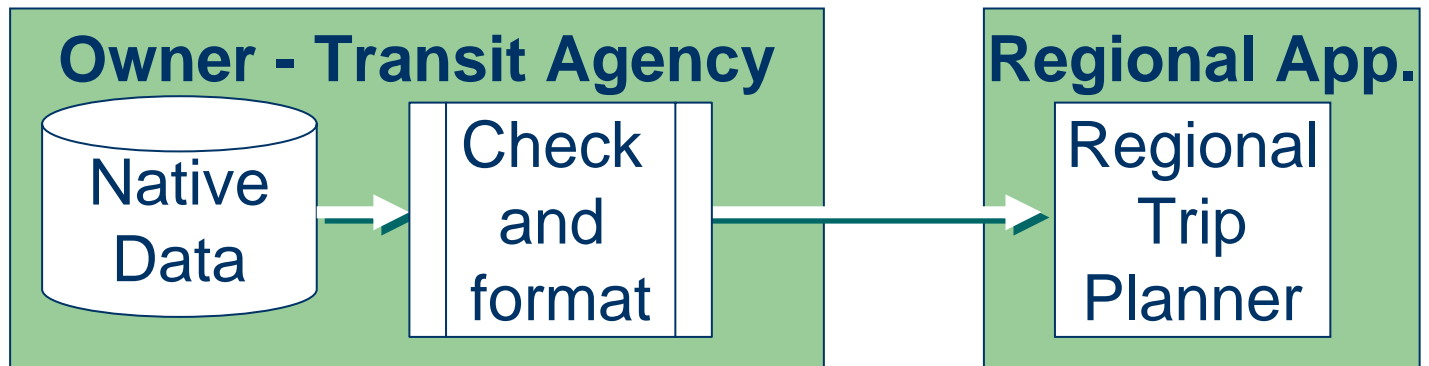
- Quality Checking
- Role of Custodian and Users
- Requirement description of downstream uses including defining data business rules, identity, persistence, regional authorities, etc.

Data Quality “Owner”

Option 1



Option 2



Quality Elements

- Semantics
 - Unambiguous data definitions
 - Relationship among concepts
 - Business Rules
- Syntax/format
- Accuracy, Lineage, Currentness, Logical Consistency, Complete

Concept to Design

Going beyond the Data Dictionary

- Semantics: “data concept meaning”
 - Model relationship of concepts,
 - Describe business rules for using data, and
 - Interpret equivalent (“Isomorphic”) relationships.
- Format Requirements
 - Relational database
 - Exchange while minimizing quality checking
 - Facilitate exchange of large files



Design to Implementation

- Various Implementation Methods
 - Physical Database Query capability (SQL)
 - XML Schema (e.g., SOAP & XML Document)
 - Comma Delimited (CSV files)
- Quality checks
 - Levels of Quality Checks
 - Syntax
 - Referential Integrity (uniqueness & logical consistency)
 - Dates
 - Other business rules related to data concepts (e.g., day type, pattern, facility, location)



TSDEA Levels of Quality Checking

- Level 1: Registration –
 - Ensures that the file contains a well formed and complete SDP XML document.
- Level 2: Authorization –
 - The file content has passed quality checks that are based on business rules and requirements. The file content is deemed logically consistent (semantically and logically accurate).
- Level 3: Regionally Consistent –
 - File content has passed tests to ensure consistency with regional naming conventions and representations.

Contact Information

Paula Okunieff

Consensus Systems Technologies Corp.

43 Jamaica Street

Boston, MA 02130 USA

Email: paula.okunieff@consystec.com

Telephone: +1 617-983-3364