Real-Time Transit Information 101

TPB MANAGEMENT, OPERATIONS, AND INTELLIGENT TRANSPORTATION SYSTEMS (MOITS) TECHNICAL SUBCOMMITTEE MEETING

August 7, 2013

Eric Randall, DTP

Objective

- Information review of Real Time Transit Data requested by TPB Regional Bus Subcommittee members following discussion at January 2013 meeting
  - Discussion of availability and permitted uses of regional transit data held by WMATA and/or available publicly from agencies

- Institutional context
  - Multiple agencies -> multiple data sources
  - Different formats, different uses
  - Liability for data accuracy & integrity
    - Use of data for purposes other than the original intent
  - Relationships with third parties (i.e., NextBus)
  - Cost of providing data

Transit Info: Real-Time and Static

What do we use transit data for?
- *Traveler Applications* built for desktop, web and mobile platforms
- *Transit Agency Operations and Planning Applications* that leverage the data to improve safety, efficiency, regulatory compliance, and customer satisfaction of transit operations
  - MWCOG uses for planning and Regional Data Clearinghouse

What transit data is there?
- *Static Data* – most static transit data is now available in electronic form provided in:
  - Trip planners via the web, and
  - In standard formats like GTFS
- *Real-time Information* – larger transit providers are now providing in:
  - Various forms via web tools and applications
  - In various formats

Static Data: Data Visualization

- Initial version presented at Jan 2013 RBS meeting.
- WMATA requested permission to provide data to developer (STL Transit) for all regional agencies.
- 2nd version published February 2013.

- *Essentially* – data version of traditional paper schedules.

http://www.youtube.com/watch?v=J8sjGQV_beo#at=31

Real-Time Data
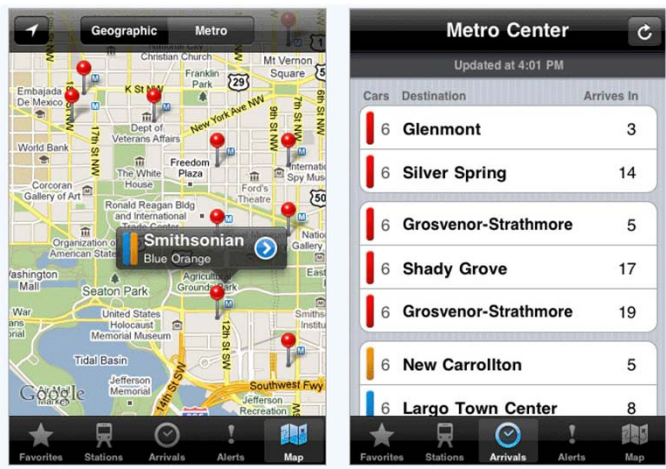
– What customers want!
- Smartphones
- Displays

– Also used for operations:
- Dispatching / service control
- Incident response

DC Metro Rails

Get real time train arrivals and rail alerts for the Washington DC Metro

Get around the city with DC Metro Rails and your Android phone. Use real time train arrivals and rail alerts to help you catch the next train. Find the nearest station entrance and get there with the map and compass. Favorite your most used stations or make a shortcut to them on your home screen.

TPB overseeing TIGER project for Real Time Passenger Information (RTPI)

Proposed Locations (225 signs total)

- **16th Street (30)**
- **Georgia Avenue (25)**
- **H St / Benning Rd (28)**
- **Wisconsin Avenue (40)**
- **Addison Road (14)**
- **University Blvd (19)**
- **Veirs Mill Road (25)**
- **VA-7 / Leesburg Pike (31)**
- **Pentagon and Franconia-Springfield Stations (13)**

- WMATA awarded contract in May 2013.
- Coordinating with road agencies, utility companies, and bus shelter contractors to install and electrify the bus shelters at proposed locations.

What data formats are out there?

<u>Static data</u>

- GTFS - general transit feed specification
  - common format for public transportation schedules and associated geographic information.
  - has emerged as a national standard for static information and for the most part is the standard in Virginia

<u>Real-time</u>   Data standards have yet to formally emerge…

- GTFS-RT – variation of above for real-time info
  - allows public transportation agencies to provide real-time updates about their fleet to application developers.
- SIRI (Service Interface for Real Time Information)
  - European standard specifically designed for public transport
- APTA Transit Communications Interface Profiles (TCIP) Standard
  - Developed in 2001. Use in industry?

What is an API?

- Application Programming Interface (API)
  - "Specifies how software components should interact with each other. In practice, most often an API is a library that includes specifications for routines, data structures, object classes, and variables."
  - Formatted data exchange between two computer systems
    - Specific data request (e.g., what time next bus at place X?)
    - Specific data response (e.g., all bus lines that serve X.)
  - Requires common format/standards
    - Not just language or format, but what terms, what order, etc., in a common "vocabulary".
    - Popular formats include: JSON, XML, Protobuf
  - *These APIs are typically proprietary to the developers*

Uses of an API

- Two kinds of API: *Everything* and *Transactional*
  - "Everything" – lots of data -> maps, displays, and high-connectivity situations.
    - e.g., online real-time map, lobby display
  - "Transactional" – limited data -> ask a specific question, useful for smart-phones or low-connectivity situations.
    - e.g., what time is next bus at by stop?

- The use for which the API is intended determines the amount of data and in turn the type of data format
  - API data requires management: WMATA uses Mashery to control data access -> which limits speed at which data can be accessed

Where is the transit industry on real-time data formats?

- NYC – modified SIRI (for bus) and GTFS-realtime (for rail)
- LA – internally developed API and Nextbus API
- CTA – Clever Devices BusTime API, separate rail API by QuicTrak
- SEPTA – internally developed API
- Muni – nextbus API
- WMATA – internally developed API that includes bus and rail

Where is our region on real-time data?

With APIs/real-time data:

- WMATA – internally developed API that includes bus and rail
  - (Note: NextBus Inc. under contract to use this data for bus predictions)
  - WMATA in process of moving to open API
- Montgomery County Ride-On – Ride-On Real Time (own API), GTFS-RT
- Arlington ART – Data available in multiple formats (Mobility Lab); contractor Connexionz also provides GTFS-RT
- Prince George's The Bus, Fairfax City CUE, DC Circulator – use Nextbus API

  ❖ Ride-On and ART also provide data to RITIS. WMATA soon (?)

All local agencies have data available to create static GTFS feeds:

- Alexandria DASH, Fairfax Connector, Fredrick TransIt, Loudoun County Transit, MTA Commuter Bus: status of AVL?

Developers | WMATA

Home   API Explorer   Documentation   FAQ   App Gallery   Change Log

Welcome to WMATA Developer Portal

The collection of data offered here allows developers to create new and innovative applications for the web or mobile devices. We encourage you to integrate Metro data into your applications and mashups to help get people the information they want about getting around. Use of our APIs and other data is free of charge, but we ask Developers to register with us to receive a development key and agree to the guidelines outlined in our Developer License Agreement. Once registered, our open data will be yours to creatively show off your passion for programming and Metro to make something great!

Metrobus: Build with Washington DC Metro API

Join our developer community and build great applications.

Register

Register and get an API key.

Metro API

Real time Metrorail & Metrobus info & status updates on elevators/escalators.

GTFS Data

Google open format with Metro schedule data.

Service Alerts

Alert & advisory feeds and press releases.

Characteristics of Data Set Definition

- Standard data sets enable subsystem and multi-agency communication

- Proprietary formats can be restrictive or cost prohibitive to convert to a non-proprietary format

- The national trend is for transit agencies and others to make static and real-time information openly available to developers at no charge

- Information clearinghouses like Regional Integrated Transportation Information System (RITIS) and VA 511 can also be data receivers

- We need to develop a better understanding of the cost implications of these choices

RITIS

Regional Integrated Transportation Information System
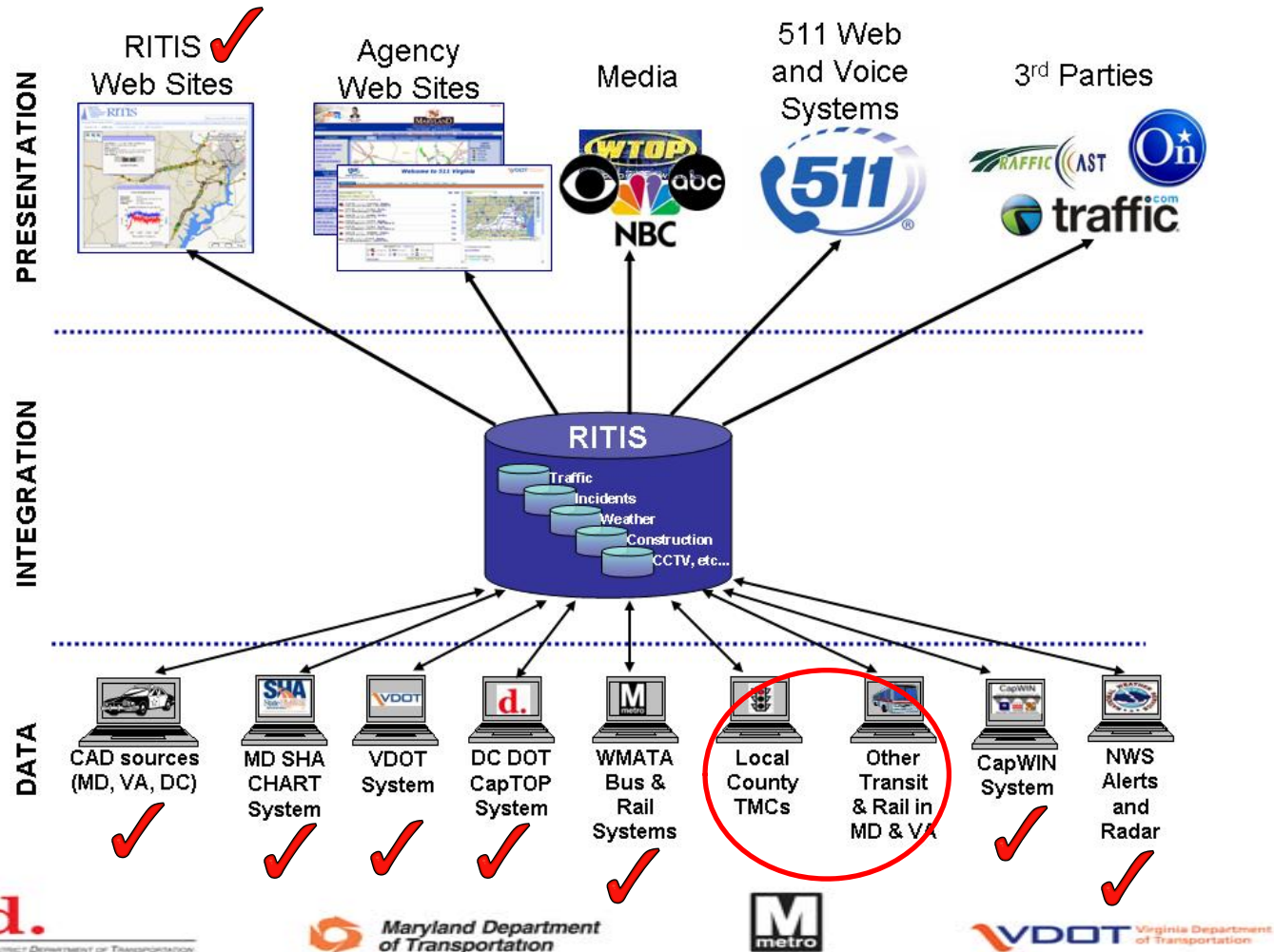
"Working together to reduce incident-related travel delays through improved coordination, cooperation, and information-sharing."

Regional Coordination in Virginia

- DRPT leading efforts for the state to make real-time and historical data available to the public and to 3rd party developers
  - Real-time and static data collection regionally is needed as much for transit planning purposes as for creation of public-facing applications.
  - Scope of this project is to develop the real time API, not the data warehouse specifications
  - Using standards working group to define transit traveler information goals and leverage regional efforts

- The potential benefits of this approach include;
  - Strengthening of standards-based sharing,
  - Out of the box interoperability,
  - Cost efficiencies to agencies by leveraging the existing investment.

Key Findings in Virginia

- Many agencies have either recently procured, currently evaluating or soon issuing RFPs for AVL technology
- All agencies are interested in participating in the real time API
- RITIS is an important stakeholder in the API development. Most agencies underscored the importance of ensuring that providing data to the API and RITIS are as similar as possible

- The three most critical issues facing local agencies in providing real time data:
    - Integrating information from the many disparate transit systems that are in place within each agency
    - Encouraging vendors to provide data in an open, standards-based format
    - **Obtaining technical help given the lack of IT resources within most transit agencies**

- **Most agencies do not have dedicated information technology departments**.
    - Heavily dependent upon city/county IT resources or outside contractors
    - Agencies need guidance from the real time API team on how to ensure AVL vendors provide data in the proper format

Virginia Regional Architecture

To implement in Virginia, need the following:

- Jurisdictional agreements, potential public policy actions

- Standard formats, standard real-time data format

- Hosting locations for static and real-time data

- It is the responsibility of the participating local agencies to:
  - Format all local data as defined in the API specification.
  - Integrate the required data with the regional format
  - Provide a location (or locations) within each agency's infrastructure for retrieving the data required for the API.
  - Periodically update the data and provide to other agencies.

Virginia Regional API Concept - Business Model Options

1. Individual agency publishes its own API
   - This is the current practice. Burden is on the application developer to pull various sources.
   - Potential regional responsibility would be to ensure that each agency is truly standards-based and interoperable

2. Have all of the regional data fed into a regional API
   - Would require investment/support for this regional API.

3. A regional directory of agency API feeds, but each agency would be responsible for building and maintaining its own API.
   - Less regional investment/support; more burden on agencies.

4. Hybrid approach in which the Regional API would aggregate data from each local API.
   - RITIS might be able to serve this role.
   - Arlington County's contractor, Redmon Group, serves a similar role today for pulling static and real-time schedule from various transit agencies and use the consolidated data to feed to the electronic displays (kiosk-type displays).

Four Regional Approaches

DRPT — Virginia Department of Rail and Public Transportation

Local Transit Agency A — API

Local Transit Agency B — API

Local Transit Agency A

Local Transit Agency B

4: Regional API would aggregate data from each local API.

2: All data aggregated and offered at regional level via Regional API

Regional Directory of API Feeds

3: Local APIs to Regional Directory

Regional Data Store and API

1: Direct from Local API to Developer with Regional standard and interoperability

Developer Community / Public Data Consumers / Transit Agencies

Considerations for Regional Coordination

- What would be the regional standard?

- Who would build and manage the infrastructure?

- How would local agencies provide the funding and technical support required to for support of the regional directory or regional data hub?

- What type of governance would be implemented?
  - Ensuring format, regional transit stop identification, periodic updates, etc.
  - Responding to public uses of data, which may alter or use information in unintended ways.

Extra Slides

References / Resources

- Marullus Williams, DRPT
  - Presentation of April 19, 2012

- David Alpert, Greater Greater Washington: Three part series: Jan 16, 18, & 23, 2013
  - http://greatergreaterwashington.org/post/17386/whats-up-with-nextbus-part-1-the-disappearing-app/
  - http://greatergreaterwashington.org/post/17402/whats-up-with-nextbus-part-2-a-pile-of-apis/
  - http://greatergreaterwashington.org/post/17460/whats-up-with-nextbus-part-3-where-ride-on-is-the-leader/

- Kurt Raschke, blogger
  - http://www.kurtraschke.com/

- WMATA PlanIt Metro
  - http://planitmetro.com/2013/02/11/one-day-of-washington-region-transit/
  - http://planitmetro.com/2013/07/19/next-generation-communications/

- Mike Harris, Kimley-Horn
- Amy Tang McElwain, VDOT

Data Set Definition

- Standard data sets foster subsystem and multi-agency communication
- Proprietary formats can be restrictive or cost prohibitive to convert to a non-proprietary format
- The national trend is for transit agencies and others to make static and real-time information openly available to developers at no charge
- Information clearinghouses like Regional Integrated Transportation Information System (RITIS) and VA 511 can also be data receivers
- Google transit information data standard, general transit feed specification (GTFS) has emerged as a national standard for static information and for the most part is the standard in Virginia
- Real-time data standards have yet to formally emerge
- The working group reviewed local existing data formats including;
  - Washington Metropolitan Transit Authority (WMATA) real-time data format,
  - SIRI – transit-specific, highly extensible, and
 - Virginia Tech Bus Tracker

GTFS

- GTFS transit feed specification defines a common format for public transportation schedules and associated geographic information. GTFS-RT is a feed specification that allows public transportation agencies to provide real-time updates about their fleet to application developers.

- GTFS Advantages:
  - Supported by Google. Google provides significant marketing resources for publicizing the availability of agencies'' GTFS data feeds. Easy for agencies to adopt standard and quickly display data via the popular Google Maps service.
  - Robust online documentation and forums to provide support to transit agencies
  - Free to connect to GTFS
  - Many transit technology vendors have adopted GTFS
  - There is a large community of developers familiar with Google's API specifications

- GTFS Disadvantages:
  - Completely dependent upon Google's support; if Google ceases support for GTFS, the standard would be in jeopardy of obsolescence
  - Google does not provide access to raw data that it collects from agencies
  - Must agree to Google's inflexible legal terms regarding indemnification
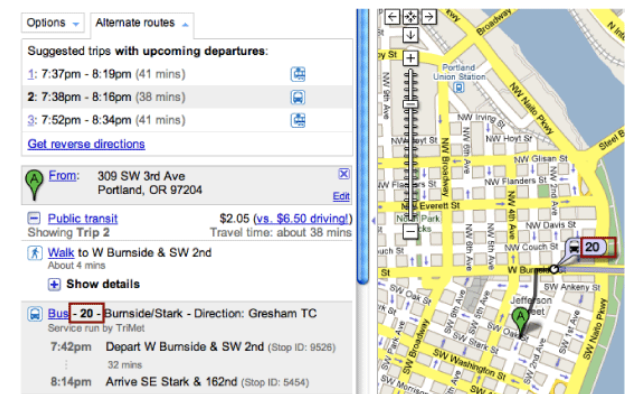
SIRI

- SIRI is managed by a CEN Working Group - TC278 WG3 SG7. SIRI allows pairs of server computers to exchange structured real-time information about schedules, vehicles, and connections, together with general informational messages related to the operation of the services.   The information can be used for many different purposes, for example:
  - To provide real time-departure from stop information for display on stops, internet and mobile delivery systems.
  - To provide real-time progress information about individual vehicles.
  - To manage the movement of buses roaming between areas covered by different servers.
  - To manage the synchronization of guaranteed connections between fetcher and feeder services.
  - To exchange planned and real-time timetable updates.
  - To distribute status messages about the operation of the services.
  - To provide performance information to operational history and other management systems

- SIRI Advantages:
  - Vendor-neutral standard
  - Supports significantly more data elements than GTFS
  - Widely used Internationally
  - Extensible; agencies can create their own custom data fields

- SIRI Disadvantages:
  - Complex to implement
  - Not used as much in the US as in Europe

Concept of Operations OV-1
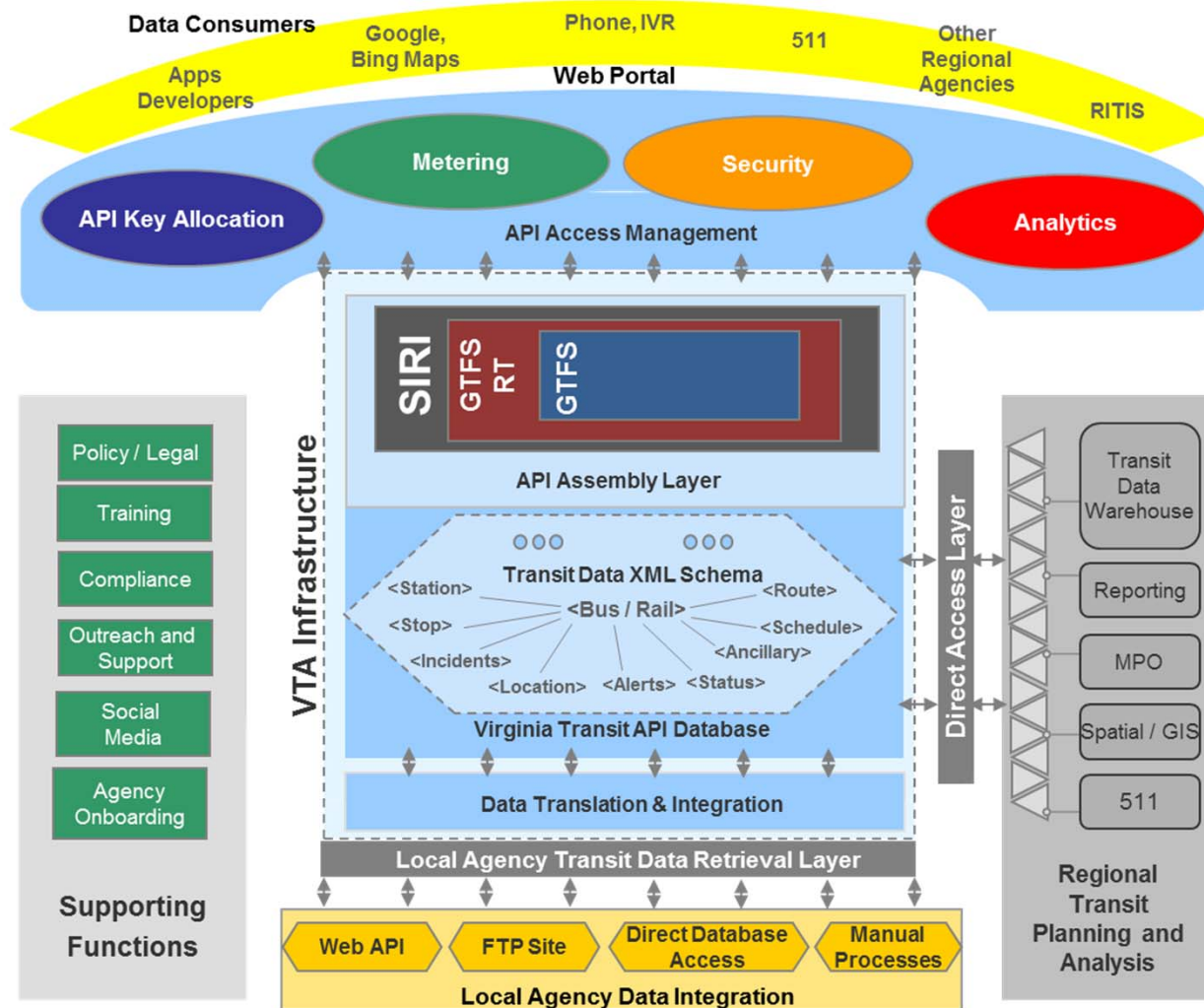
Rolling Out a Successful API Project

- Agreeing to data sets to be published
- Implementing a standards-based approach
- Connecting all required data elements to the API
- Creating a fast, reliable infrastructure by leveraging cloud services and API-specific solutions like Mashery
- Publicizing the API
- Communicate regularly with the developer community
- Building an API forum / community using tools such as Facebook, Twitter
- Managing updates to the API. Good documentation is key.
- Identifying and managing all legal, policy and security risks.
- Monitoring the use of transit data by developers and the public.

API Assembly Layer

- The Feed Assembly Layer packages data that will be provided to Data Consumers. The interface to this layer will be HTTP-based REST protocol, which will respond in one of the supported output formats, SIRI and GTFS/GTFS-RT.

- This layer will have specific modules for converting to XML, Protobuf and JSON formats depending on the request. Protocol buffers (Protobuf) is a binary format used by GTFS-RT and is a flexible, efficient, automated mechanism for serializing structured data.
  - It is smaller, faster, and simpler than XML. JSON is also a small

API Management

- An API Management tool like Mashery would provide the following benefits to the API:
    - Eliminates the need to internally develop API gatekeeping functionality
    - Well-supported and currently employed by WMATA, Best Buy, Netflix, Cnet and others to support publication of APIs for third-party developer use
    - Provides API registration, access and self-service provisioning
    - Provides key issuance and credential management
    - Allows usage control: throttling and limiting tied to key, user, method or group
    - Caches frequently used calls
    - Supports business rules configuration based on filters, parameters, and methods
    - Provides real-time insight to all activity and data export available for independent analysis
    - Provides reports that measure uptime, track errors, and show cache activity
    - Provides API usage information including call volumes, top method calls, and top user activity
    - Includes content management, versioning and documentation change control