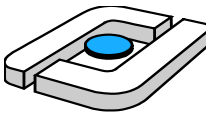


Continuous Information Quality Assessment in Stream Based Smart City Frameworks

19. VDE/ITG Fachtagung Mobilkommunikation
Osnabrück, 22.05.2014

Thorben Iggena

Tel.: +49 541/969-2723
E-Mail: t.iggena@hs-osnabrueck.de



Einleitung

- CityPulse

Stand der Technik

- Quality of Information
- Reputation

Konzept & Architektur

- Smart City Framework
- Komponenten zur Validierung von Zuverlässigkeit und Vertrauenswürdigkeit
- Semantische Stream Annotation

Dynamische Bewertung von Informationsqualität und Reputation

- Funktionsweise
- Beispiel

Zusammenfassung

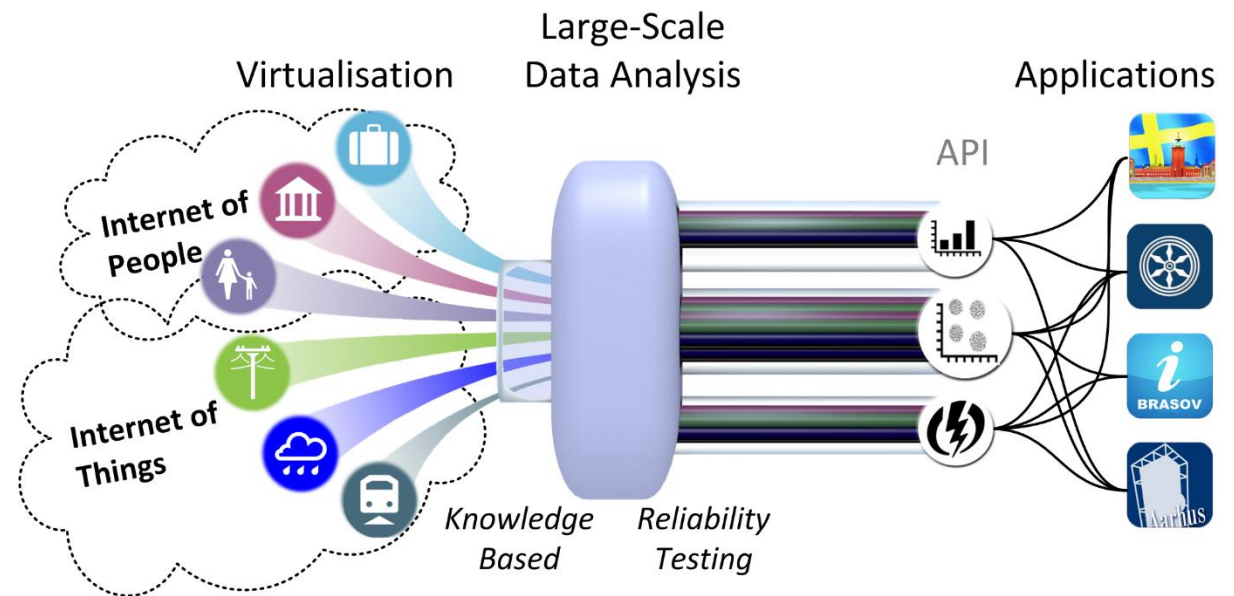


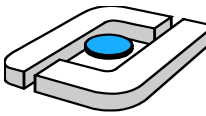
□ Entwicklung

- Wachsende Anzahl von Sensoren in Städten
 - Verkehr, Luftverschmutzung, Müllbehälter, Parkplätze, ...
- Sensorsysteme meist als Silo-Architekturen implementiert
 - Keine Verknüpfung untereinander
 - Heterogene Schnittstellen/Standards

□ Ziele von CityPulse

- Framework als einheitliche Plattform für neue Anwendungen
- Domänenunabhängige Zusammenführung von Datenquellen
- Berücksichtigung von Echtzeit-Anforderungen
- Von besonderer Bedeutung
 - Informationsqualität und
 - Vertrauenswürdigkeit eines Datenstroms





Einleitung

- CityPulse

Stand der Technik

- Quality of Information
- Reputation

Konzept & Architektur

- Smart City Framework
- Komponenten zur Validierung von Zuverlässigkeit und Vertrauenswürdigkeit
- Semantische Stream Annotation

Dynamische Bewertung von Informationsqualität und Reputation

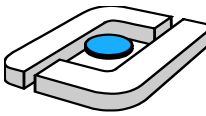
- Funktionsweise
- Beispiel

Zusammenfassung



- ❑ Fehlerhafte Daten können negative Auswirkungen haben
 - Nicht funktionierende Anwendungen
 - Falsch gezogene Rückschlüsse
 - Beispiel: verpasster Bus durch unpräzise Zeitangabe
- ❑ Quality of Information
 - Beschreibt die Qualität einer Information
 - Kann zur Auswahl einer Informationsquelle dienen
 - Ist oft anwendungsspezifisch:
 - Aktualität
 - Besser: Alter
 - Daher Unterteilung
 - Quality of Information Alter, Präzision, Vollständigkeit
 - Value of Information Aktualität, Relevanz, Neuheit, ...





❑ Problemstellung

- Zwei beieinander liegende Temperatursensoren

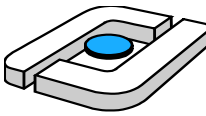
Sensor 1	Sensor 2
5°C	21°C

- Welcher Sensor liefert den korrekten Wert?
 - Gleicher Standort (innen/außen)?
 - Korrekte Funktion der Sensoren?

❑ Reputation

- Herkunft einer Datenquelle von besonderer Bedeutung
- Beschreibt die Glaubwürdigkeit einer Datenquelle
 - Ermöglicht die Bewertung von Datenquellen
 - Aussortierung fehlerhafter/manipulierter Datenquellen möglich





Einleitung

- CityPulse

Stand der Technik

- Quality of Information
- Reputation

Konzept & Architektur

- Smart City Framework
- Komponenten zur Validierung von Zuverlässigkeit und Vertrauenswürdigkeit
- Semantische Stream Annotation

Dynamische Bewertung von Informationsqualität und Reputation

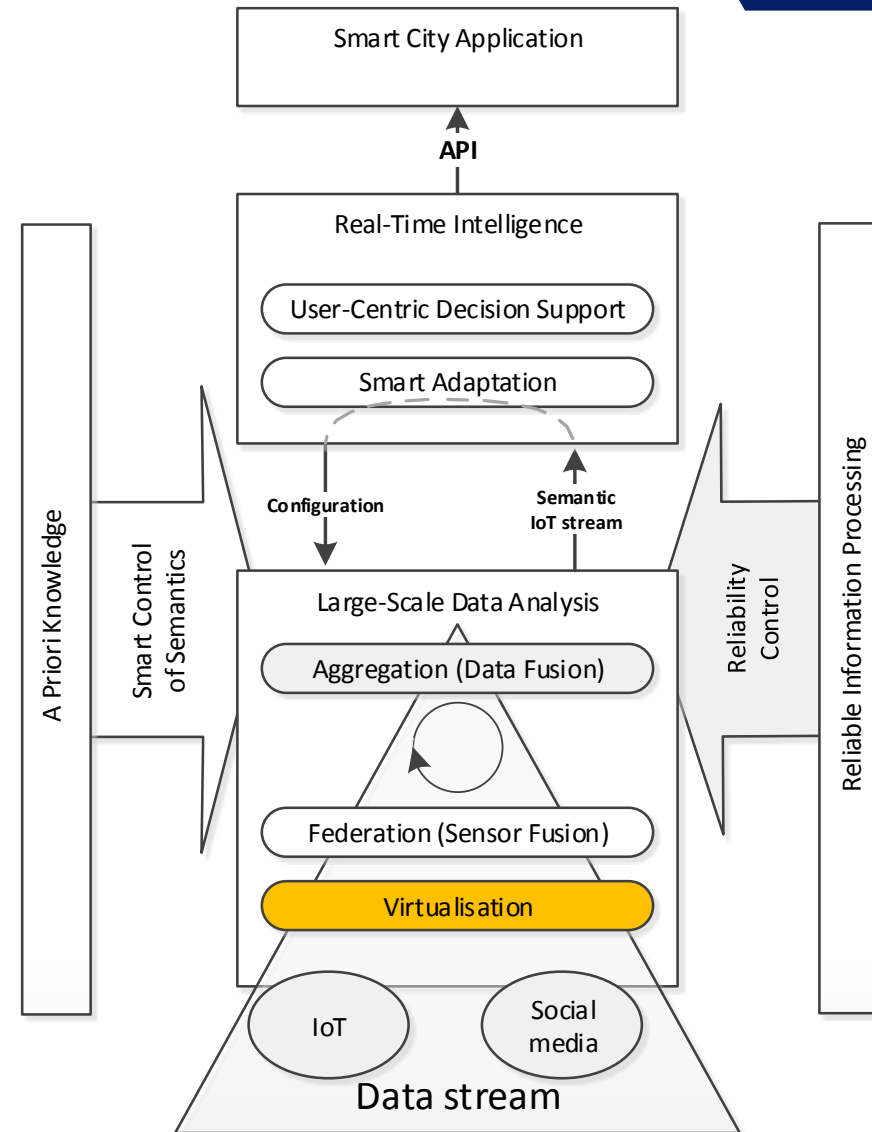
- Funktionsweise
- Beispiel

Zusammenfassung



Virtualisation

- Verknüpfung von Silo-Architekturen
- Verschiedenste Datenquellen
 - Personen
 - Mobiltelefon, ...
 - Social Media
 - Öffentlich verfügbare Dienste
 - Wetterdaten, ...
- Annotation von Datenströmen



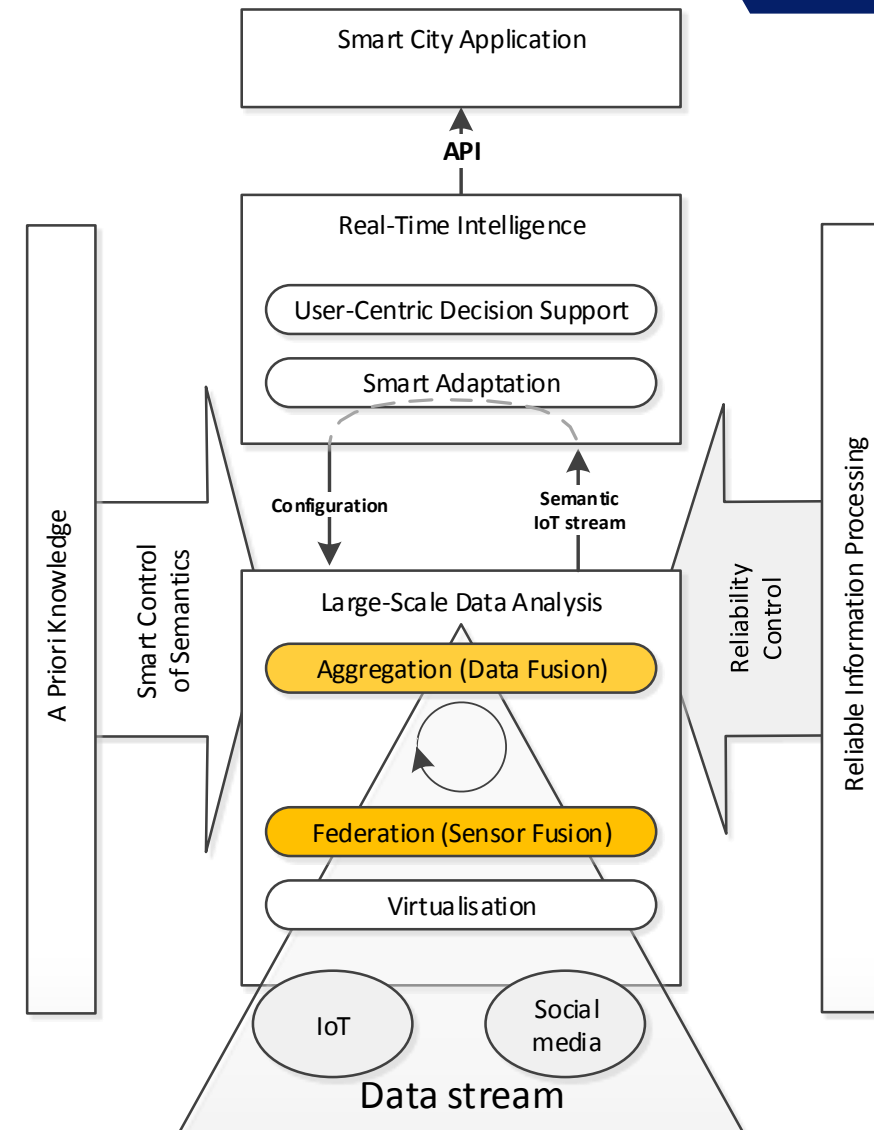
Federation & Aggregation

❑ Sensorfusion

- IoT unterliegt häufigen Änderungen
 - Ausfall von Sensoren
 - Neue Sensoren
- Verbindet heterogene Gruppen von Datenströmen

❑ Datenfusion

- Reduzierung der Datenmenge
 - Clustering
 - Filterung
 - Mustererkennung
- Zusammen mit Federation
 - Extraktion relevanter Informationen
 - Ereigniserkennung

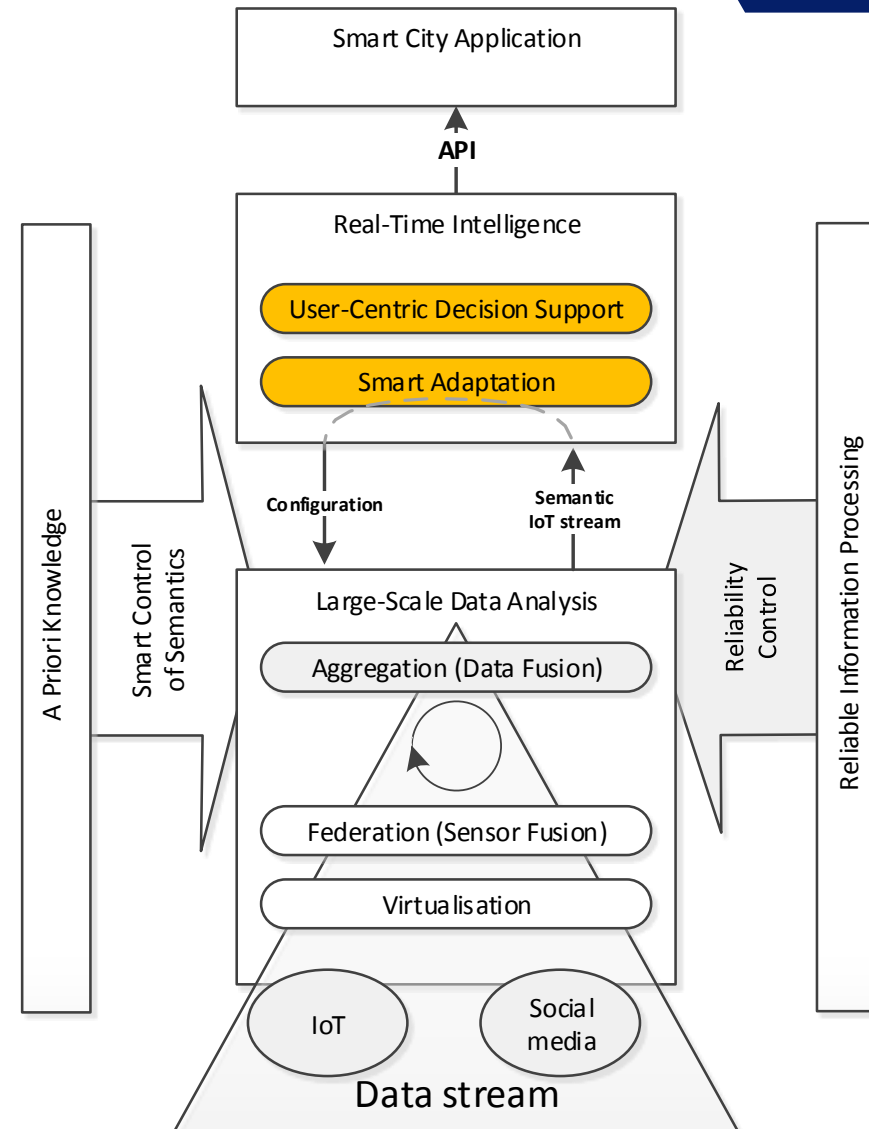


Smart Adaption

- ❑ Interpretation zur Verfügung stehender Daten
- ❑ Umwandlung von Informationen unterer Schichten zu übergeordneten Abstraktionen
- ❑ Adaption der Datenverarbeitung zur Laufzeit mittels wissensbasiertem Ansatz

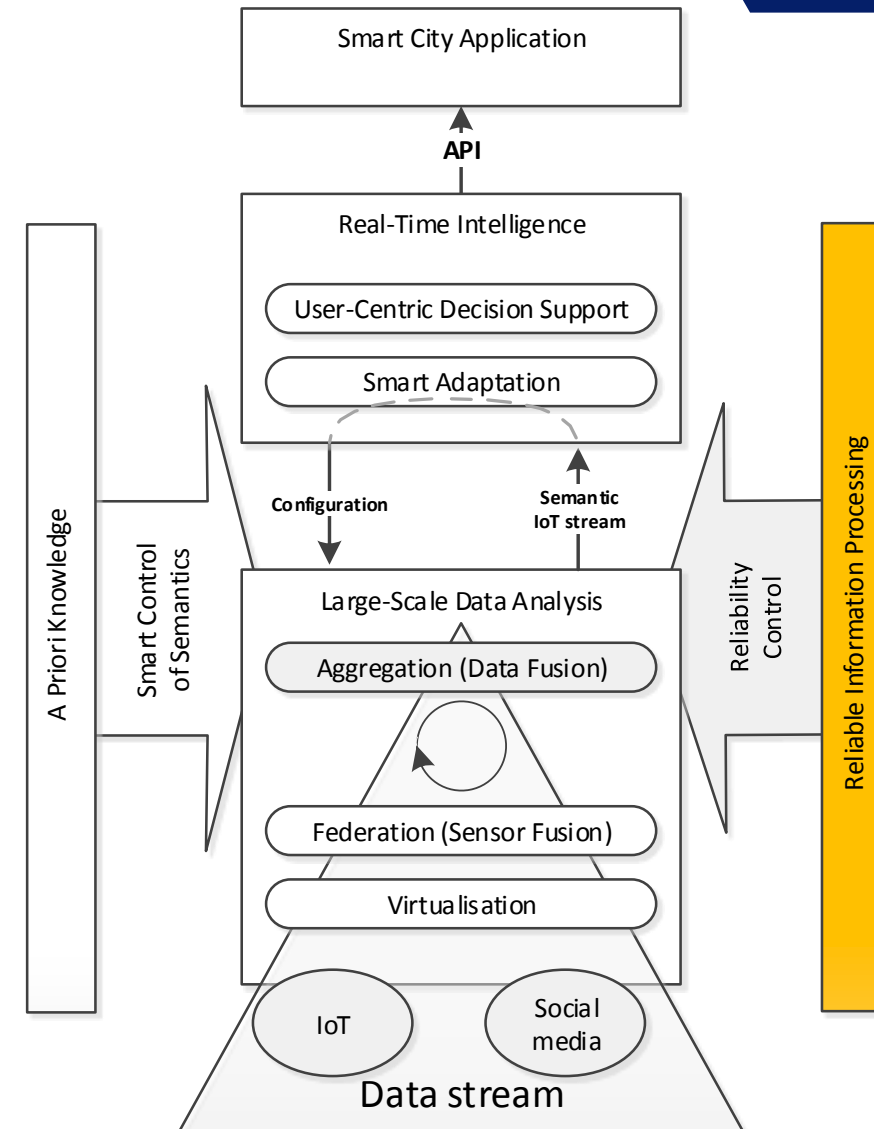
User-Centric Decision Support

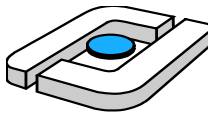
- ❑ Anwender
 - Bürger, Unternehmen, Stadtverwaltungen
- ❑ Angabe persönlicher Präferenzen
- ❑ Kontextanalyse zur Bereitstellung von Daten
 - Präferenzen und Attribute
 - Standort, Zeit, ...



Reliable Information Processing

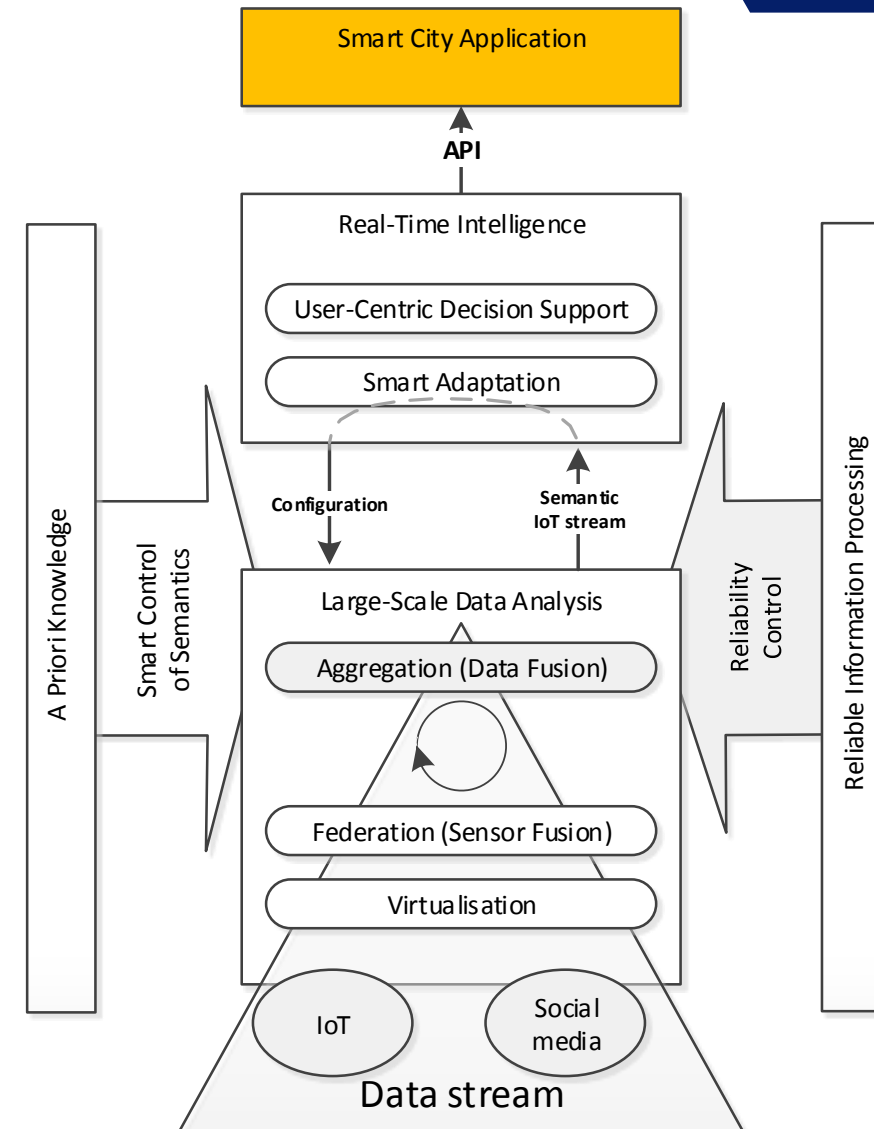
- Dynamische IoT Umgebungen
- Datenquellen unterschiedlicher Herkunft
- Genauigkeit und Vertrauen muss berücksichtigt werden
- Analyse der Streams
 - Raumzeitliche Zusammenhänge
 - Entitätsbasierte Zusammenhänge
- Konfliktlösungsstrategien
 - Fehlfunktionen
 - Verschwinden von Datenquellen



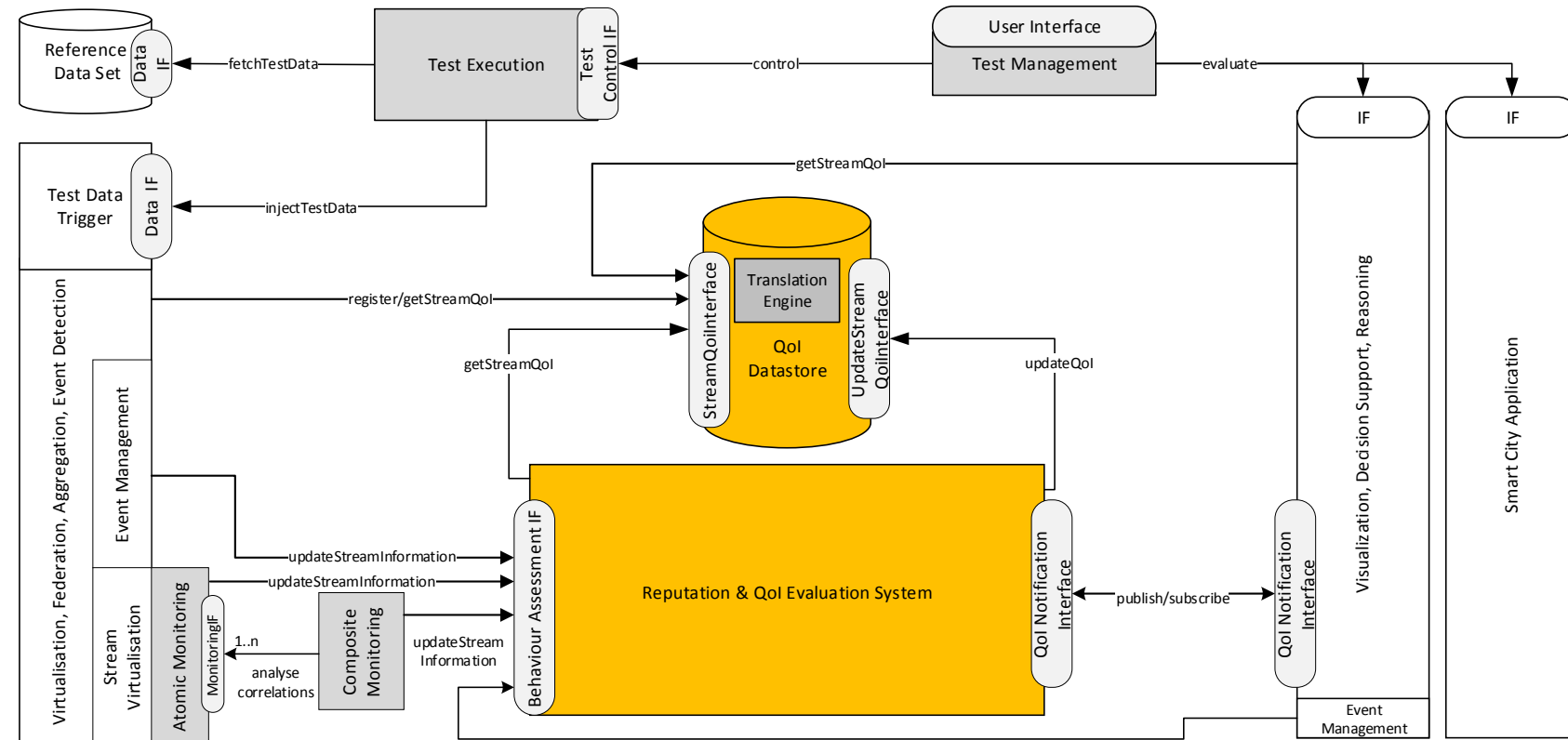


Smart City Application

- ❑ Häufig ändernde Anforderungen in Smart Cities
- ❑ APIs für Anwendungsentwickler
 - Beschleunigte Entwicklung von Prototypen neuer Anwendungen
 - Gesicherter Zugang zum CityPulse Framework



- ❑ Datastore
 - Speichert Qol und Reputation zu Datenströmen
- ❑ Reputation & Qol Evaluation System
 - Aktive Neubewertung von Qol & Reputation
 - Updates durch Events

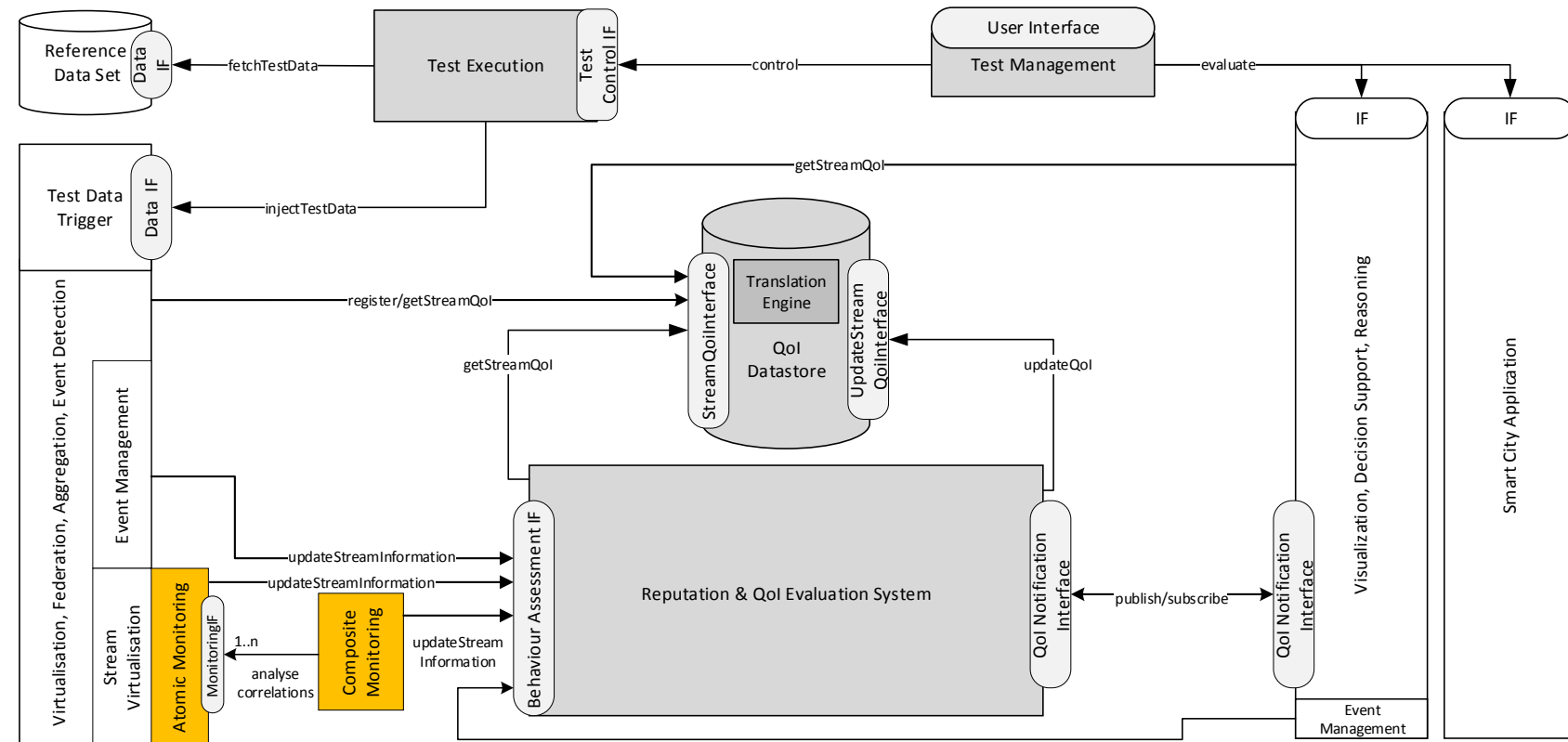


❑ Atomic Monitoring

- Direkte Überwachung der Stream Daten während Virtualisierung auf plausible Werte
- Beispiel: Temperatursensor im Außenbereich zeigt mehr als 80°C an

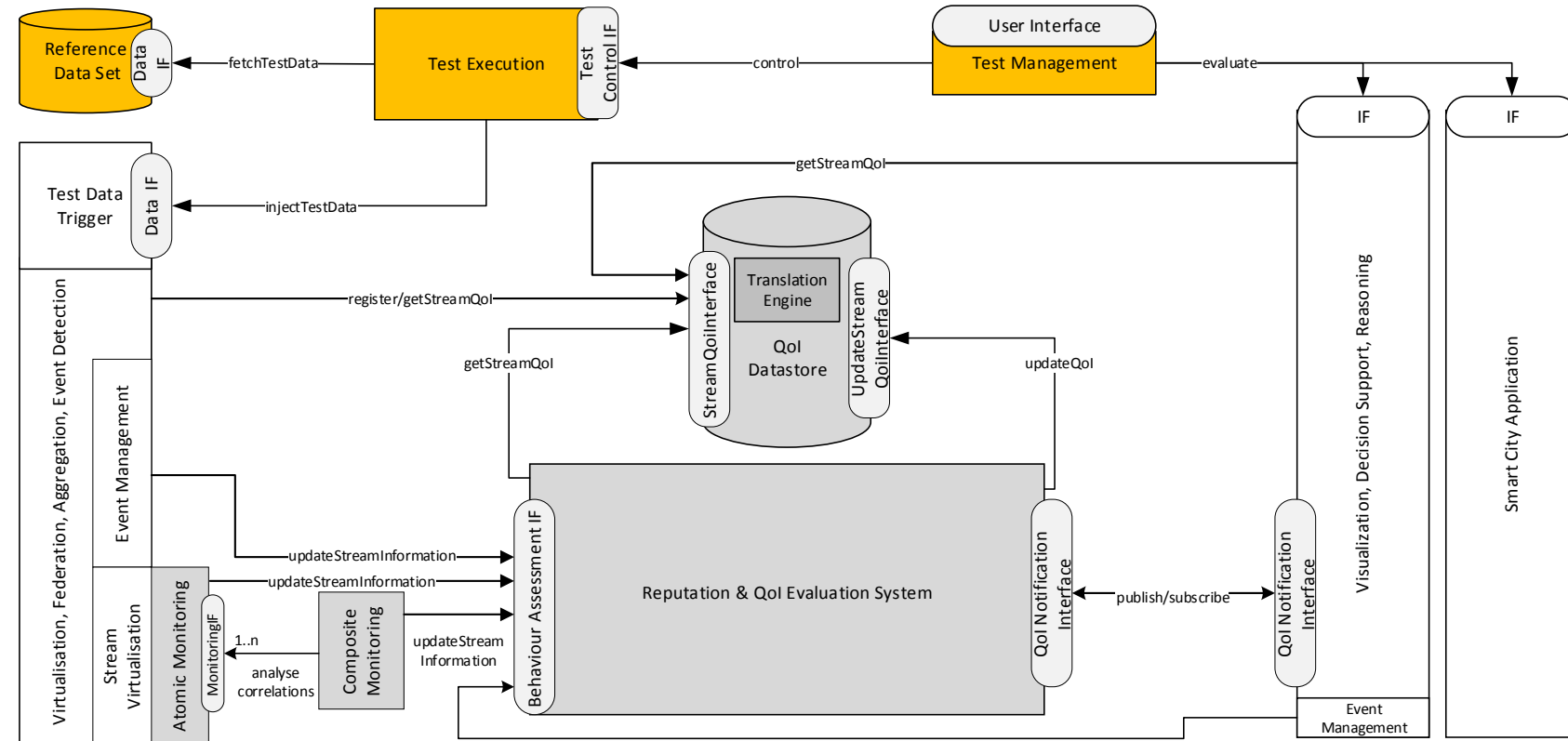
❑ Correlation Monitoring

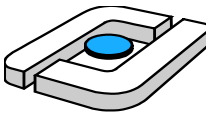
- Prüft raumzeitlich oder entitätsbasiert zusammenhängende Datenströme



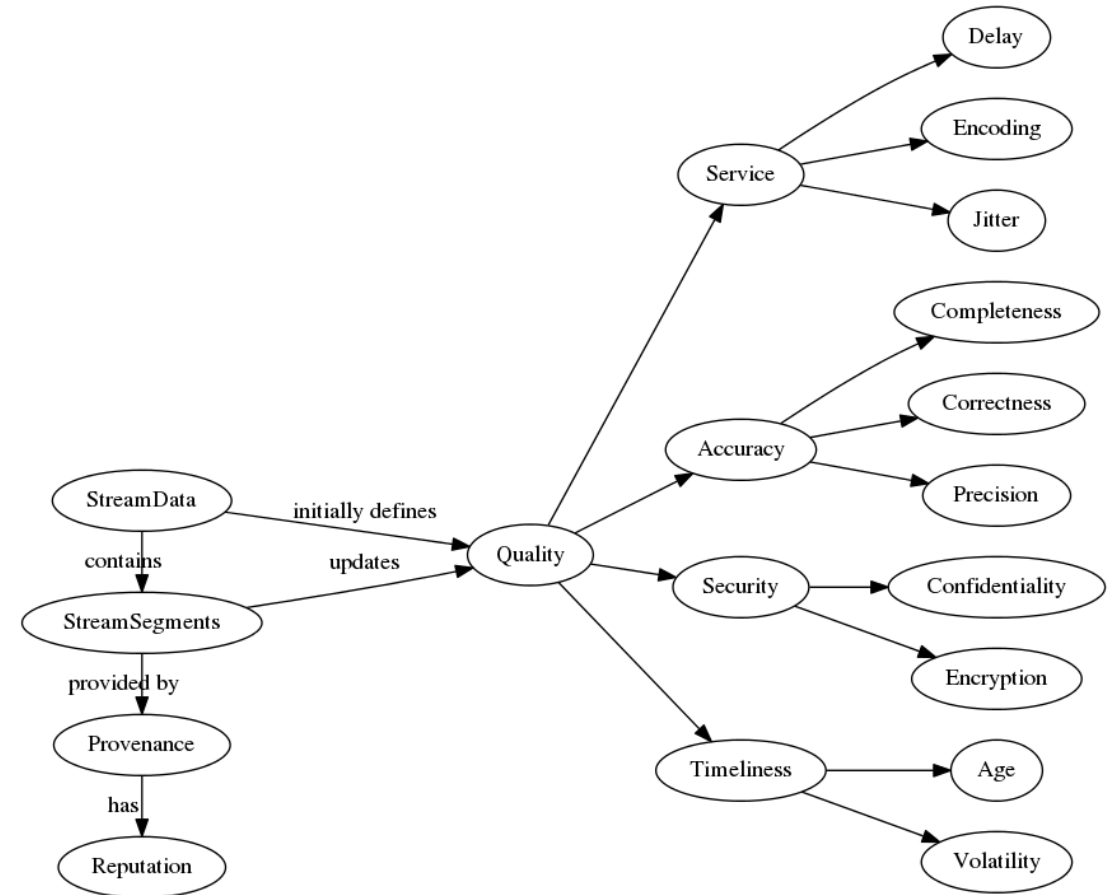
□ Test Management & Execution

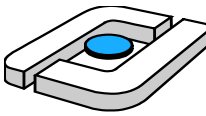
- Validierung neuer Smart City Anwendungen durch Referenzdatensätze
- Einspeisung durch Test Execution





- ❑ Semantische Stream Annotation
- ❑ Ontologie vereint QoS und QoI
- ❑ Nutzung eines Triple-Store zur Speicherung
- ❑ Wiederverwendung bestehender Ontologien
 - SSN
 - Semantische Sensor Netze
 - PROV-O
 - Herkunfts-Informationen
 - Timeline
 - Beschreibung temporärer Informationen





Einleitung

- CityPulse

Stand der Technik

- Quality of Information
- Reputation

Konzept & Architektur

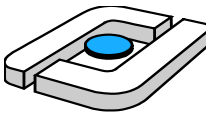
- Smart City Framework
- Komponenten zur Validierung von Zuverlässigkeit und Vertrauenswürdigkeit
- Semantische Stream Annotation

Dynamische Bewertung von Informationsqualität und Reputation

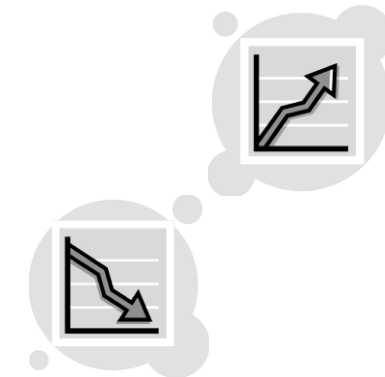
- Funktionsweise
- Beispiel

Zusammenfassung





- ❑ Verknüpfung von QoI und Reputation
 - Reputation unterstützt die Auswahl einer Datenquelle
 - Je höher die Reputation einer Datenquelle desto “wertvoller” ist die gelieferte QoI bei der Zusammenführung von einzelnen Quellen
- ❑ Eingruppierung neuer Datenquellen
 - Niedrige Start-Reputation
 - Zuweisung zur einer Gruppe bereits bestehender Quellen
 - Personengruppe
 - Bestimmter Hersteller für einen Sensortyp
- ❑ Funktionsweise
 - Bei Bereitstellung korrekter Daten → Reputation erhöhen
 - Bei Bereitstellung falscher Daten → Reputation senken



□ Beispiel: Meldung eines verspäteten Busses

- Offizieller Fahrplan: „Bus ist pünktlich.“
- Person(en): „Bus hat Verspätung.“

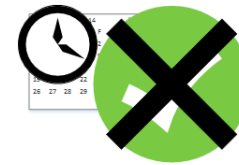
▪ Fall 1

Eine Person meldet Verspätung

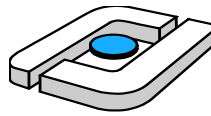


▪ Fall 2

Mehrere Personen melden Verspätung



- Offizieller Fahrplan hat höhere Reputation als einzelne Person
- Aggregierter „Stream“ von Personen hat höhere Reputation
 - $\text{Reputation}_{\text{Personen}} > \text{Reputation}_{\text{OffFahrplan}} \rightarrow \text{Bus ist unpünktlich}$

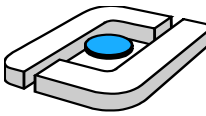


Dynamische Bewertung von QoI und Reputation

- ❑ Einsatz der Reputation bei Zusammenführung von Datenquellen
- ❑ Aggregation von Daten
 - QoI kann ebenfalls zusammengeführt werden
 - Gewichtung der einzelnen QoIs auf Basis der jeweils gültigen Reputation
 - Beispiel: Präzision (Pr) einer Datenquelle

$$Pr_{gesamt} = \sum_{k=1}^n Pr_k \cdot Rnorm_k$$

- ❑ Reputation einer zusammengefassten Quelle kann ebenfalls erhöht werden
 - Identische Werte aus verschiedenen Datenquellen
 - Reputation der einzelnen Quellen kann erhöht werden



Einleitung

- CityPulse

Stand der Technik

- Quality of Information
- Reputation

Konzept & Architektur

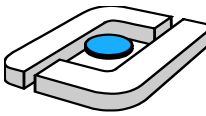
- Smart City Framework
- Komponenten zur Validierung von Zuverlässigkeit und Vertrauenswürdigkeit
- Semantische Stream Annotation

Dynamische Bewertung von Informationsqualität und Reputation

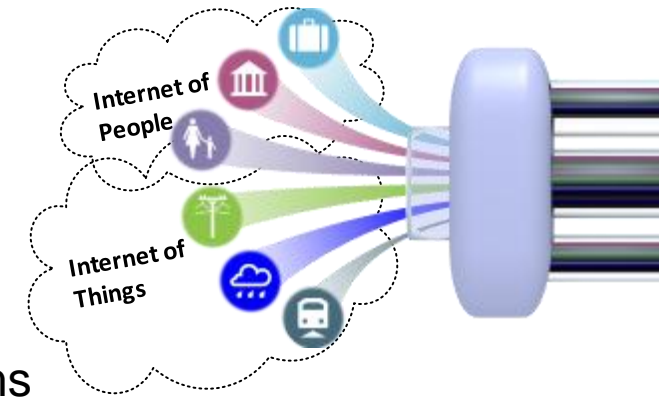
- Funktionsweise
- Beispiel

Zusammenfassung

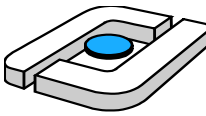




- ❑ Framework
 - Zusammenführung unterschiedlichster Datenquellen
 - Zeitnahe Annotation von Informationen
- ❑ Auswahl geeigneter Datenquelle
 - Quality of Information
 - Unterstützt Auswahl mit erweiterten Angaben zur Qualität eines Datenstroms
 - Reputation
 - Unterstützt die Auswahl auf Basis einer Angabe zur Vertrauenswürdigkeit
 - Unterstützt eine gewichtete Aggregation von einzelnen Datenströmen



Ende



Vielen Dank für Ihre Aufmerksamkeit!

Für Fragen stehe ich Ihnen gerne zur Verfügung