




University of Applied Sciences  
Osnabrück

# Zuverlässige Datenverarbeitung für SmartCity-Applikationen

Daniel Kümper  
Hochschule Osnabrück & Universität Osnabrück

## GiN Forum Mobile Systeme - 14.07.2015

GiN Forum Mobile Systeme - 14.07.2015 1 Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation



University of Applied Sciences  
Osnabrück

## Inhalt

1. Einleitung
2. CityPulse Projekt
3. Beurteilung von Datenqualität / Quality of Information(QoI)
4. Distanz und Interpolation in der Stadt
5. Testansätze
6. Zusammenfassung/Quellen

GiN Forum Mobile Systeme - 14.07.2015 2 Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

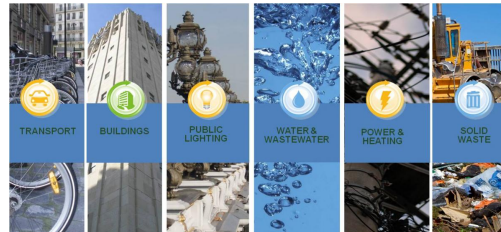
## Herausforderungen



University of Applied Sciences  
Osnabrück

### Context-basierte Smart City Anwendungen:

- Heterogene Datenquellen und eine Vielzahl von proprietären Datenformaten verhindern Wiederverwendbarkeit.
- Silo Architekturen auf Technologie- und Domänenebene blockieren Nutzung existierender Infrastruktur.
- Unbekannte Qualität der Informationen durch verteilte Herkunft.
- Skalierbarkeit bei Echtzeitverhalten muss gewährleistet werden.
- Semantische Lücke zwischen Daten und Bedeutung





## Inhalt



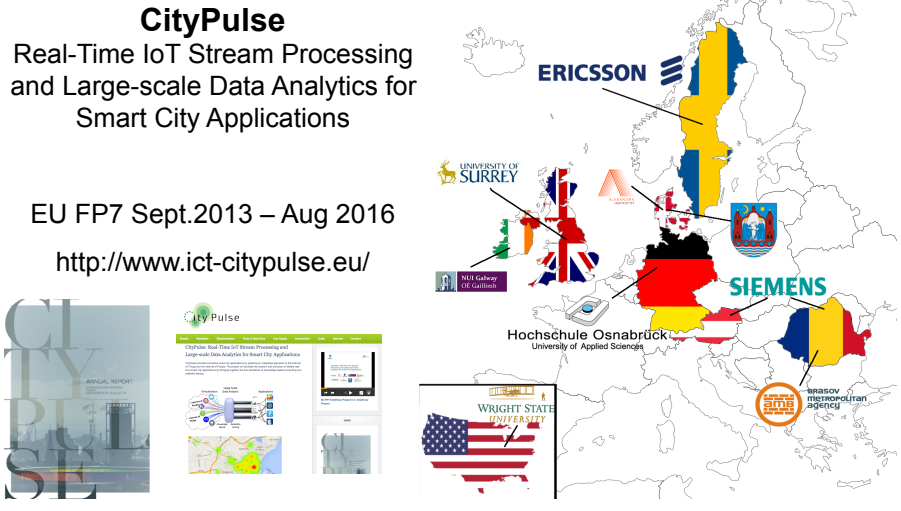
University of Applied Sciences  
Osnabrück

1. Einleitung
2. **CityPulse Projekt**
3. Beurteilung von Datenqualität / Quality of Information(QoI)
4. Distanz und Interpolation in der Stadt
5. Testansätze
6. Zusammenfassung/Quellen

 **Project** 


**CityPulse**  
Real-Time IoT Stream Processing  
and Large-scale Data Analytics for  
Smart City Applications

EU FP7 Sept.2013 – Aug 2016  
<http://www.ict-citypulse.eu/>



GiN Forum Mobile Systeme - 14.07.2015 5 Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

**Smart City Data** 

**Rahmenbedingungen für eine  
Smart City Plattform:** 

- Heterogene Daten verschiedener Anbieter:
  - Stadt
  - Globale Provider
  - Bürger (Smart Citizens)
- Daten sind häufig verrauscht und unvollständig
- Abhängig von Ort und Zeit
- Dynamische Verfügbarkeit mit wechselnder Qualität
- Keine verfügbare "Ground Truth" die unbedingt wahr ist
- Echtzeitverarbeitung von Daten ist notwendig
- Privatsphäre und Sicherheit sind bedeutende Teilgebiete

GiN Forum Mobile Systeme - 14.07.2015 6 Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

## An Integrated Approach

University of Applied Sciences  
Osnabrück

Wiederverwendbare  
Komponenten

GiN Forum Mobile Systeme - 14.07.2015
7
Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

## Semantische Annotation

University of Applied Sciences  
Osnabrück

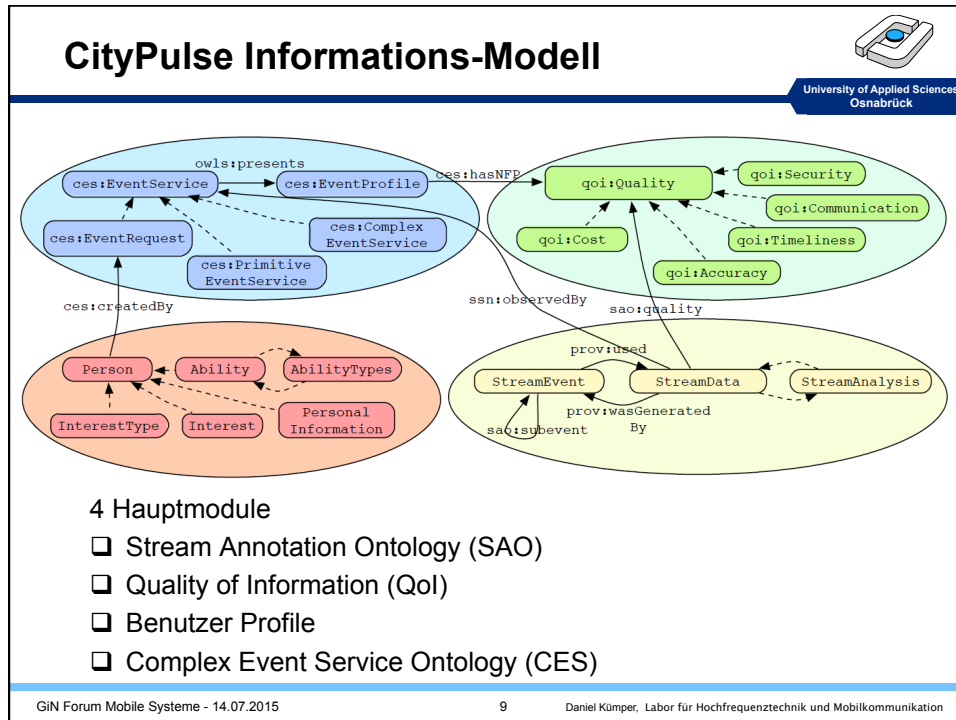
**Problem:** Rohdaten ohne spezifische Metainformation sind bedeutungslos.  
 ⇒ Keine automatisierte Verarbeitung möglich

**Ansatz:** SAO

- Repräsentation aggregierter Datenströme und zeitlicher Bedeutung
- Basiert auf der SSN Ontology und Timeline Ontology

GiN Forum Mobile Systeme - 14.07.2015
8
Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation






## Inhalt

1. Einleitung
2. CityPulse Projekt
3. **Beurteilung von Datenqualität / Quality of Information(QoI)**
4. Distanz und Interpolation in der Stadt
5. Testansätze
6. Zusammenfassung/Quellen

GiN Forum Mobile Systeme - 14.07.2015 10 Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation


## „Quality“-Annotation der Datenströme



University of Applied Sciences  
Osnabrück


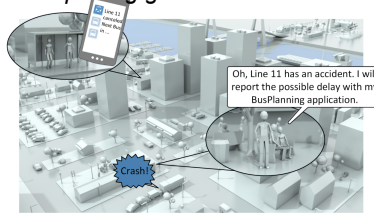
Quality Ontology (QoS & QoI)	Parameter-name	subcategory	abstraction level	Measurement unit	
	Accuracy	Correctness		information	Probability that information is within the range of precision and completeness
Precision		Resolution	information	absolute value in sensing unit	
		Deviation (max)	information	maximum deviation percentage	
Completeness			information	Probability that all stream data sets contain the defined values and are updated in their defined frequency	
Communication		Network Performance			
		Packet Loss		technical	Ratio / Error Rate
		Bandwidth (Bitrate)		technical	Bits per second
		Latency		technical	(milli, micro)seconds
		Jitter		technical	(Milli)Seconds
		Throughput		technical	Bits per second
		Queuing			
		Queuing Type		technical	Queue Type
		Ordered		technical	Probability that data sets arrive in the defined queuing order
		Cost	Energy Consumption		technical
Monetary Consumption				operational	Defined per information or per operating time
Network Consumption			technical	Defined per information or per operating time	
Security	Confidentiality (reuse of rights ontology, e.g. <a href="http://creativecommons.org/ns">http://creativecommons.org/ns</a> )				
	licence def.		operational	Reference to Licence class, e.g. <a href="http://creativecommons.org/ns#Licence">http://creativecommons.org/ns#Licence</a>	
	may be used		operational	Reference to Permission class, e.g. <a href="http://creativecommons.org/ns#Permission">http://creativecommons.org/ns#Permission</a>	
	may be published		operational	Reference to Permission class, e.g. <a href="http://creativecommons.org/ns#Permission">http://creativecommons.org/ns#Permission</a>	
	Encryption		operational	Encryption method, authority for key management	
	Signing		operational	Encryption method, authority for key management	
Timeliness	Age		information	maximum time between measurement and publication	
	Volatility		information	Average Duration how long the information is usable, measured in seconds	
	Frequency		technical	Maximum timespan between two data sets	
GIN				communication	

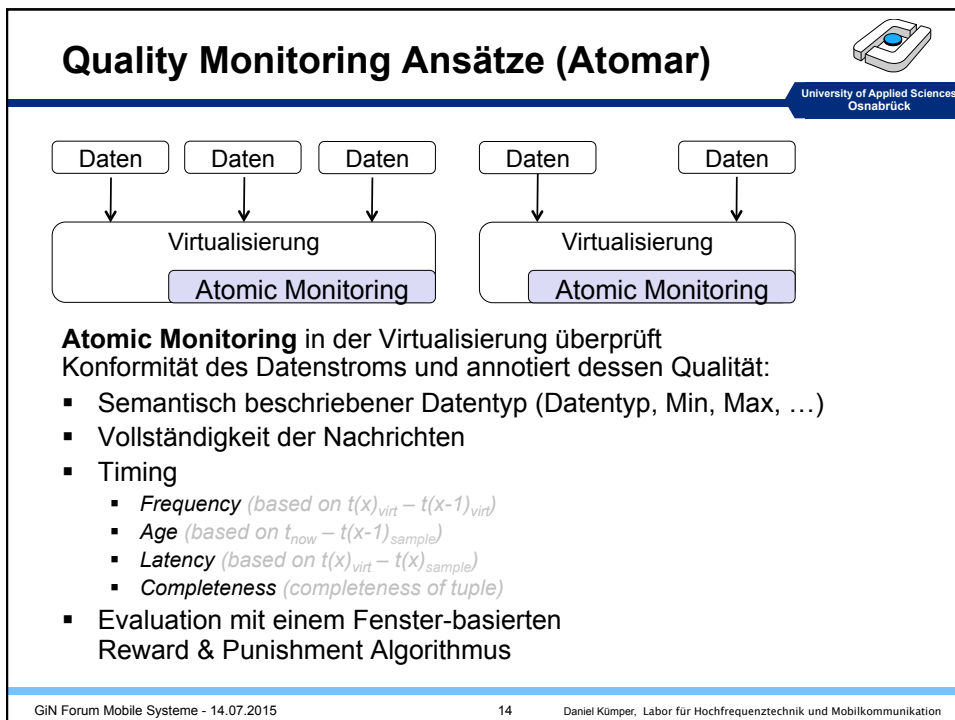
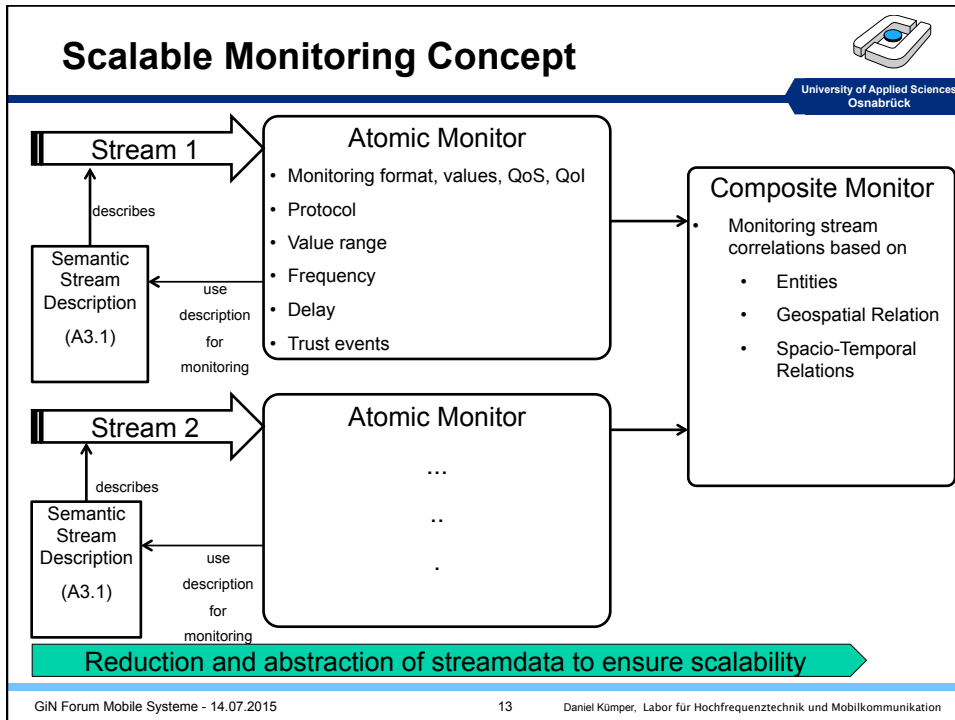
## Problem: Unzuverlässige Daten




University of Applied Sciences  
Osnabrück

- Unzuverlässige Datenverfügbarkeit
- Veraltete Daten
- Ungenaue Daten
- Widersprüchliche Daten
  - Einzelne Datenquellen verbreiten fehlerhafte Informationen und können nicht einfach erkannt werden.
- *Beispiel:*
  - *Bus Fahrplan gibt an, dass der Bus pünktlich ist.*
  - *Einzelne Benutzer geben an, dass es eine Verspätung gibt.*
- Herkunft der Daten (Reputation)
  - Eigene Infrastruktur
  - 3rd Party Provider
  - Einzeluser / mehrere User
  - Crowdsourcing

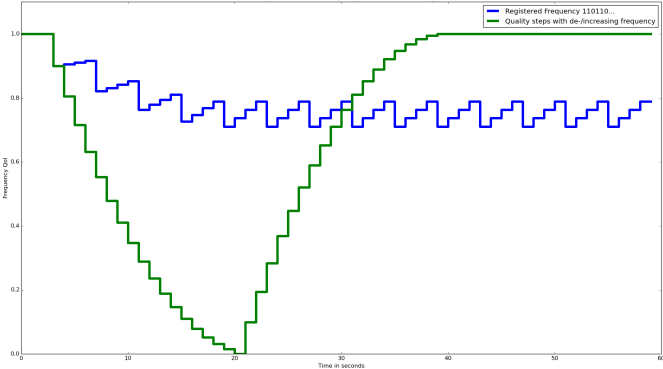


## Quality Monitoring Ansätze




University of Applied Sciences  
Osnabrück

- Reward and Punishment [3]
  - Wenn Qualitätsmetriken nicht eingehalten werden wird dessen Qualitätswert/Reputation heruntergesetzt (bestraft)
  - Quality sinkt/steigt logarithmisch

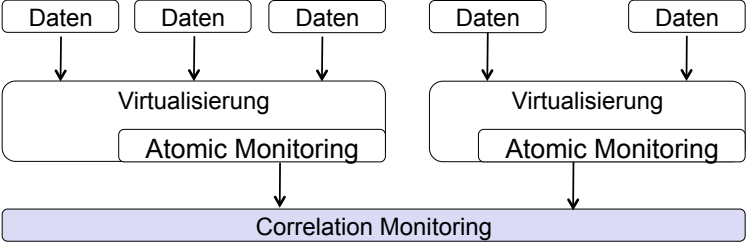


GiN Forum Mobile Systeme - 14.07.2015
15
Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

## Quality Monitoring Ansätze (Correlation)



University of Applied Sciences  
Osnabrück



**Correlation Monitoring** überprüft ob aktuelle Events durch Information aus individuellen Datenströmen nachvollzogen werden können.

- Validierung von Staumeldungen durch Verkehrsflusssensoren an zuführenden Straßen
- Zusammenhang zwischen Luftverschmutzung und Verkehr
- Social Media/Twitter Meldungen
- Veranstaltungskalender, Parkhausauslastung,...

GiN Forum Mobile Systeme - 14.07.2015
16
Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation



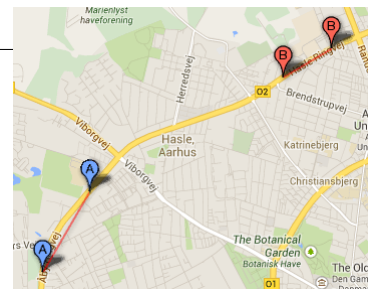
## Raumzeit (Spatio-Temporal) Einfluss



University of Applied Sciences  
 Osnabrück

Bei der Betrachtung der  
 Abhängigkeiten ist auch die Zeitliche  
 Verschiebung zu betrachten.

Time Elapsed	Relative Entropie/ Mutual Information	Pearson Korrelation
	Average MI daily windows	Average PC daily windows
-5min	1.0899	.4624
0min	1.1317	.4938
+5min	<b>1.1480</b>	<b>.5195</b>
+10min	1.1099	.4963
+15min	1.0900	.4740



GiN Forum Mobile Systeme - 14.07.2015

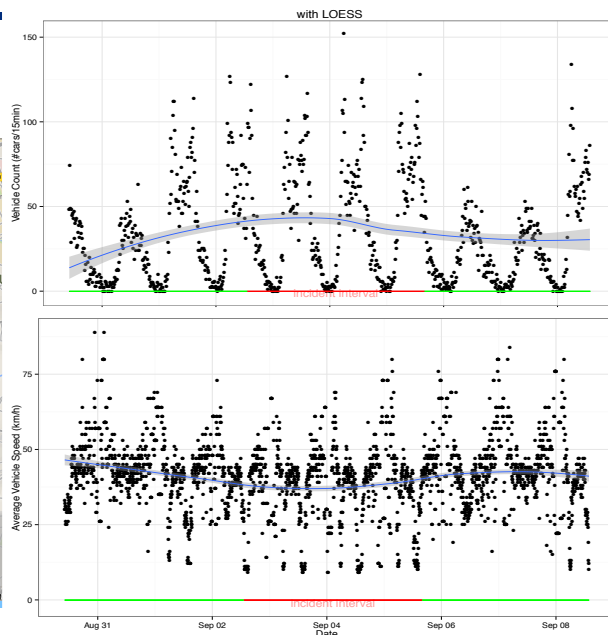
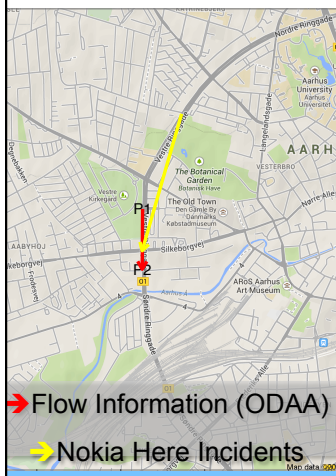
19

Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

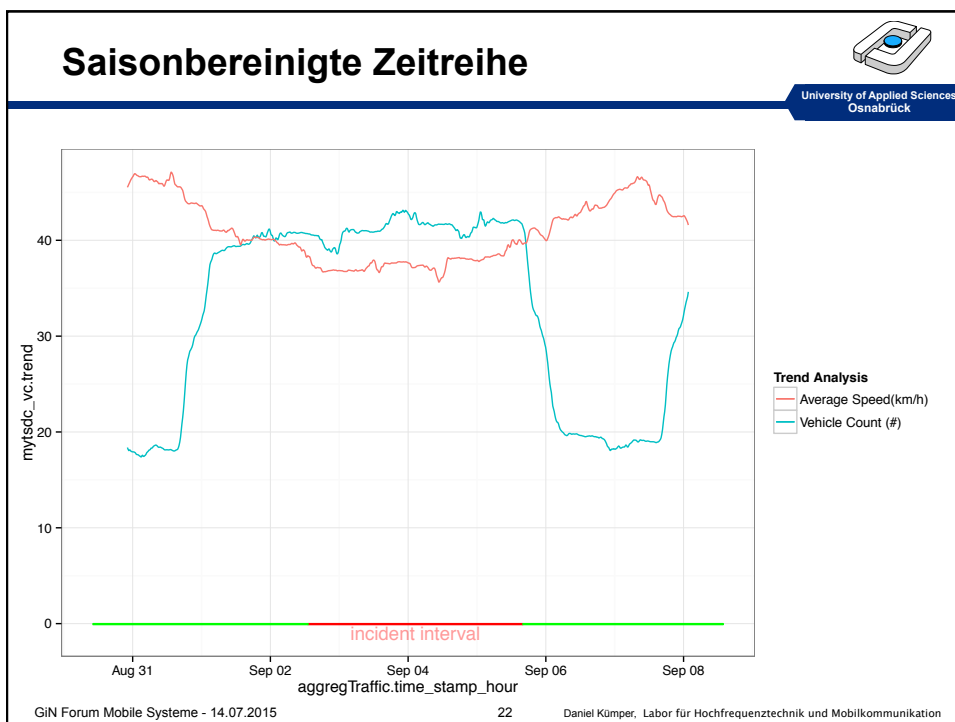
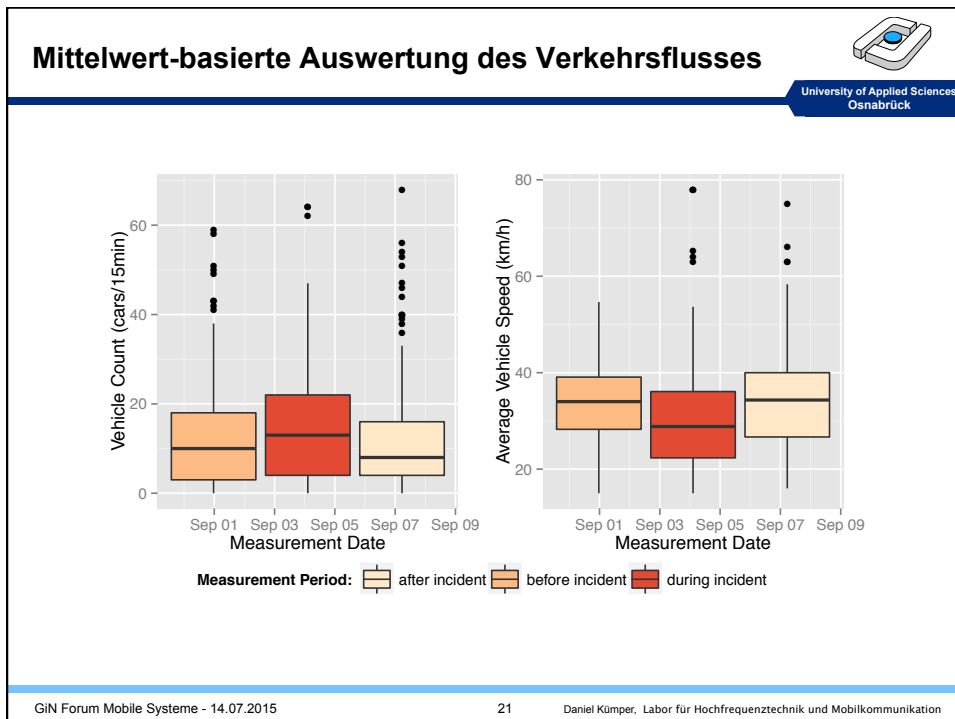
## Evaluation von Verkehrsmeldungen durch Zeitreihenanalyse (Multiple Stream, Multiple Feature, Multiple Sources)

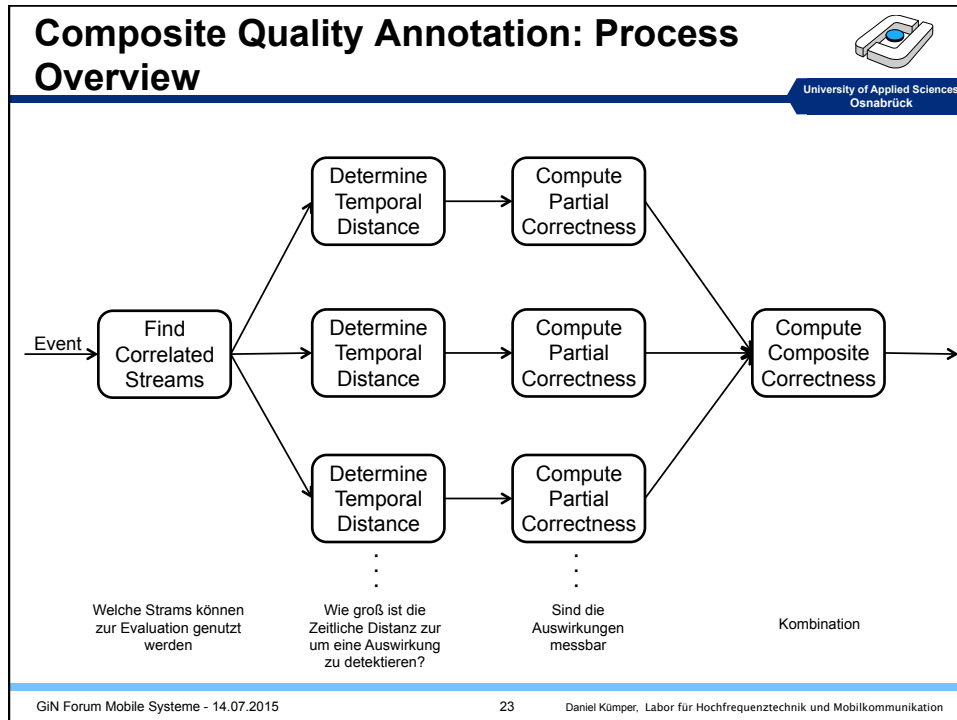


ODAA Report\_ID: 181034  
 Street: Vestre Ringgade  
 ODAA Distance: 447M



GIN Forum Mobile Systeme - 14.07.2015






### Inhalt

1. Einleitung
2. CityPulse Projekt
3. Beurteilung von Datenqualität / Quality of Information(QoI)
4. **Distanz und Interpolation in der Stadt**
5. Testansätze
6. Zusammenfassung/Quellen

GiN Forum Mobile Systeme - 14.07.2015 24 Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

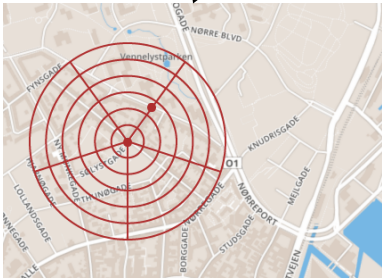



## Modell zur Distanzbestimmung und Interpolation




University of Applied Sciences  
Osnabrück

	Distance	Sight	Way	Track/Vehicle
<b>Propagation</b>	Radial, direkte Distanz	Radiale Distanz mit Abschattung	Ebene definiert durch einen Graph	Eingeschränkte Ebene durch Graph
<b>Example</b>	Luftverschmutzung	Licht	Straßen System	U-Bahn / Zug
<b>Modelbildung</b>	Einfach	Sehr Komplex	Mittel	Mittel

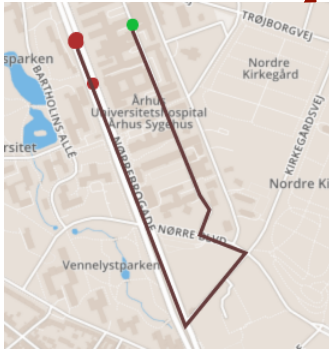
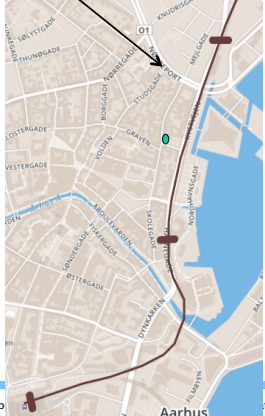
GiN Forum Mobile Systeme - 14.07.2015
25
Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

## Modell zur Distanzbestimmung und Interpolation




University of Applied Sciences  
Osnabrück

	Distance	Sight	Way	Track/Vehicle
<b>Propagation</b>	Radial, direkte Distanz	Radiale Distanz mit Abschattung	Ebene definiert durch einen Graph	Eingeschränkte Ebene durch Graph
<b>Example</b>	Luftverschmutzung	Licht	Straßen System	U-Bahn / Zug
<b>Modelbildung</b>	Einfach	Sehr Komplex	Mittel	Mittel

GIN Forum Mobile Systeme - 14.07.2015
26
Daniel Kümper, Lab




University of Applied Sciences  
Osnabrück

## Inhalt

1. Einleitung
2. CityPulse Projekt
3. Beurteilung von Datenqualität / Quality of Information(QoI)
4. Distanz und Interpolation in der Stadt
5. **Testansätze**
6. Zusammenfassung/Quellen

GiN Forum Mobile Systeme - 14.07.2015
27
Daniel Kümpfer, Labor für Hochfrequenztechnik und Mobilkommunikation



University of Applied Sciences  
Osnabrück

## Testansätze

- Nutzung funktionaler **Schnittstellenbeschreibungen**  
 Beschreibung zur Testerzeugung der Framework Komponenten
- Beschreibung von **Testfällen in Business Prozessen**
- Verwendung von **Testdatensätzen** zur Validierung der  
 Funktionalität und Performance
- Datensätze **öffentlich verfügbar**  
 über die Projektwebseite

**CITYPULSE**

Home

**Datasets**

Model Primer

Software

Evaluation (KPI)

Contact


Description	Duration	Location (Provider)	Type
<a href="#">Vehicle Traffic Data</a>	2/2014 - 6/2014 8/2014 - 9/2014 10/2014 - 11/2014	Aarhus, Denmark (Open Data Aarhus)	Real
<a href="#">Pollution Data</a>	8/2014 - 10/2014	Aarhus, Denmark (Open Data Aarhus)	Generated
<a href="#">Weather Data</a>	2/2014 - 6/2014 8/2014 - 9/2014	Aarhus, Denmark (Open Data Aarhus)	Real
<a href="#">Cultural Event Data</a>	5/2014 - 1/2015	Aarhus, Denmark (Open Data Aarhus)	Real
<a href="#">Twitter Data</a>	8/2013 - 12/2013 10/2014 - 12/2014	Aarhus, Denmark (Twitter)	Real
<a href="#">Social Event Data</a>	6/2012 - 6/2014	Surrey, UK (Municipality RSS)	Real
<a href="#">Library Event Data</a>	10/2013 - 6/2015	Aarhus, Denmark	Real
<a href="#">Parking Data</a>	5/2014 - 11/2014	Aarhus, Denmark	Real
<a href="#">Meeting Rooms Data</a>	Ongoing (live)	Ericksson Research, Kista	Real

The next table allows for a quick download of all Aarhus datasets (raw format only) and can be used for lampson combination. Clicking on the description provides information on how to retrieve the annotated datasets as well as more information on the datasets themselves.

Description	Duration																			
	2013			2014						2015										
	9	10	11	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5
<a href="#">Vehicle Traffic Data</a>				batch 1	batch 2															
<a href="#">Pollution Data</a>																				
<a href="#">Weather Data</a>				batch 1	batch 2															
<a href="#">Parking Data</a>																				
<a href="#">Cultural Event Data</a>																				
<a href="#">Library Event Data</a>																				
<a href="#">Twitter Event Data</a>				batch 1											batch 2					

GiN Forum Mobile Systeme - 14.07.2015
28

## Beispiel: BPMN getriebene Testfälle

  
University of Applied Sciences  
Osnabrück

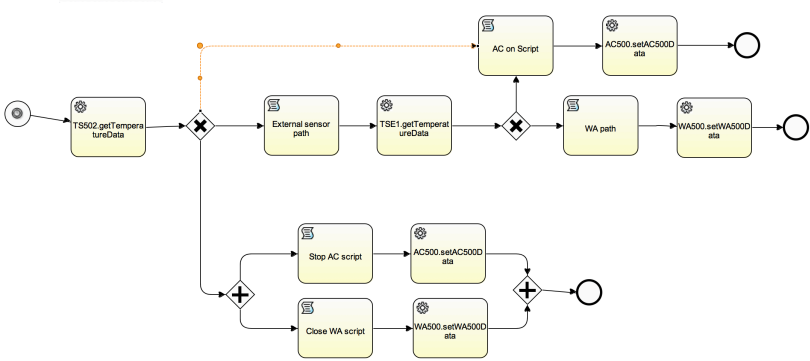
Outline Variables Breakpoints Expressions Debug Properties

General Label width (50-500)


Main config Condition

```
return ((int) ((SmartSensorStatus) contentOutput).getSensorValue()) > 30;
```

Package Explorer Servers TempRegulator



## Inhalt

  
University of Applied Sciences  
Osnabrück

1. Einleitung
2. CityPulse Projekt
3. Beurteilung von Datenqualität / Quality of Information(QoI)
4. Distanz und Interpolation in der Stadt
5. Testansätze
6. **Zusammenfassung/Quellen**

GIN Forum Mobile Systeme - 14.07.2015 30 Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

## Zusammenfassung



University of Applied Sciences  
Osnabrück

Was wird benötigt um ein zuverlässiges, offenes Framework für Smart City Applikationen bereitzustellen?

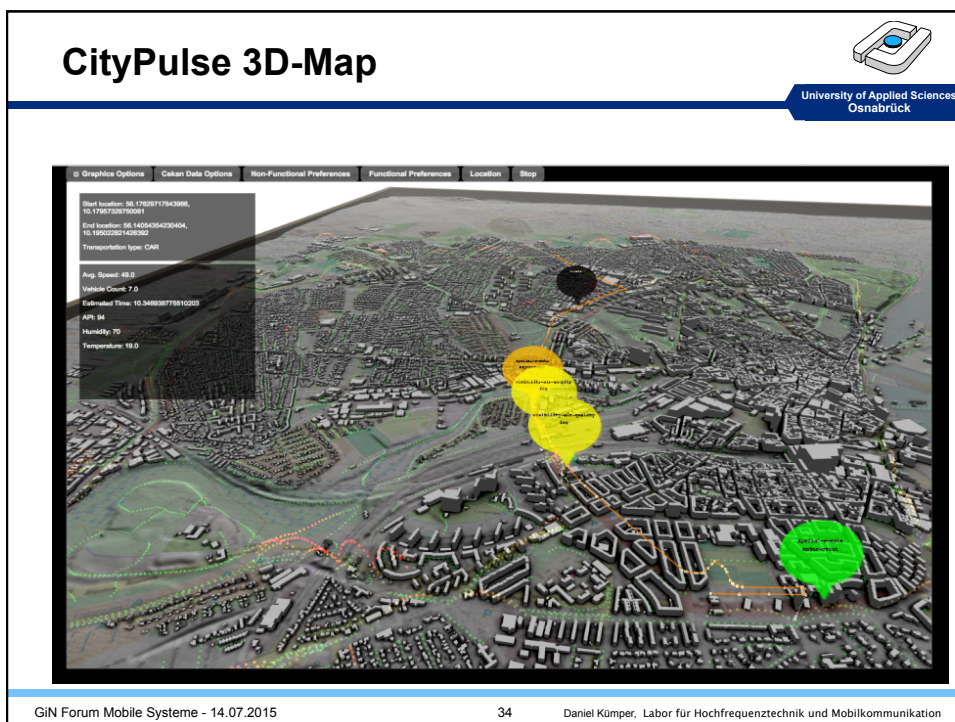
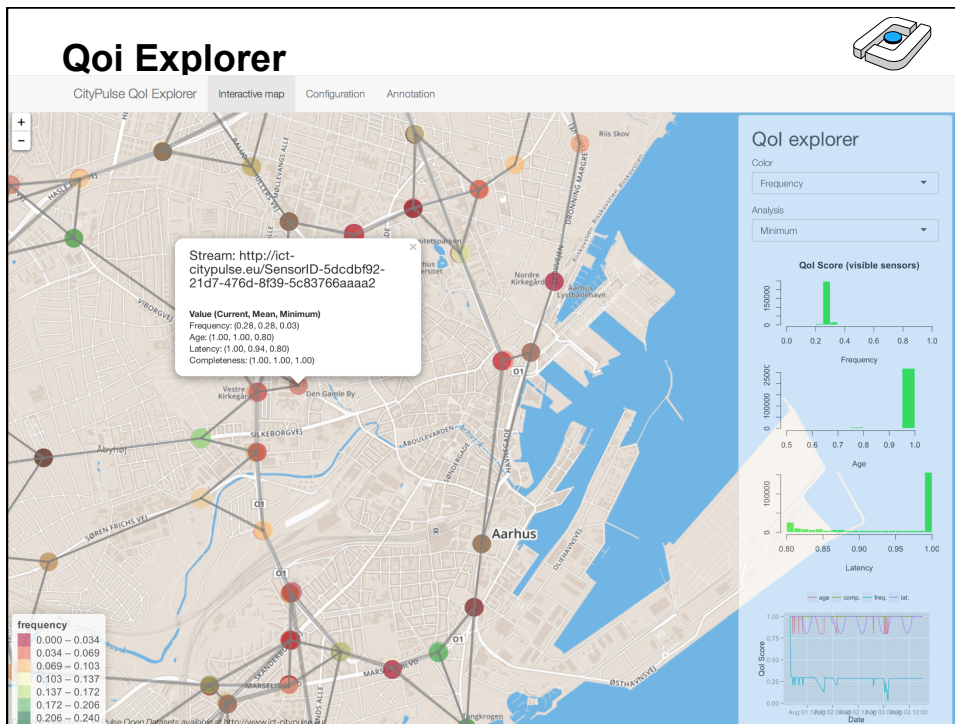
- ❑ **Daten Bedeutung verleihen**
  - Semantische Annotierung zur automatisierten Verarbeitung
- ❑ Berücksichtigung der **Zuverlässigkeit und Genauigkeit**
  - Modellierung und Verarbeitung von Vertrauenswürdigkeit und Genauigkeit
- ❑ Offene **Datenverfügbarkeit und Referenzdatensätze** zur Evaluation
  - Freier Zugriff auf öffentliche Information ermutigt Entwickler
  - Referenzdatensätze ermöglichen den Vergleich von Lösungen
- ❑ **Skalierbare Ansätze**

## Quellen




University of Applied Sciences  
Osnabrück

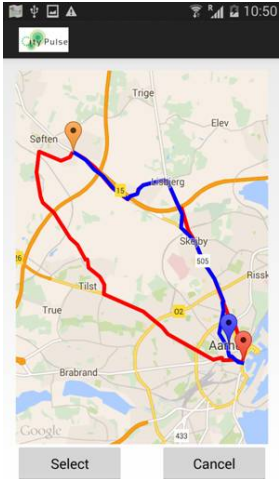
- (1) <http://wikibon.org/blog/big-data-statistics/>
- (2) <http://expandedramblings.com/index.php/march-2013-by-the-numbers-a-few-amazing-twitter-stats/10/>
- (3) Hossain, M. Anwar et al. "Modeling and assessing quality of information in multisensor multimedia monitoring systems." ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP) 7.1 (2011): 3. page 11.
- (4) CityPulse D2.3: Reference Data Sets and Benchmark Measures.
- (5) CityPulse D4.1 :Measures and Methods for Reliable Information Processing.

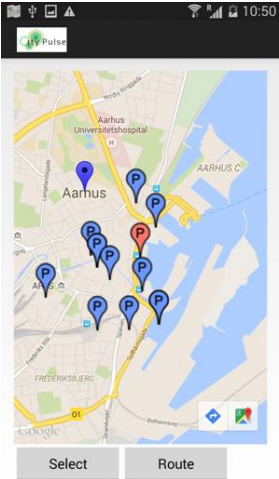


## Mobile Anwendungen




University of Applied Sciences  
Osnabrück



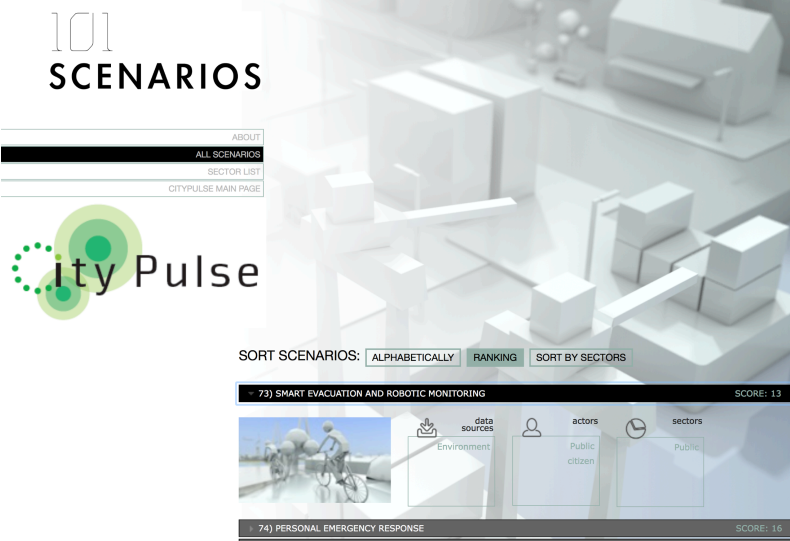


GIN Forum Mobile Systeme - 14.07.2015
35
Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation

## 101 Scenarios



University of Applied Sciences  
Osnabrück




GIN Forum Mobile Systeme - 14.07.2015
36
Daniel Kümper, Labor für Hochfrequenztechnik und Mobilkommunikation






• **Vielen Dank!**

• **EU FP7 CityPulse Project:**

 <http://www.ict-citypulse.eu/>

 @ictcitypulse

 d.kuemper@hs-osnabrueck.de

```
status,avgMeasuredTime,avgSpeed,extID,medianMeasuredTime,TIME
STAMP,vehicleCount_id,REPORT_ID
OK,92,91,646,92,2014-02-13T11:30:00,35,189978,158446
OK,84,100,646,84,2014-02-13T11:35:00,26,190427,158446
OK,78,108,646,78,2014-02-13T11:40:00,31,190876,158446
OK,79,106,646,79,2014-02-13T11:45:00,28,191325,158446
```