



**Big Data in der Praxis:
Smart Traffic und Visualisierung von
Einbruchsdaten**

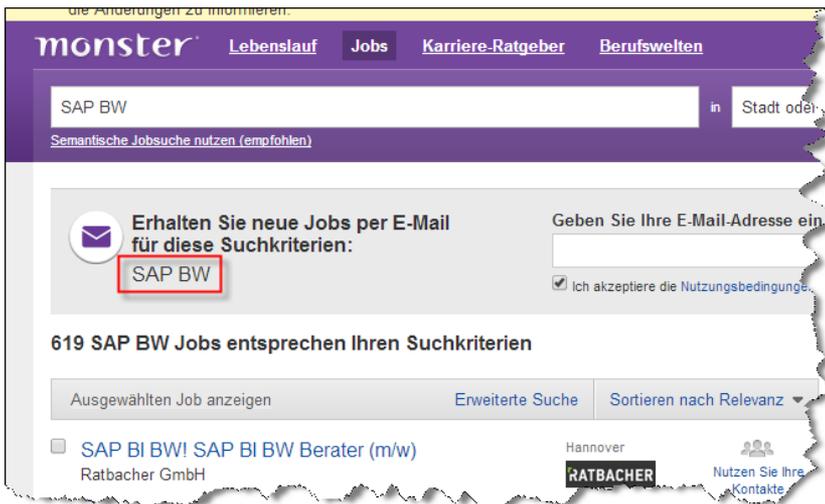
**Prof. Dr. Klaus Freyburger, Hochschule
Ludwigshafen am Rhein**



- **Prof. Dr. Klaus Freyburger**
- **1982 - 1989 Studium der Mathematik und Betriebswirtschaftslehre an der Universität Mannheim und University of Massachusetts at Amherst / USA**
- **1991 Promotion in Mathematik**
- **1991 - 2002 beschäftigt bei der SAP SE in Walldorf
Tätigkeitsschwerpunkt: Entwicklung von Software zur Unternehmensplanung**
- **Seit 2002 Professur für Wirtschaftsinformatik an der Hochschule Ludwigshafen am Rhein**
- **Lehre und Forschung:**
 - Programmierung
 - Business Intelligence / Big Data / Analytics mit SAP, Microsoft und Open Source (z.B. Python, R)

- **4.300 Studierende**
- **87 Hauptamtlich Lehrende**
- **250 Lehrbeauftragte**
- **Große Auswahl an Studiengängen der Betriebswirtschaftslehre und der Sozialen Arbeit**
 - **Bachelor Wirtschaftsinformatik**
 - **Master Wirtschaftsinformatik (Information Management & Consulting)**
 - **Dualer Bachelor Wirtschaftsinformatik**
- **Angewandt, praxisorientiert, innovativ und international**

- Obwohl Hochschulen zunächst auf Unabhängigkeit gegenüber den großen Softwareherstellern bedacht sein sollten, ...
- ... kann die steigende Nachfrage nach Absolventen mit Kenntnissen über Architektur und Anwendung einer führenden Standardsoftware im Sinne einer praxisgerechten Ausbildung nicht einfach ignoriert werden.



die Änderungen zu informieren.

monster [Lebenslauf](#) [Jobs](#) [Karriere-Ratgeber](#) [Berufswelten](#)

SAP BW in Stadt oder...

Semantische Jobsuche nutzen (empfohlen)

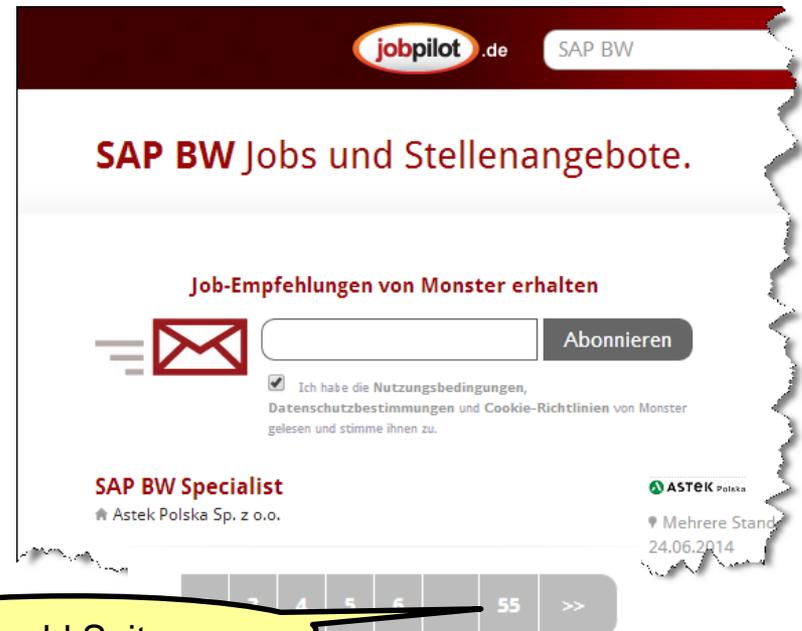
 Erhalten Sie neue Jobs per E-Mail für diese Suchkriterien: Geben Sie Ihre E-Mail-Adresse ein

SAP BW Ich akzeptiere die Nutzungsbedingungen

619 SAP BW Jobs entsprechen Ihren Suchkriterien

[Ausgewählten Job anzeigen](#) [Erweiterte Suche](#) [Sortieren nach Relevanz](#)

SAP BI BW! SAP BI BW Berater (m/w) Hannover  Nutzen Sie Ihre Kontakte
Ratbacher GmbH



jobpilot .de SAP BW

SAP BW Jobs und Stellenangebote.

Job-Empfehlungen von Monster erhalten

 [Abonnieren](#)

Ich habe die Nutzungsbedingungen, Datenschutzbestimmungen und Cookie-Richtlinien von Monster gelesen und stimme ihnen zu.

SAP BW Specialist
★ Astek Polska Sp. z o.o.

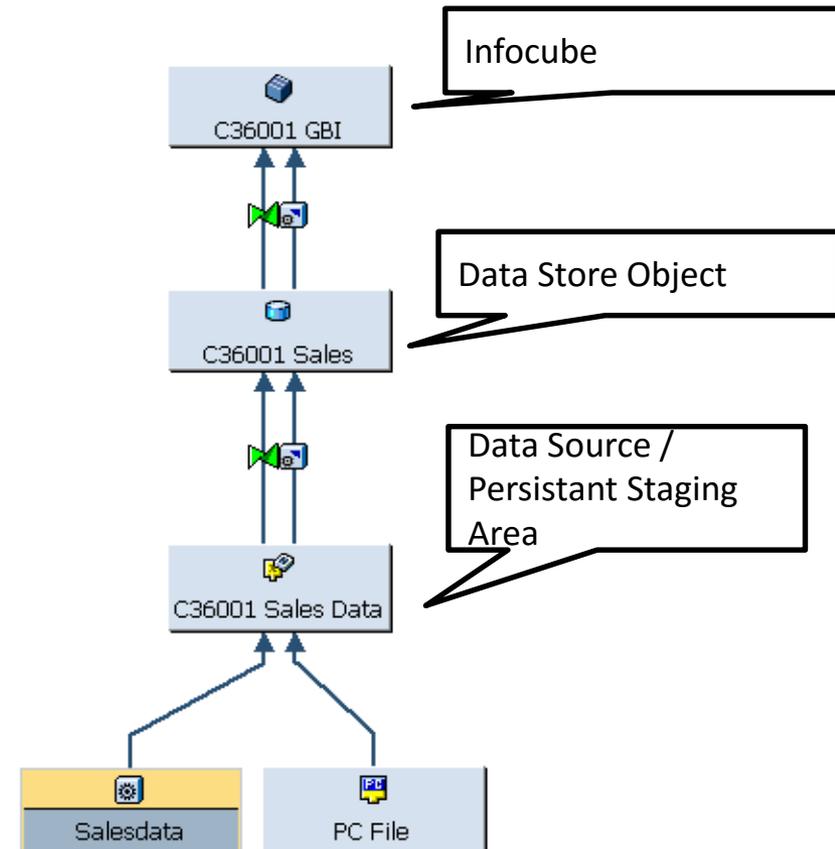
 Mehrere Standorte
24.06.2014

55 >>

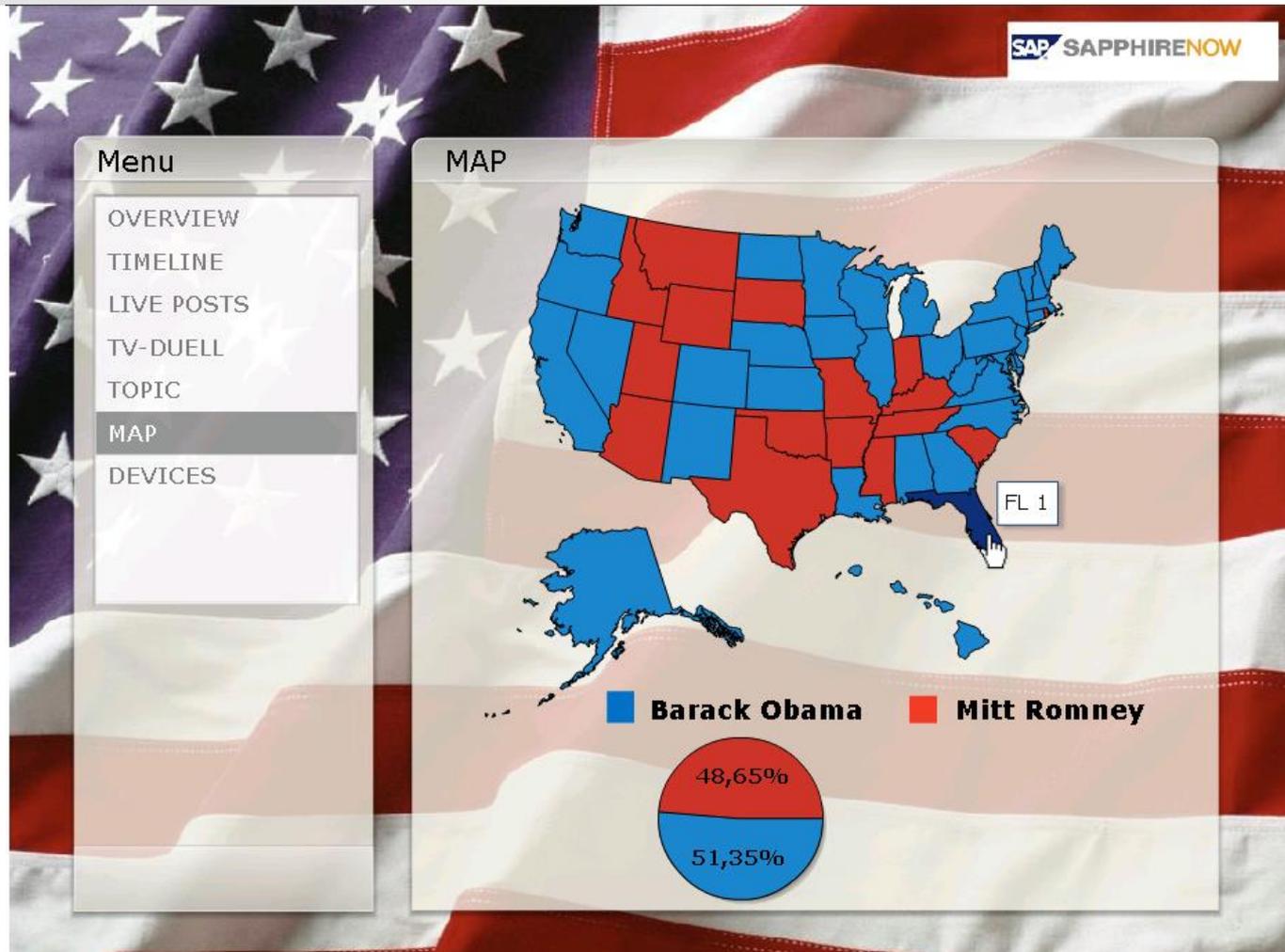
Anzahl Seiten

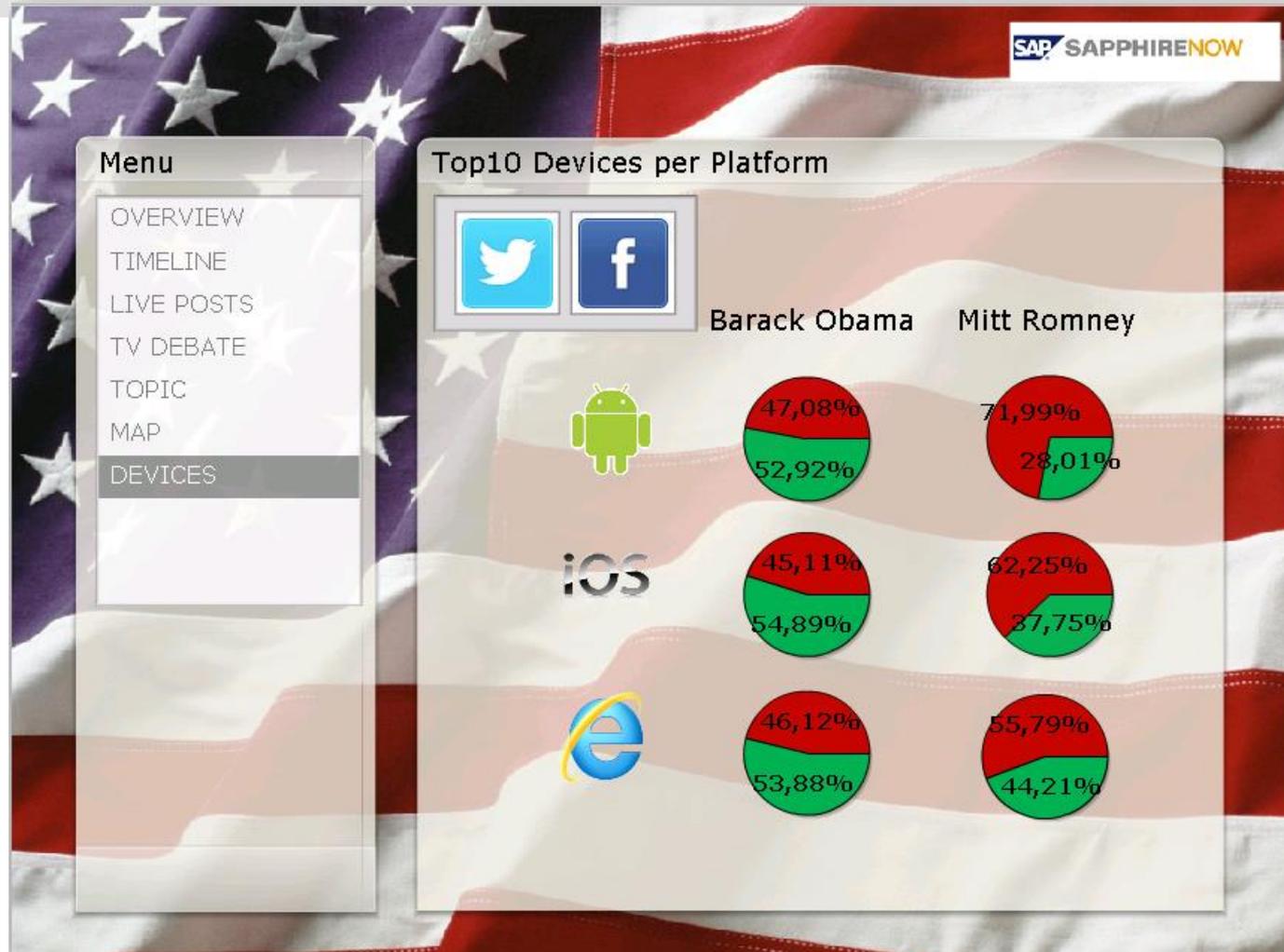
„Brot und Butter Geschäft“: Jeder Studierende baut einen Case im SAP BW System

	Revenue	Count	Revenue %
Product	\$		%
Professional Road Bike (Shimano)	7.377.321,90	1	13
Men's Off Road Bike Fully	6.133.319,81	2	24
Professional Touring Bike (silver)	5.857.365,23	3	34
Deluxe Touring Bike (silver)	5.169.373,07	4	44
Deluxe Road Bike (Shimano)	4.931.751,51	5	52
Men's Off Road Bike Hard Tail (SRAM)	4.655.870,79	6	61
Men's Off Road Bike Hard Tail (Shimano)	2.855.490,94	7	66
Professional Touring Bike (black)	2.783.657,95	8	71
Women's Off Road Bike Fully	2.721.869,85	9	76
Professional Touring Bike (red)	2.535.383,94	10	80
Deluxe Touring Bike (black)	2.405.581,35	11	84
Deluxe Touring Bike (red)	2.312.287,53	12	89
E-Bike Tailwind	1.998.759,41	13	92
Professional Road Bike (Campagnolo)	1.584.721,97	14	95
Deluxe Road Bike (SRAM)	1.396.619,89	15	97
Professional Road Bike (SRAM)	1.334.643,08	16	100
City Bike Max	101.413,19	17	100
Fixed Gear Bike Plus	19.231,99	18	100
Overall Result	56.174.663,40	18	100









- **SAP HANA for Humanity: How can SAP HANA contribute to improve the world a bit**
- **People should be encouraged to capture their dreams**



- **Challenges**
 - **GUI**
 - **Textanalysis**
 - **Analytics**
- **Design Thinking**
- **Launch / Go-Live with students at Sapphire 2013 in Orlando / Florida**



- **Projekte wurden angefertigt im WS15/16 im Rahmen der Veranstaltung „Realisierung Business Intelligence“**
- **3. Semester Master**
- **Innerhalb von 12 Wochen**
- **5 Credits entspricht Workload von 150 Stunden**
- **Gruppe zu je 3- 4 Studierende bearbeiten eine Aufgabenstellung**
- **Leistungsnachweis**
 - 2 Meilensteine
 - Abschlusspräsentation
 - Projektbericht

Projekt 1: Visualisierung von Einbruchdaten

In Zusammenarbeit mit mayato



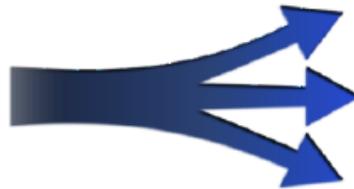
Aufgabe:

- **Visualisierung der in Deutschland in den letzten Jahren erfolgten Einbrüche. Dazu müssen die in Fließtext auf Webservern vorliegenden Informationen (z.B. <http://www.presseportal.de/blaulicht/>) extrahiert, sinnvoll gefiltert und in einer strukturierten Datenbank abgelegt werden.**
- **Diese Daten sollen dann in geeigneter Form visualisiert werden. So soll z.B. mit einer Heatmap die geographische Verteilung von Einbrüchen in verschiedenen Stadtteilen dargestellt werden.**
- **Mögliche Tools (beispielhaft):**
 - Visualisierung: Tableau oder SAP Lumira
 - Datenbank: Exasol oder SAP HANA
 - Extraktion: Talend oder SAP Data Services

Bearbeitet durch 2 Gruppen mit je 4 Studierenden



Presseportal.de



ETL



Database



Visualisierung



Talend Open Studio for Big Data (kostenlos)

Exasol Free Edition

Tableau (Vollversion über Academic Program)

POLIZEIPRÄSIDIUM MANNHEIM

POL-MA: Mannheim: Einbruch in der Draisstraße

05.02.2016 – 09:47

Mannheim (ots) - Eingebrochen wurde Freitagnacht in eine Wohnung in der **Draisstraße**. Zwischen 2.30 und 4.30 Uhr hebelte ein bislang Unbekannte die Wohnungstüre auf und durchsuchte sämtliche Schränke und Kommoden. Dem Einbrecher fielen Bargeld, Schmuck sowie ein Handy im Gesamtwert von mehreren tausend Euro in die Hände. Der Sachschaden an der Türe schlägt mit weiteren hundert Euro zu Buche. Zeugen, die Hinweise hierzu geben können, werden gebeten, sich beim Polizeirevier Mannheim-Neckarstadt, Tel.: 0621/3301-0, zu melden.

Rückfragen bitte an:

Polizeipräsidium Mannheim

Öffentlichkeitsarbeit

Ulrike Mathes

Telefon: 0621 174-1106

E-Mail: mannheim.pp.stab.oe@polizei.bwl.de

<http://www.polizei-bw.de/>

POLIZEIPRÄSIDIUM MANNHEIM

POL-MA: Schwetzingen/Rhein-Neckar-Kreis: Bei Einbruch in Apotheke Bargeld erbeutet - Zeugen gesucht

05.02.2016 – 13:35

Schwetzingen/Rhein-Neckar-Kreis (ots) - Mehrere hundert Euro Bargeld erbeutete ein bislang unbekannter Täter bei einem Einbruch in eine Apotheke in der **Carl-Benz-Straße**. In der Nacht von Donnerstag auf Freitag wurde die Eingangstüre aufgehebelt und auseinander geschoben. Der Täter hatte es offenbar nur auf Bargeld abgesehen und nahm das Wechselgeld aus zwei Kassen. Das Polizeirevier Schwetzingen nahm die Spurensicherung vor. Zeugen, die im fraglichen Zeitraum verdächtige Beobachtungen gemacht haben werden gebeten sich unter Telefon 06202/2880 zu melden.

Rückfragen bitte an:

Polizeipräsidium Mannheim

Öffentlichkeitsarbeit

Dieter Klumpp

Telefon: 0621 174-1105

E-Mail: mannheim.pp@polizei.bwl.de

<http://www.polizei-bw.de/>

05.02.2016 – 12:37

POLIZEIPRÄSIDIUM MANNHEIM

POL-MA: Hockenheim/Rhein-Neckar-Kreis: Firmeneinbrüche** in der Nacht zum Donnerstag - Polizei ermittelt und bittet um...**

Hockenheim/Rhein-Neckar-Kreis (ots) - In ein Firmengebäude in der Altriper Straße brachen in der Nacht zum Donnerstag bislang unbekannte Täter ein. Durch Einwerfen von Fenstern bzw. Aufbrechen von Türen verschafften sich die Täter Zugang und durchsuchten verschiedene Räumlichkeiten und Behältnisse beider ...

05.02.2016 – 10:45

POLIZEIPRÄSIDIUM MANNHEIM

POL-MA: Heidelberg-Wieblingen: **Versuchter Einbruch**

Mannheim (ots) - Im Sandwingert versuchte ein bislang unbekannter Täter am Donnerstag zwischen 12:30 Uhr und 18:50 Uhr, die Terrassentür eines Einfamilienhauses aufzuhebeln. Der Täter ließ jedoch aus unbekanntem Gründen von seinem Vorhaben ab. Zeugen, denen zu dieser Zeit etwas Ungewöhnliches aufgefallen ist, werden gebeten, sich beim Kriminaldauerdienst, Tel.: 0621/174-5555, zu melden. ...

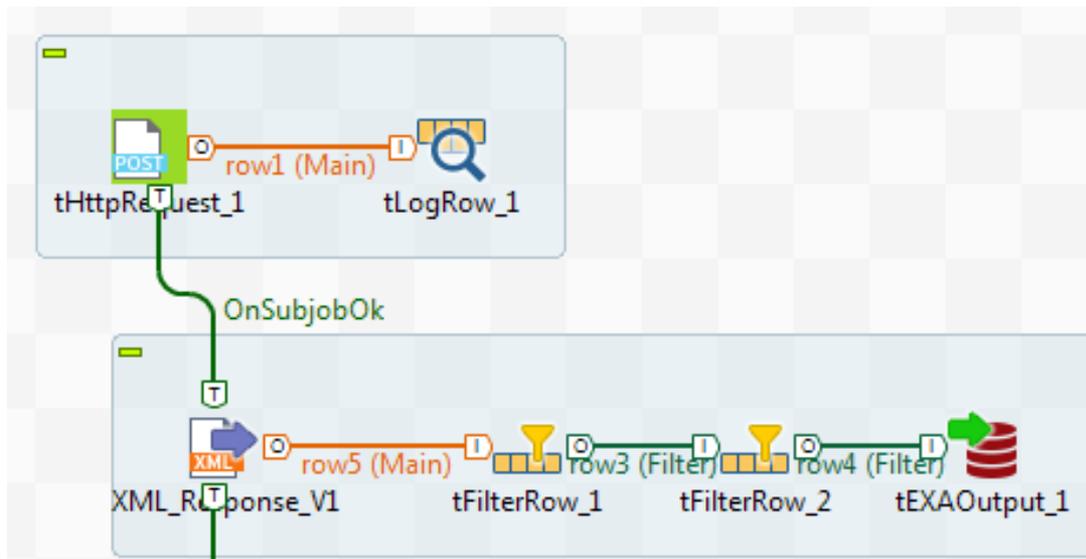
Ungenaue Ortsangabe

POLIZEIPRÄSIDIUM MANNHEIM

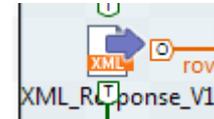
**POL-MA: Leimen: Einbruch in Clubhaus und
Leistungszentrum - Möglicherweise besteht
Tatzusammenhang - Polizei sucht Zeugen**

04.02.2016 – 12:36

Leimen/Rhein-Neckar-Kreis (ots) - Leimen: In die Gaststätte des Tennisclubs s...



XML Parser Pflege durch Wizard



Datei - Schritt 4 von 5

Add a Metadata File on repository
Define the setting of the parse job



Quell-Schema

- office
 - id
 - url
 - name
 - logo
- content
 - story
 - id
 - url
 - title
 - body
 - published
 - keywords
 - highlight
 - short

Ziel-Schema

Xpath loop expression

Absoluter XPath-Ausdruck	Loop li...
/response/content/story	50

Fields to extract

Relativer oder absoluter XPath-Ausdruck	Spaltenname
id	id
url	url
title	title
body	body
published	published

Filter durch Java Code

 **tFilterRow_1**

Basic settings Schema

Advanced settings Logischer Operator zur Verknüpfung der Bedingungen *

Dynamic settings Bedingungen

Input column	Funktion	Operator
--------------	----------	----------

Erweiterten Modus verwenden

Erweitert

```
input_row.title.contains("einbruch") ||
input_row.title.contains("Einbruch") ||
input_row.title.contains("einbrüche") ||
input_row.title.contains("Einbrüche") ||

input_row.body.contains("einbruch") ||
input_row.body.contains("Einbruch") ||
input_row.body.contains("einbrüche") ||
input_row.body.contains("Einbrüche")
```

Man benötigt Geoinformationen zur Mitteilung

- **Straßenverzeichnis von Mannheim** [http://rhein-neckar-wiki.de/Stra%C3%9Fenverzeichnis \(Mannheim\)](http://rhein-neckar-wiki.de/Stra%C3%9Fenverzeichnis_(Mannheim))
- **Vervollständigung mit Geoinformationen**
<http://www.findlatitudeandlongitude.com/batch-geocode>
- **Matching durch Javacode mit Stringsuche in SQL auf Exasol**

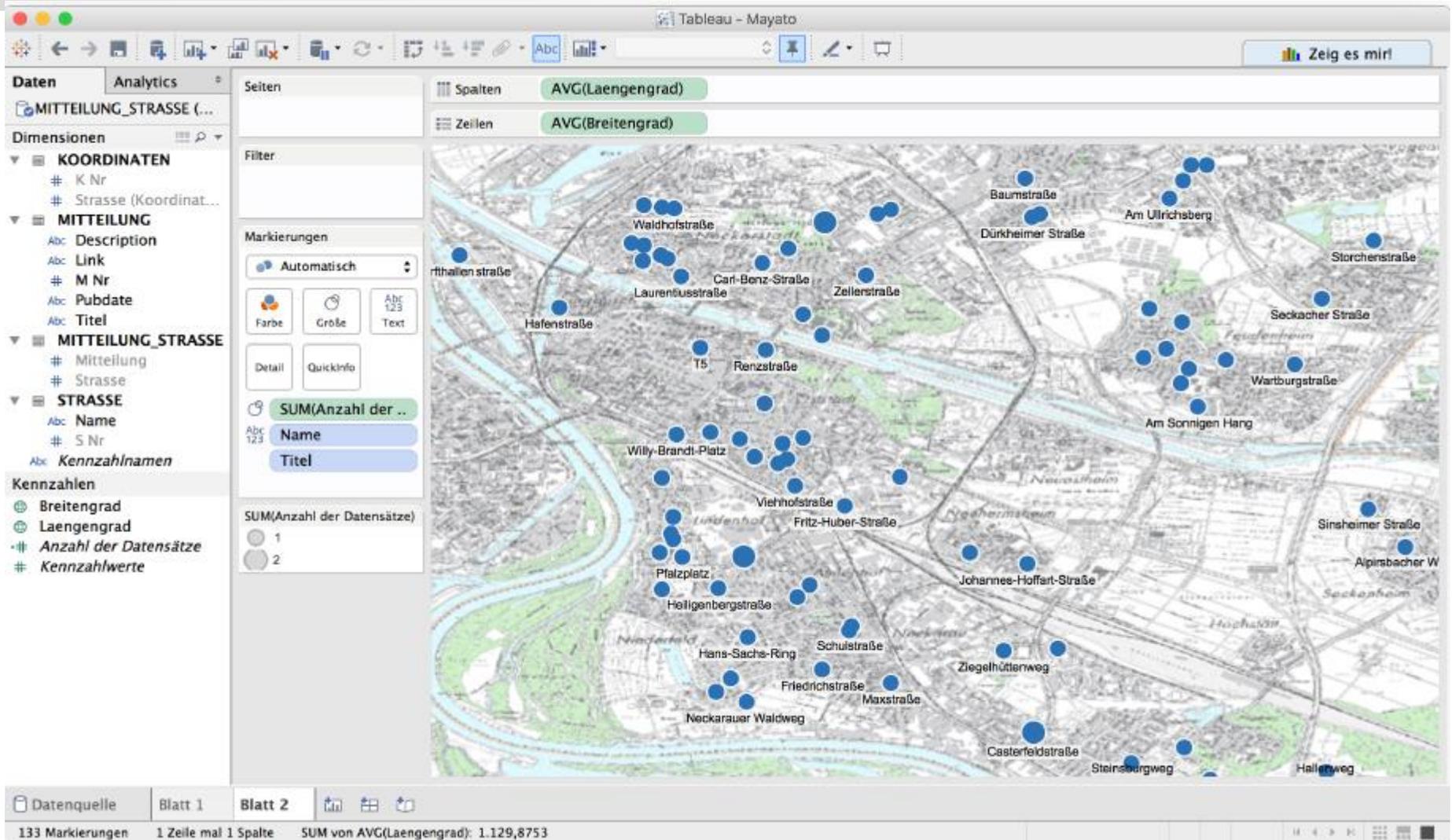
```
System.out.println("Creating statement...");
stmt = conn.createStatement();
String sql = "SELECT M_NR, S_NR, NAME FROM MITTEILUNG, STRASSE WHERE DESCRIPTION like CONCAT('%', NAME, '%)";
System.out.println("SQL: " +sql);
ResultSet rs = stmt.executeQuery(sql);
System.out.println("Resultset filled");
```

Straßen sollten angezeigt werden

- **OpenStreetMap**
 - Kein Kartenmaterial zu Straßen
- **Google**
 - kein API zur Einbindung in andere Produkte
 - Speicherung der Karten laut AGB unzulässig

Lösung:

- **WMS Kartendienst von Mannheim**
 - https://www.gis-mannheim.de/mannheim_service/mod_wms/wms_list.php
 - Einbindung in Tableau



Keine Beschränkung auf Mannheim

- **Problem: 50 Treffer pro Aufruf des APIs des Presseportals, Sperrung nach 60 Aufrufen**
- **Lösung: Loop in Talend, Verwenden mehrerer API Keys und Verwaltung in Kontexten von Talend**

Verschiedene Ansätze zur Textanalyse

- **Talend**
 - tReplace, tReplaceList (entfernen Sonderzeichen, Füllwörter)
- **R**
 - Pakete „NLP“ und „tm“ (Entfernen Sonderzeichen, Umwandeln in Kleinschreibung, entfernen Stoppwörter von vorgefertigter Liste, die mit eigenen Wörtern ergänzt wurde)

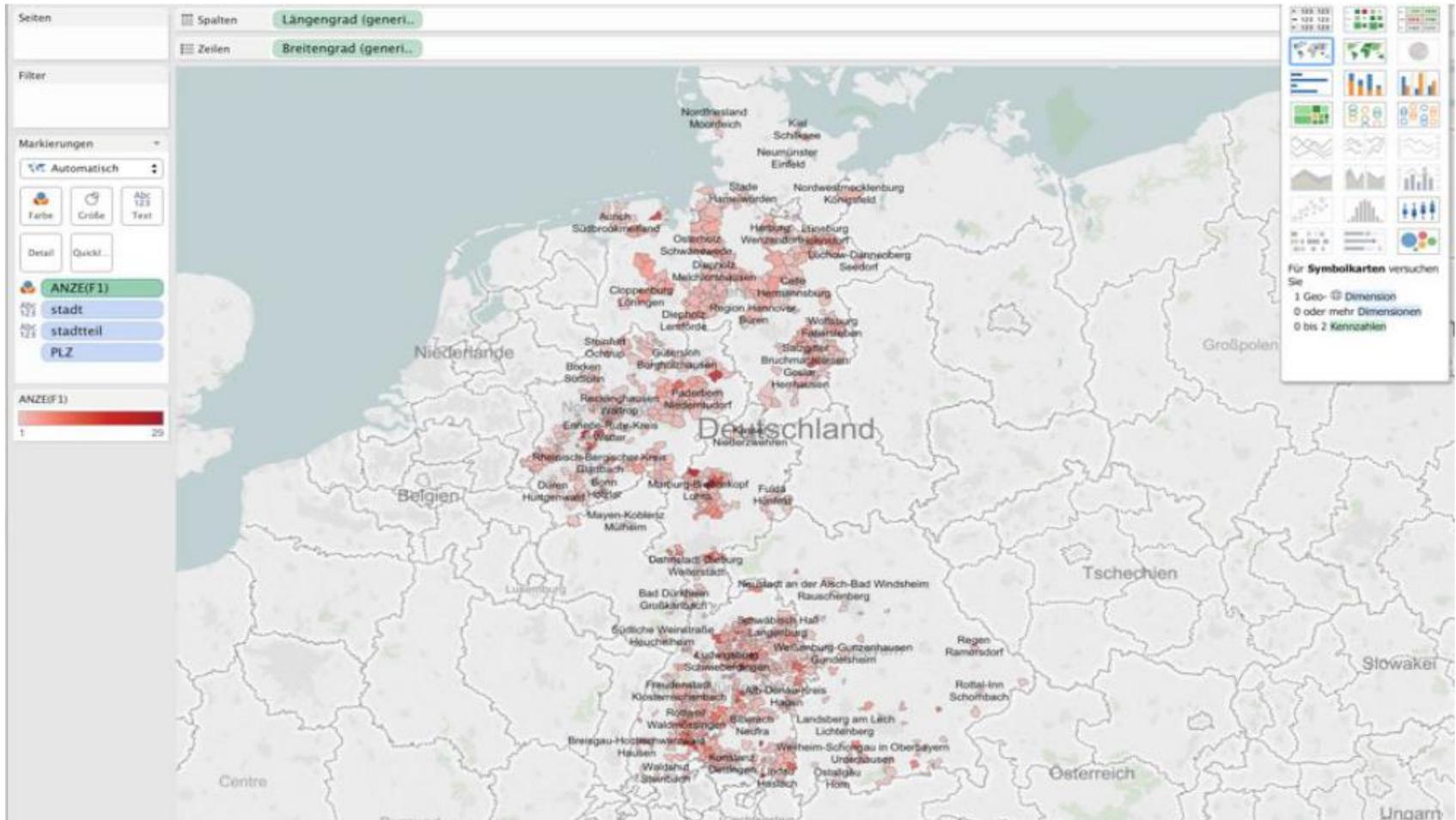
Datenspeicherung

- **Zum verteilten Arbeiten sollte Exasol Cloud verwendet werden**
 - Performanceprobleme beim Schreiben der Daten aus Talend → nicht gelöst → in der Folge mit CSV gearbeitet

- Datenbank mit Geo Koordinaten für ganz Deutschland
<http://opengeodb.giswiki.org/wiki/OpenGeoDB>
- Durch Joins entstand eine Tabelle mit Stadtteilen und Koordinaten

loc_id	stadtteil	PLZ	stadt	regierungsbezirk	bundesland	lat	lon
81559	Neckarstadt	68167	Mannheim	Karlsruhe	Baden-Württemberg	49.4883	8.46472
81559	Neckarstadt	68169	Mannheim	Karlsruhe	Baden-Württemberg	49.4883	8.46472
81560	Wohlgelegen	68167	Mannheim	Karlsruhe	Baden-Württemberg	49.4883	8.46472
81562	Schwetzinger...	68161	Mannheim	Karlsruhe	Baden-Württemberg	49.4883	8.46472

- **Matching in Talend**
 - tNormalize (Split der Meldung in einzelne Wörter)
 - tMap (LookUp)



SAP Traffic Control

Efficient Mobility for Sustainable Growth

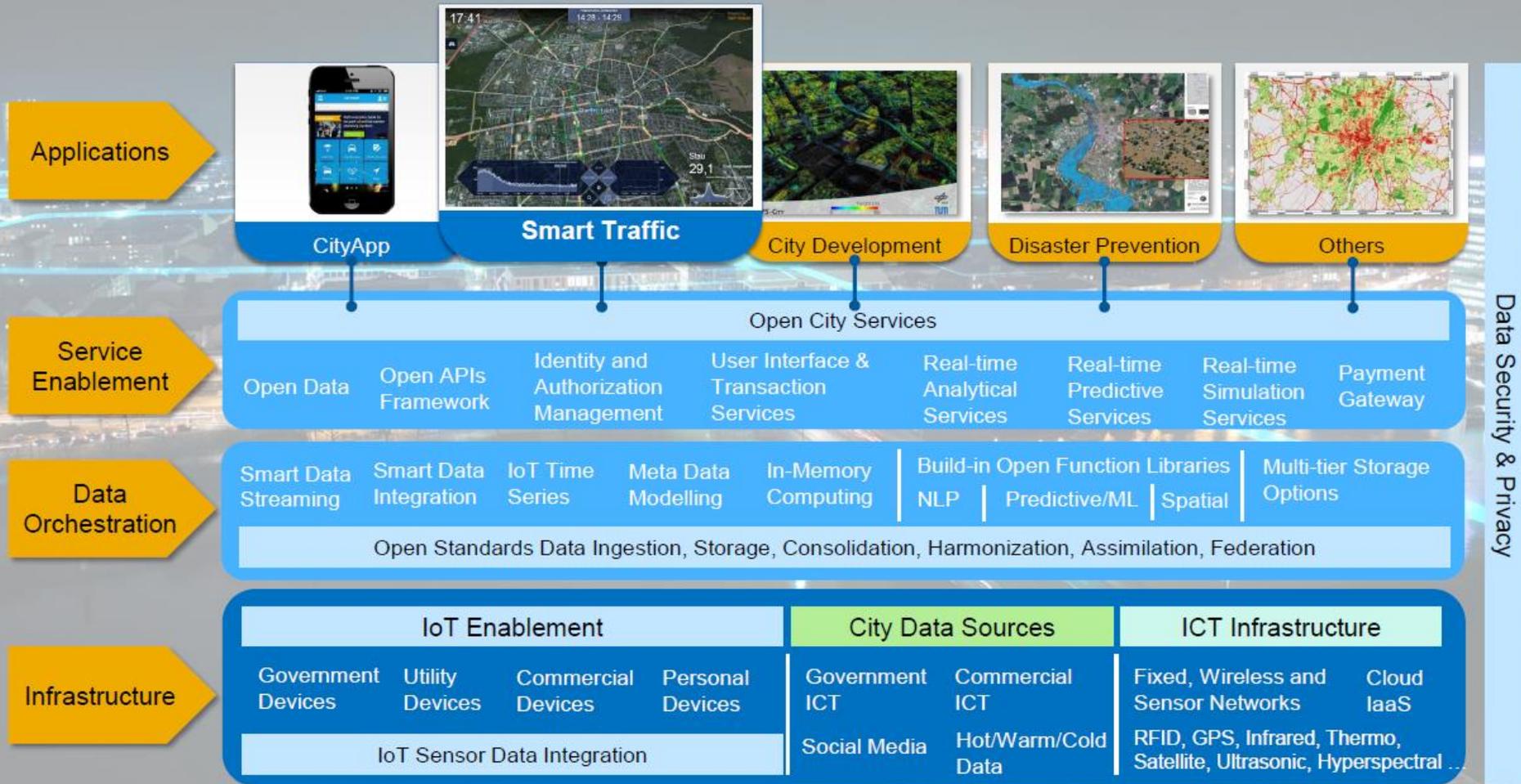
Innovation Center Network, Research & Innovation, SAP SE



Dr. Norbert Koppenhagen, SAP SE
ICN, Research & Innovation



Empowering the city eco-system with SAP HANA



Hauptziel:

- **Bessere Steuerung des Verkehrsflusses (durch Telematikmodelle)**

Unterziele:

- **Steigerung der Mobilität**
- **Vermeidung von Staubildung**
- **Reduzierung der Luftverschmutzung**
- **Optimale Ausnutzung gegebener Kapazitäten der Verkehrsinfrastruktur (auch: satellitengesteuerte Navigationstechnik)**
- **Ableitung von Handlungsempfehlungen für den gezielten Ausbau der Verkehrsinfrastruktur sowie bessere Organisation der öffentlichen Verkehrsmittel**

Mögliche Langfrist-Ziele:

- **Förderung von (Car-)Sharing-Modellen**
- **Unterstützung autonomen Fahrens durch Fahrzeuge**

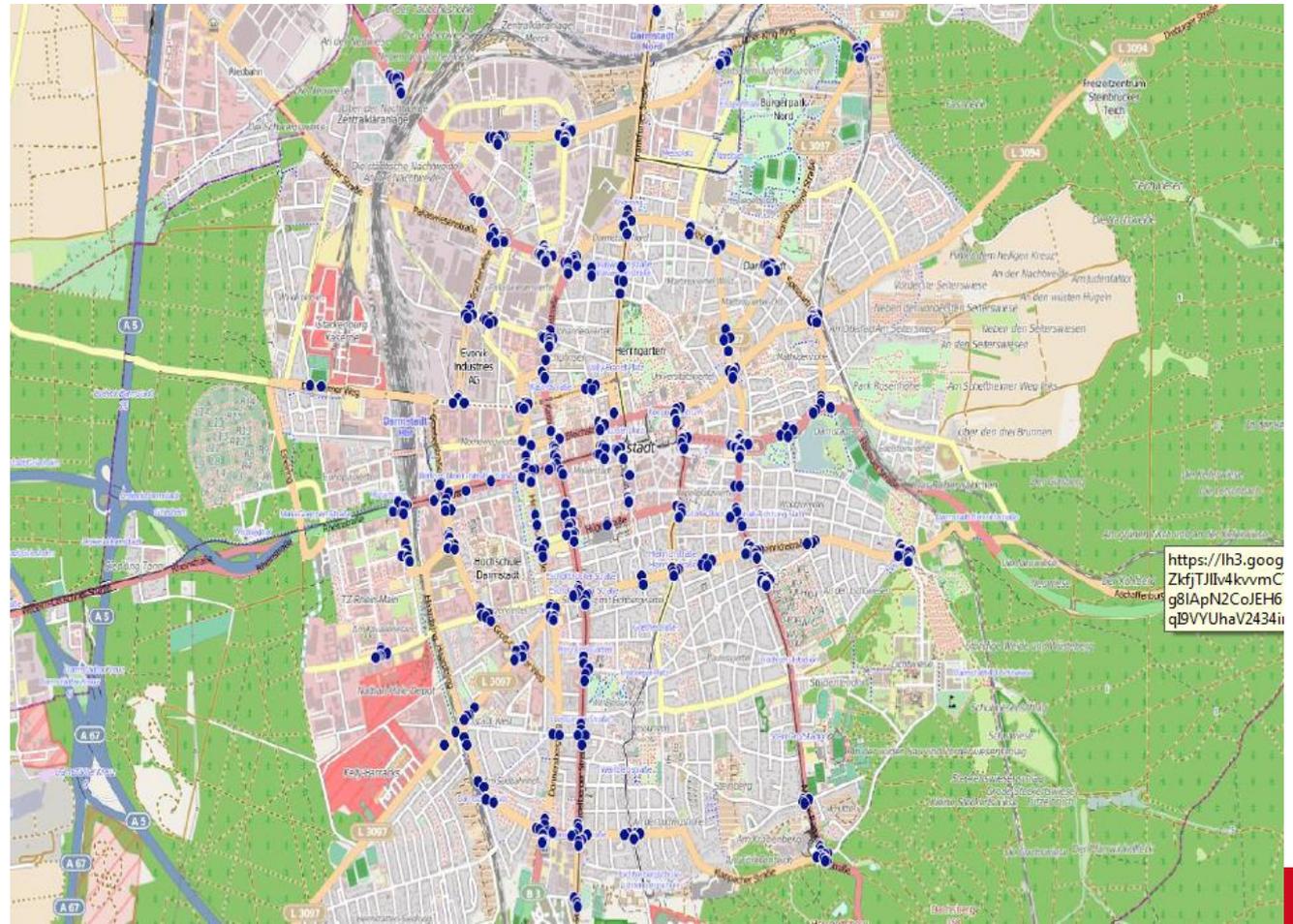
Voraussetzungen:

Digitalisierung und (Echtzeit-)Auswertung von Daten sowie sofortige Umsetzung entsprechender Maßnahmen zur Verkehrssteuerung.

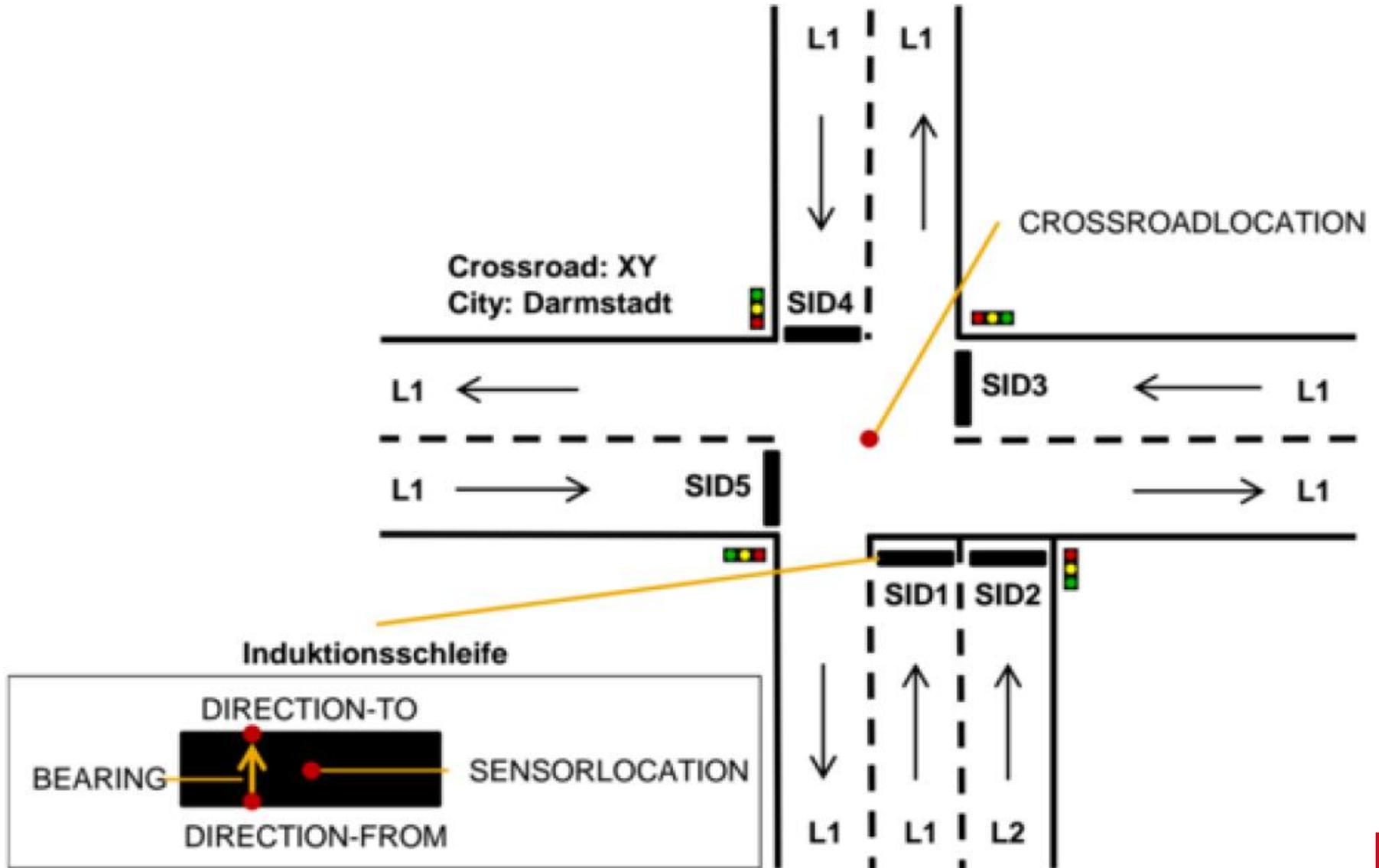
- Knoten und bewertete Kanten
- Kürzester Weg (Zeit oder Entfernung)
 - Dijkstra Algorithmus
- Kapazität
 - Min-Cut Max-Flow Theorem



- Daten von Darmstadt über die UrbanPulse-Plattform bereitgestellt
- angereichert mit Geoinformationsdaten von Open Street Map
- 96 Kreuzungen
- 858 Sensoren



Beispiel einer Straßenkreuzung



Sensor Events

"SAP_NIC_TRAFFIC_PULSE".EVENTS	
10	E_DATE
10	E_LOAD
12	E_COUNT
12	E_CONGESTION
12	E_SPEED
10	E_INTERVAL
10	E_SID
10	E_MESSAGETYPE
10	E_MESSAGE_TIMESTAMP

Sensoren

"SAP_NIC_TRAFFIC_PULSE".SENSORS	
10	S_SID
10	S_TYPE
10	S_CITY
10	S_CROSSROAD
12	S_LANE
12	S_SENSORLOC_LAT
12	S_SENSORLOC_LNG
12	S_SENSORDIR_BEARING
12	S_SENSORDIR_FROM_LAT
12	S_SENSORDIR_FROM_LNG
12	S_SENSORDIR_TO_LAT
12	S_SENSORDIR_TO_LNG
12	S_CROSSROADLOC_LAT
12	S_CROSSROADLOC_LNG
10	S_MESSAGETYPE
10	S_MESSAGE_TIMESTAMP
10	S_DISTRICT_ID

"SAP_NIC_TRAFFIC_PULSE".GIS.SENSOR_SEGMENT	
10	SENSOR_ID
12	SEGMENT_ID

Knoten & Kanten

"SAP_NIC_TRAFFIC_PULSE".GIS.SEGMENTS	
12	SEGMENT_ID
12	FROM_LAT
12	FROM_LNG
12	TO_LAT
12	TO_LNG
12	ONE_WAY
12	LENGTH
12	WAY_ID
12	WAY_SUB_SEQ
12	SPLIT_SEQ
12	FROM_ND
12	TO_ND
12	WAY_TYPE
10	WAY_NAME
12	STRUCT_TYPE
12	LAYER

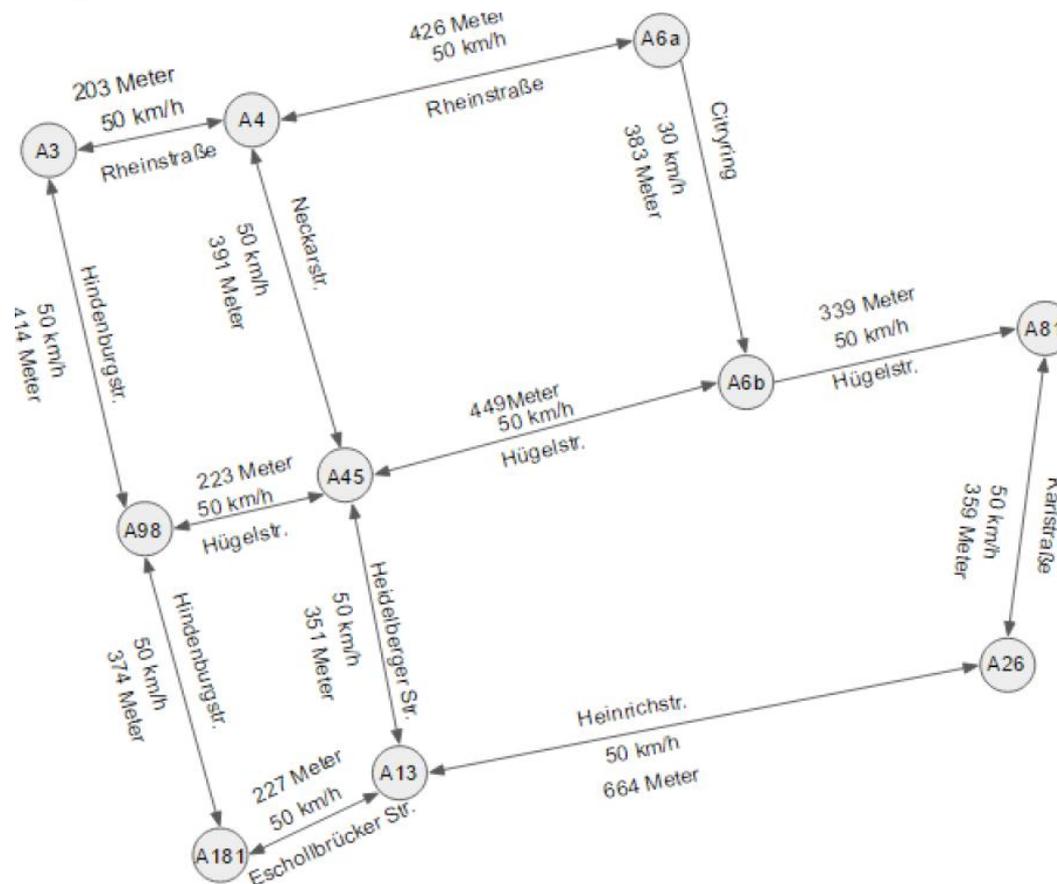
Bewertung der Kanten: Validierung E_SPEED in Tabelle EVENTS

- E_COUNT: Anzahl der Detektionen (Fahrzeuge) eines Sensors
- E_LOAD: Prozentuale Belegung eines Sensors innerhalb des Intervalls in %
 - COUNT 1 und LOAD 100 → Stau
- E_CONGESTION: Berechnete Verkehrsbelastung (in Prozent, Black Box)

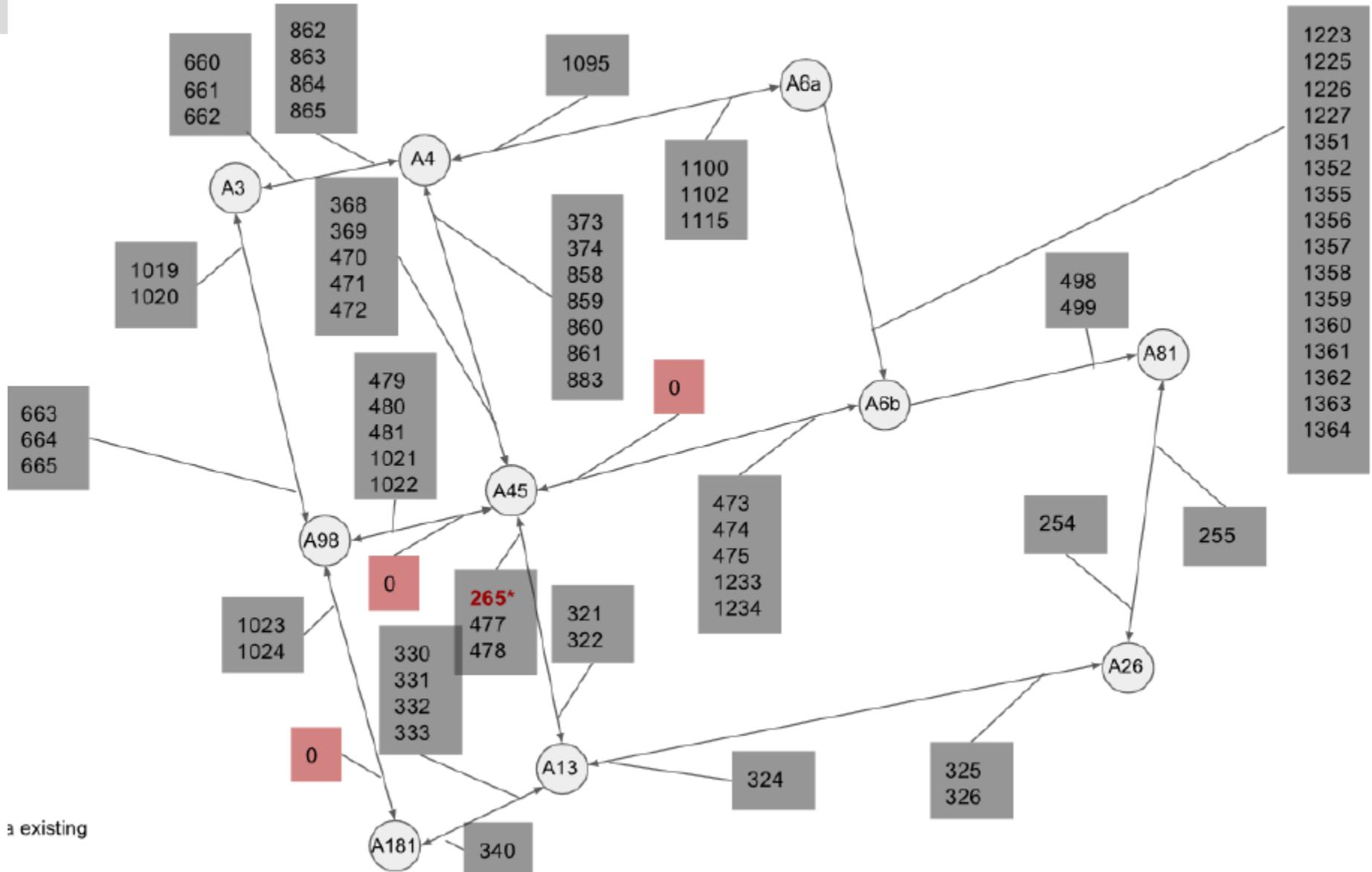
```
SELECT * FROM "SAP_NIC_TRAFFIC_PULSE"."EVENTS" WHERE E_SID = '1095' AND E_DATE BETWEEN '2015-09-11 15:00' AND '2015-09-11 17:00'
```

	E_DATE	E_LOAD	E_COUNT	E_CONGESTION	E_SPEED	E_INTERVAL	E_SID	E_MESSAGE TYPE	E_MESSAGE TIMESTAMP
8	Sep 11, 2015 4:19:00.0 PM	0	0	49.684861179079036	?	1	1095	SensorEvent	Sep 11, 2015 4:22:01.927 PM
9	Sep 11, 2015 3:28:00.0 PM	0	0	39.88638881769859	?	1	1095	SensorEvent	Sep 11, 2015 3:31:59.731 PM
10	Sep 11, 2015 3:42:00.0 PM	0	0	26.973610954462657	?	1	1095	SensorEvent	Sep 11, 2015 3:45:00.322 PM
11	Sep 11, 2015 3:28:00.0 PM	0	0	40.71208326503013	?	1	1095	SensorEvent	Sep 11, 2015 3:30:59.685 PM
12	Sep 11, 2015 3:26:00.0 PM	38	2	39.91263881827601	?	1	1095	SensorEvent	Sep 11, 2015 3:28:59.639 PM
13	Sep 11, 2015 3:22:00.0 PM	41	2	40.28013882635989	?	1	1095	SensorEvent	Sep 11, 2015 3:24:59.482 PM
14	Sep 11, 2015 4:10:00.0 PM	70	3	41.73819442926389	?	1	1095	SensorEvent	Sep 11, 2015 4:13:01.458 PM
15	Sep 11, 2015 4:02:00.0 PM	20	2	31.29055547859106	?	1	1095	SensorEvent	Sep 11, 2015 4:05:01.199 PM
16	Sep 11, 2015 3:52:00.0 PM	35	1	34.38333326329788	?	1	1095	SensorEvent	Sep 11, 2015 3:55:00.751 PM
17	Sep 11, 2015 4:36:00.0 PM	10	1	35.64333329101404	?	1	1095	SensorEvent	Sep 11, 2015 4:40:02.658 PM

- War mit dem Datenmaterial nicht automatisiert möglich (Segmente sehr kleinteilig)
- Beschränkung auf Ausschnitt



Zuordnung der Sensoren



a existing

Import in HANA und Berechnung Congestion

"C5234760".KANTEN_GRAPH	
12	KANTEN_ID
12	FROM_ND
12	TO_ND
AB	FROM_ND_CR_ID
AB	TO_ND_CR_ID
12	MAX_SPEED_KM/H
12	LENGTH
AB	STREET
12	S_SID

"SAP_NIC_TRAFFIC_PULSE".EVENTS	
12	E_DATE
12	E_LOAD
12	E_COUNT
12	E_CONGESTION
12	E_SPEED
12	E_INTERVAL
AB	E_SID
AB	E_MESSAGE_TYPE
12	E_MESSAGE_TIMESTAMP

SQL Result

```
SELECT Graph."KANTEN_ID", MEDIAN(Graph.median) FROM
  (SELECT "C5234760"."KANTEN_GRAPH"."KANTEN_ID", "C5234760"."KANTEN_GRAPH"."S_SID", MEDIAN(event2.E_CONGESTION) as median
    FROM "C5234760"."KANTEN_GRAPH"
  LEFT JOIN
    ( SELECT * FROM
      ( SELECT E.SID, TO TIME(E.DATE) as Uhrzeit.
```

	KANTEN_ID	MEDIAN(MEDIAN)
4	26	40.216250136028...
5	24	39.558227697272...
6	12	34.802115641083...
7	3	49.122516325729...
8	25	51.271428609532...
9	14	44.816989171011...
10	18	44.816989171011...

- **Programmiersprache und Entwicklungsumgebung für statistische Berechnungen und visuelle Darstellung**
- **Open Source**
- **Industriestandard**
- **in kommerzielle Lösungen integriert**
 - SAP HANA
 - Microsoft Machine Learning
- **Package iGraph**

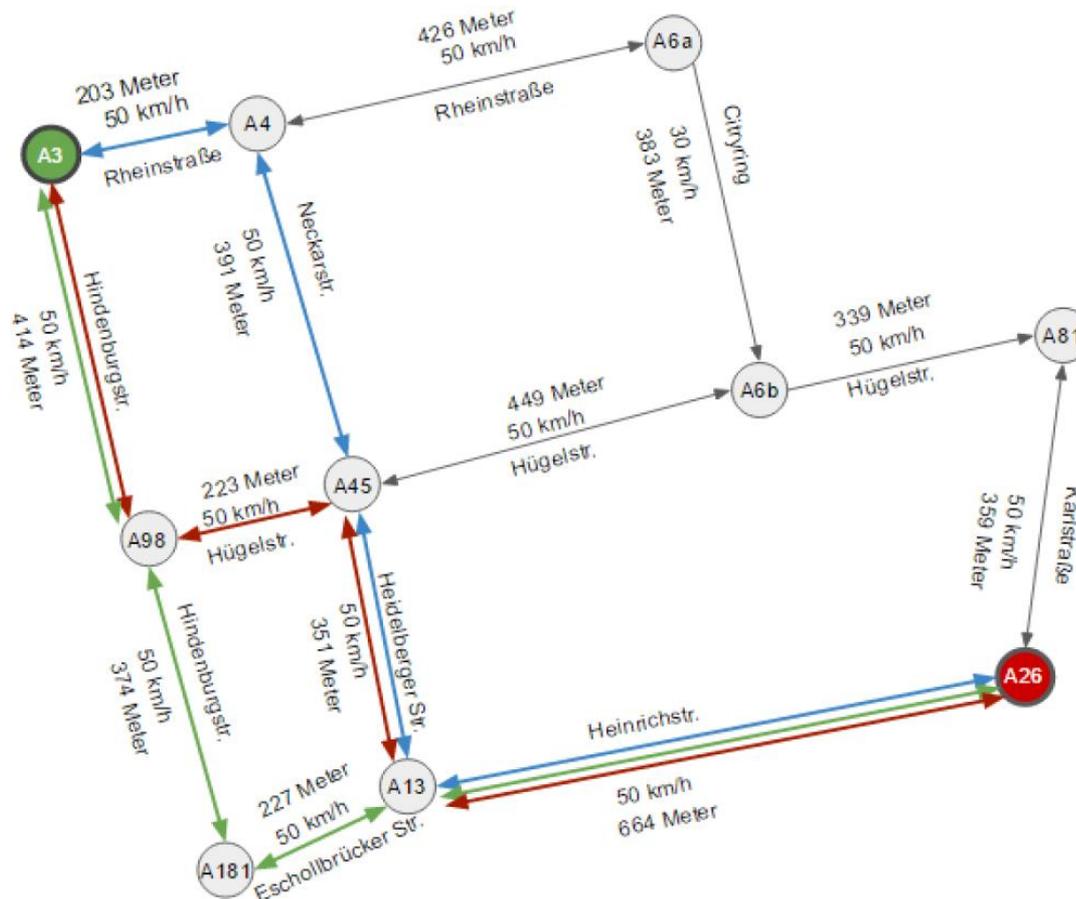
Kürzester Weg mit Dijkstra Algorithmus



```
##BERECHNUNGEN
```

```
##GraphFrAVG
```

```
distances(GraphFrAVG, v = 'A3', to = 'A26', weights = (E(GraphFrAVG)$LENGTH * 3.6 * 100 /  
(E(GraphFrAVG)$MAX_SPEED_KM.H * (100 - E(GraphFrAVG)$X2015.08.14..7.9))), mode = 'out')  
shortest_paths(Graph, from = 'A3', to = 'A26', weights = (E(GraphFrAVG)$LENGTH * 3.6 * 100 /  
(E(GraphFrAVG)$MAX_SPEED_KM.H * (100 - E(GraphFrAVG)$X2015.08.14..7.9))), mode = 'out')
```



- **Ziel wurde erreicht**
- **Einige „Handarbeit“ und Beschränkung auf Ausschnitt**
- **Bewertung der Kanten mit Heuristik (Black Box Congestion)**

Ausblick

- **Längerer Zeitraum**
- **Weitere Datenquellen (z.B. Wetterdaten)**
- **Betrachtung Kapazität mit Min-Cut Max-Flow Theorem**

Dank an Projektpartner

Mayato: Marcus Dill



SAP: Niraj Sing und Bengt Mertens



Danke!

Fragen?



**Institut für Wirtschaftsinformatik
an der Hochschule Ludwigshafen**



Prof. Dr. Klaus Freyburger
Professor für Wirtschaftsinformatik
Tel.: 0621-5203-219
Email: klaus.freyburger@hs-lu.de