



# StreamPipes: Offene Daten einfach analysieren

mFUND-Konferenz, Berlin

Dr. Dominik Riemer  
16.10.2018



# mFUND-Projekt WEKOVI

Werkzeuge für die einfache Erstellung komplexer Vergleichsindizes



## Beispiel

Bewertung eines Bauplatzes bezüglich Lärm und Mobilität

### Bike Sharing

**Anzahl Bike Sharing-Stationen  
in Bezug auf Einwohnerzahlen**

### Luftqualität

**Durchschnittliche Luftqualität in  
der Umgebung**

### Lärmbelastung

**Gibt es eine vielbefahrene  
Straße (Landstraße, Autobahn)  
in der Nähe des Bauplatzes?**

# mCLOUD & Open Data-Portale

Viele Daten haben Raum- und Zeitbezug, Heterogenität

## Echtzeitdaten

z.B. Parkhausbelegungen,  
Umweltsensoren,  
Verkehrskameras,  
Verkehrsfluss

## Zeitbasierte Datensätze

z.B. Verkehrszählungen,  
Bußgelder,  
Carsharing,  
Taxinutzung

# Ziel

Einfacher Zugang, Harmonisierung und Analysen



## Datenzugang

Generische Adapter  
Spezifische Adapter  
Metadaten  
Datenströme & Historisch



## Datenharmonisierung und -analyse

Vorverarbeitung  
Filterung/Aggregation  
Mustererkennung  
Maschinelles Lernen



## Verwertung

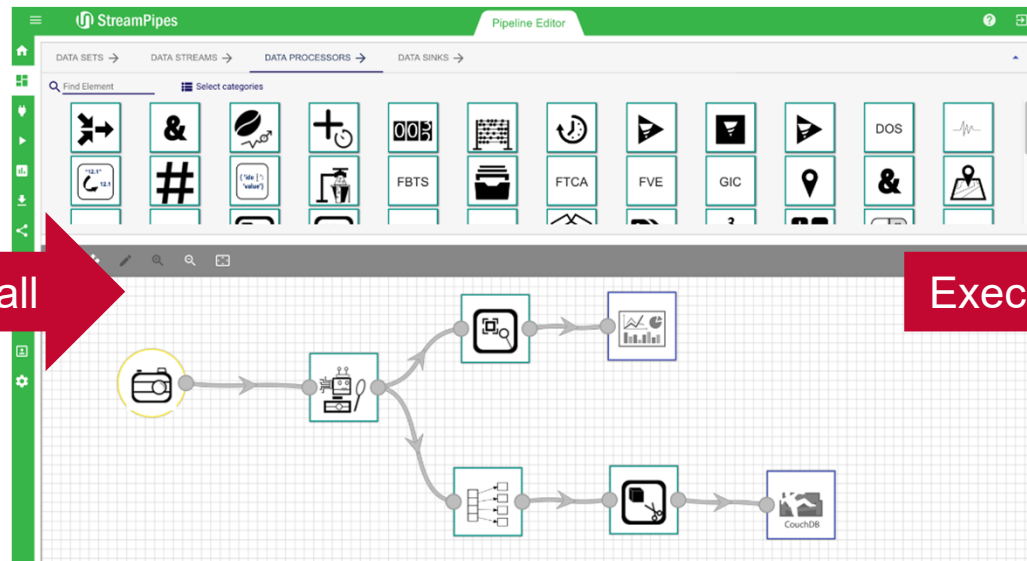
Situationserkennung  
Bereinigte Datensätze  
Einfache APIs/Visualisierung  
Drittanwendungen

# Self-Service Datenanalysen mit StreamPipes

Entwicklung hochverteilter Big Data-Anwendungen durch Fachanwender

Pipeline  
Element  
Microservices

Install



Execute

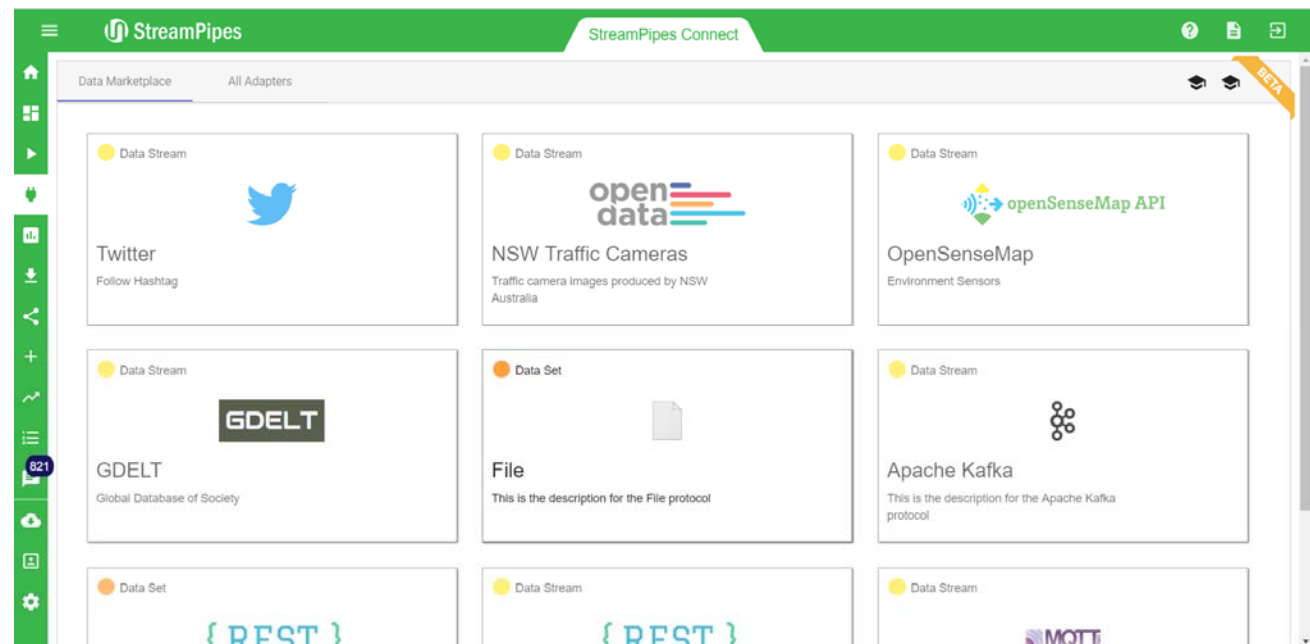
Big Data  
Infrastructure

# Datenanbindung

"Data Marketplace" für beliebige Protokolle und Formate, spezifische Adapter

## Highlights & Innovationen

- Anbindung von Datenströmen und Datensätzen
- Generische Adapter für häufig genutzte Protokolle (z.B. REST, MQTT) und Formate (z.B. JSON, XML, GeoJSON)
- Automatische Datenerkennung durch maschinelles Lernen
- RDF-basierte Metadaten (z.B. Qualitätseigenschaften, Semantik)



# Datenharmonisierung und -analyse

## Erweiterbare Toolbox



### Highlights & Innovationen

- Erweiterbare Toolbox für Vorverarbeitungs- und Analysefunktionen
- Zustandslos und zustandsbehaftet
- Semantische Konsistenzprüfung vereinfacht Modellierung
- Einbindung von Lernverfahren wie neuronalen Netzen möglich

The screenshot shows the StreamPipes Pipeline Editor interface. At the top, there's a green header with the StreamPipes logo and the title 'Pipeline Editor'. Below the header, there are tabs for 'OVERVIEW', 'STATISTICS', 'ERRORS', and 'QUICKEDIT'. The main workspace is a grid where a pipeline is being built. The pipeline starts with a source icon, followed by a filter icon, then an 'Image Enricher' icon, and finally a sink icon. Below the grid, there's a 'Quick Edit' panel with a search bar and a list of components. The components listed are 'Image Enricher', 'Image Classification', and 'Array Width Bounding Boxes'. To the right of the grid, there's a large toolbox containing various icons for different components, such as 'A', '&amp;', '+', '000', 'GIC', 'G', 'GR', 'IR', 'NF', 'SGE', and others. The interface is clean and modern, with a green and white color scheme.

# Datenverwertung & Anwendungen

Transfer zu Drittsystemen, Bereitstellung, Benachrichtigungen



Benachrichtigungen bei hohen Schadstoffbelastungen in der Umgebung

Trendanalysen von Verkehrsdaten

Entfernung datenschutzrelevanter Inhalte bei kontinuierlichen Bilddaten

Durchschnittliche Geschwindigkeitsdaten auf bestimmten Strecken

Ranking von Bike-Sharing-Stationen pro Stadt

Vorverarbeitung von Fachdaten zu "einfach nutzbaren" Datenformaten



# Übertragbarkeit auf andere Projekte

## Einatzmöglichkeiten



1 **Anforderung:** Zeitbasierte Daten und Sensordaten auf einfache Weise beobachten, harmonisieren und auswerten

2 **Anforderung:** Integrierte Sicht auf Live-Daten unterschiedlicher Datenquellen

3 **Anforderung:** Ergebnisse komplexer Datenanalysen in einem einzigen Tool unterschiedlichen Zielgruppen zugänglich machen

StreamPipes ist Open Source und unternehmensfreundlich lizenziert!

# Vielen Dank!



[streampipes.org](https://streampipes.org)



[docs.streampipes.org](https://docs.streampipes.org)



[github.com/streampipes](https://github.com/streampipes)

[riemer@fzi.de](mailto:riemer@fzi.de)