

# Machine Learning and Automated Cognitive Capabilities

PREVISION Project

Presenter: Uwe Zeltmann

# Machine Learning and Automated Cognitive Capabilities

- Tools for converting large volumes of heterogeneous data into actionable intelligence
- Discovery of implicitly preexisting information and relations between data entities
- Detection of trends regarding future evolution of criminal activities

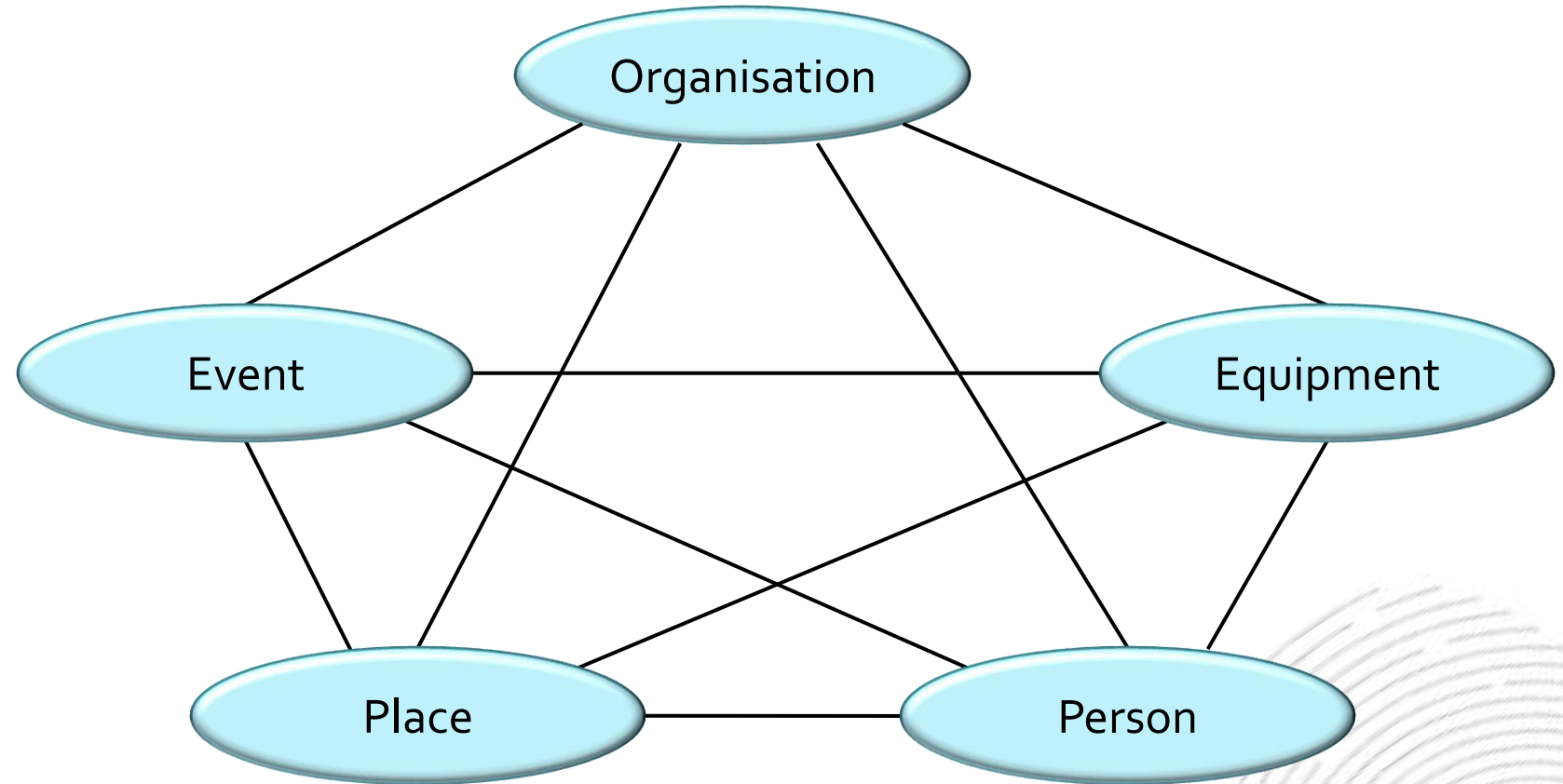
# Semantic Information Processing

# Semantic Information Processing

## - Overview

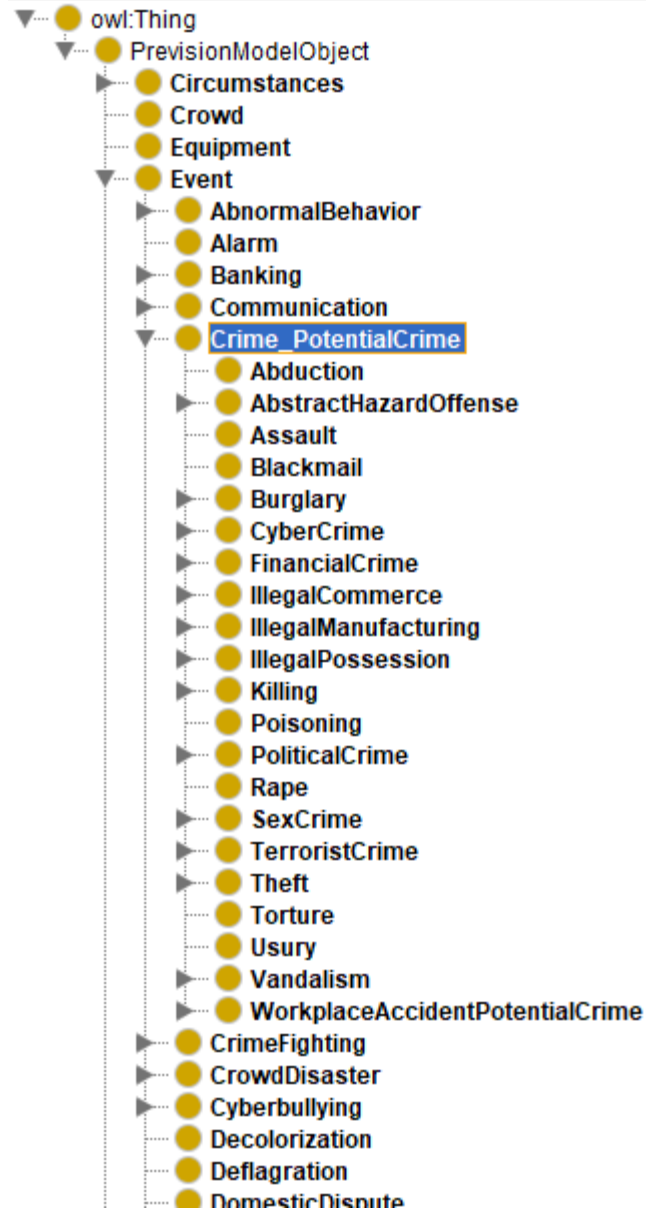
- Framework to deal with knowledge in a formalized way
- Data collected by several platform components
  - formulated in terms of an ontology
  - stored in a knowledge base
- Infer knowledge by predicate logical and probabilistic reasoning
  - using rules modelled by domain experts

# Semantic Information Processing - Main Ontology Concepts



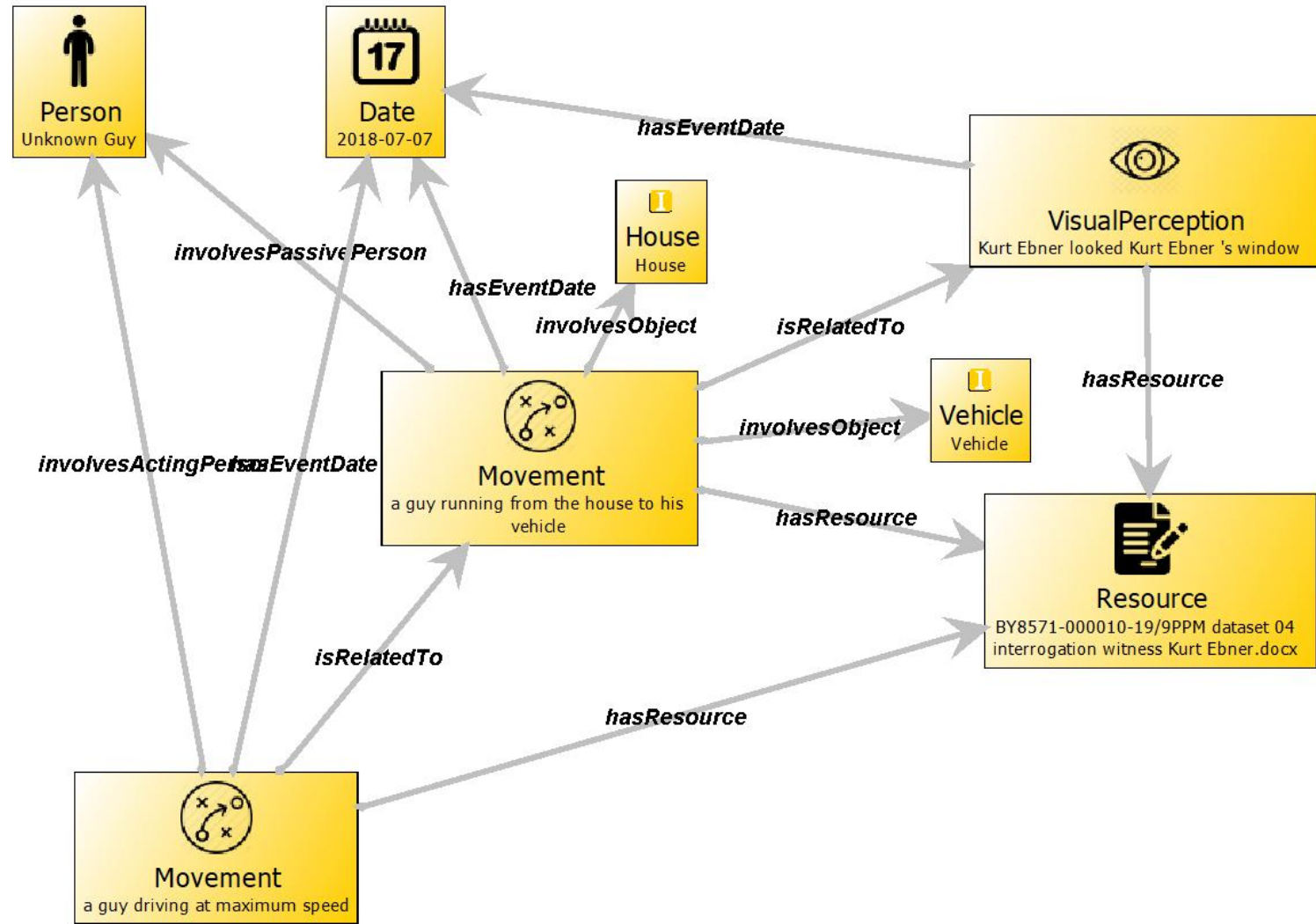
# Semantic Information Processing

## - Ontology Class Hierarchy



# Semantic Information Processing

## - RDF Data Structure



# Semantic Information Processing - Technical Setup

- SPARQL endpoint: Jena Fuseki
- REST API providing ontology-proof write access
- HMI permitting manual data editing
- Markov Logic Network implementation Tuffy

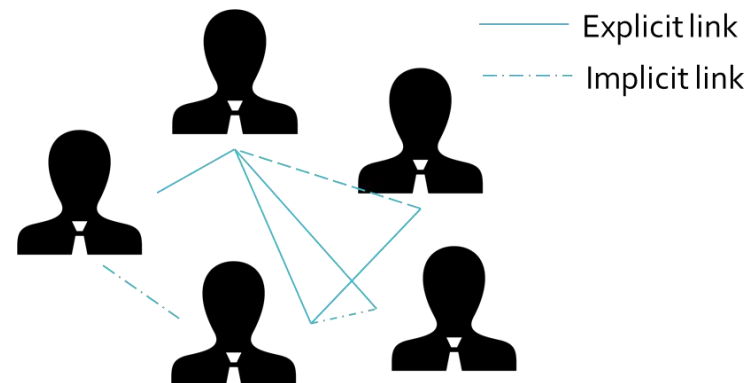


# Smart Text and Social Network Analysis

# Smart Text and Social Network Analysis

- Person Interaction

- Component to help fill the gaps in people interactions. Common topics/ jargon and interactions are combined.
- Starting points: Community detection algorithms & Data from use cases
- Aim: linking unconnected communities (unconnected due to missing data, e.g. missing explicit interactions)
- Means: Using common jargon from use-case oriented lexicons



# Smart Text and Social Network Analysis

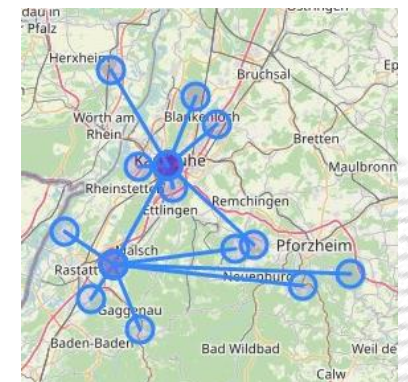
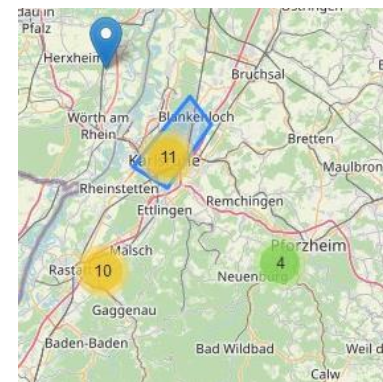
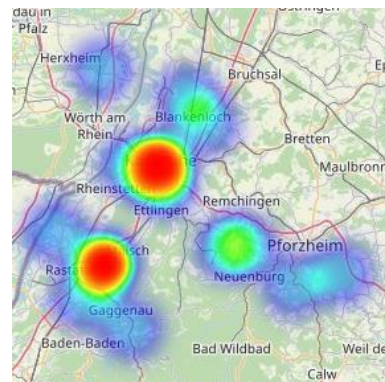
- Radicalization  
Assessment

- Predictive Analytics
  - Analysis of biographies of known offenders
  - Semantic and keyword analysis of textual messages
- Personal monitoring
- Trend Analysis

# Visual Analytics

# Visual Analytics - Geospatial Analysis

- Tool for the spatio-temporal visual analysis of geographical data, currently:
  - Heatmaps
  - Time series heatmaps
  - Marker clusters
- (Online) maps based on OpenStreetMap
- Offline usage may be possible by setting up:
  - An OpenStreetMap tile server
  - A Nominatim search engine for OpenStreetMap data



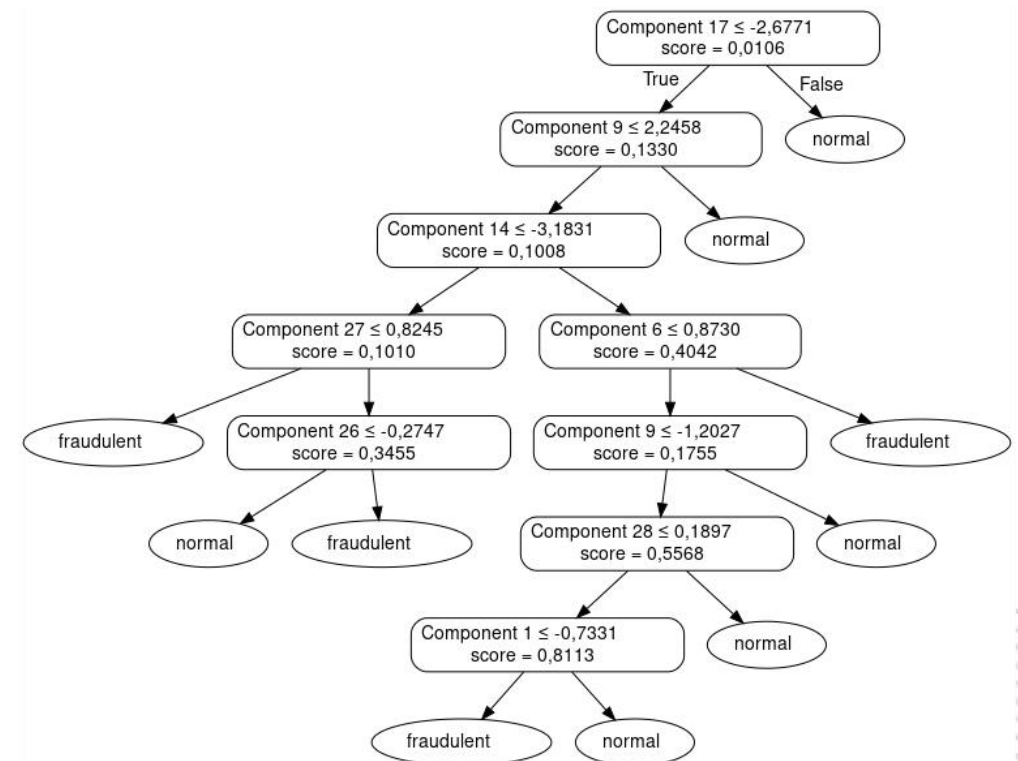
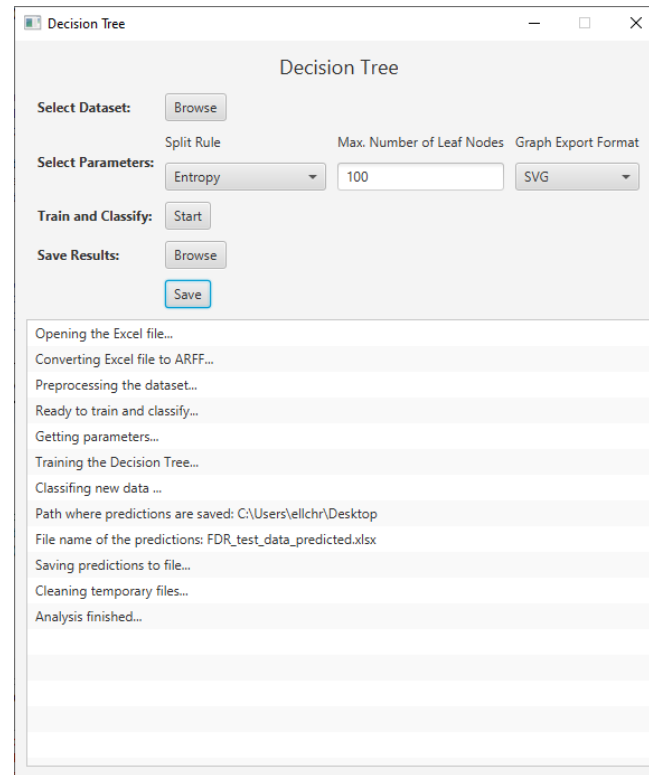
# Behavioural Analysis and Anomaly Detection

# Behavioural Analysis and Anomaly Detection

- Services providing automated capabilities for detecting behaviour changes, abnormality, outliers, and trends
- Algorithms/methodologies for:
  - Data classification based on decision trees
  - Clustering techniques for analysing telecom and financial data
  - Behavioural analysis and anomaly detection for videos

# Behavioural Analysis and Anomaly Detection - Classification Based on Decision Trees

- Supervised learning tool for multi-class classification
- Optimised for ease of use by the end user:
  - MS Excel in- and outputs
  - Automatic preprocessing
  - Crucial parameters can be adjusted
- Visualisation of the trained tree for the explainability of decisions





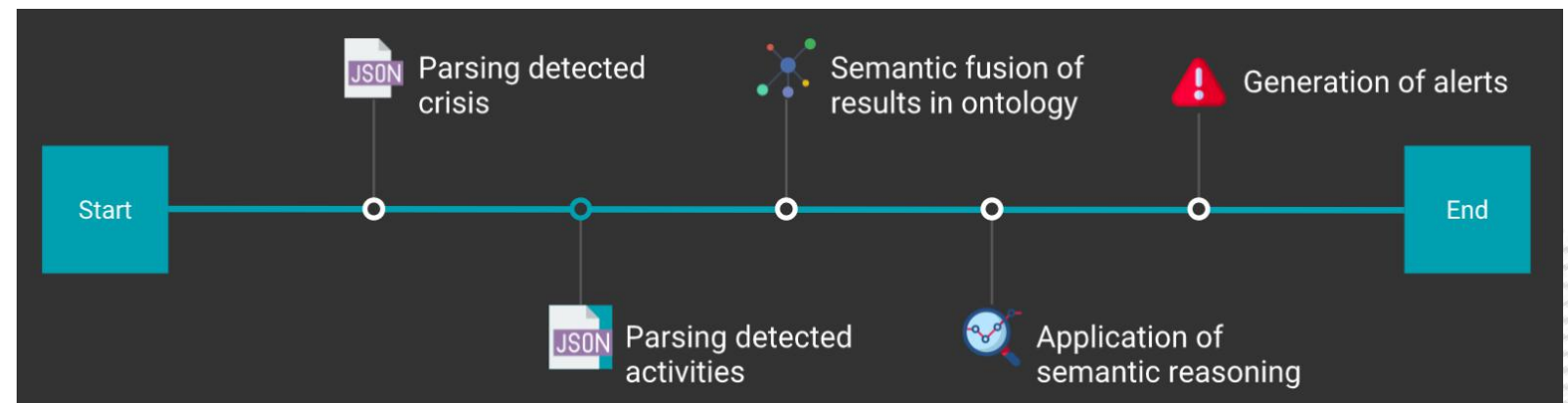
# Behavioural Analysis and Anomaly Detection - Clustering of Telecom and Financial Data

- Financial Data
  - Compute outliers in financial data records.
  - K-nearest neighbor algorithm
  - Similarity measures: Euclidean/ Manhattan/ Minkowski/ Mahalanobis
- Telecom Data
  - Compute outliers call data records (telecommunication data).
  - CBLOF (Cluster-Based Local Outlier Factor) algorithm → identify small or sparse clusters and declare the objects in those clusters to be outliers
    - clusters a data set, and sorts clusters according to decreasing size
    - CBLOF is calculated as the product of the size of the small cluster and the similarity between the centers of math of the small cluster and the closest large cluster

# Behavioural Analysis and Anomaly Detection - Videos

## Crisis event detection and severity classification

- Detection of activities and crises (e.g. fire) in videos
- Semantic fusion of detections in the PREVISION ontology
- Application of rule-based semantic reasoning
- Identification of anomalies with the spatio-temporal combination of multiple video sources
- Classification of the crisis severity level
- Generation of alerts



# Further Tools Developed in PREVISION

# Further Tools Developed in PREVISION

- Smart data fusion
- Missing data visual analytics
- Regression trees
- Fake news detection methods for alarm signal checking
- Detection of cyber attacks against deep neural networks

Thank you!

