

# GeoVisual Analytics Platform for large-scale Multi-Agent Simulations

20.12.2016

Janus Dybulla

Hauptseminar - HAW Hamburg - Dept. Informatik



Hochschule für Angewandte  
Wissenschaften Hamburg  
*Hamburg University of Applied Sciences*



MARS GROUP

| Multi-Agent Research & Simulation

# Agenda

- ▶ Motivation
  - ▶ Rückblick
  - ▶ 1.Prototyp
- ▶ Cloud-basierte Plattformen für MAS
- ▶ Integration von GeoVisual Analytics
- ▶ Laufende Arbeit
- ▶ Ausblick und Zielsetzung
- ▶ Chancen und Risiken

# MARS Group - Multi-Agent Research & Simulation

- MARS 2.0: Modellierung und Simulation von Millionen Agenten
- Verhalten von Individuen werden einzeln berücksichtigt
  - statische / dynamische Attribute
- u.a. Geospatiales Environment (GPS)
- Cloud-basierte Microservice-Architektur nach IaaS, PaaS und SaaS mit Docker
- Verschiedene AnwenderInnen
- Simulationsauswertung und Visualisierung

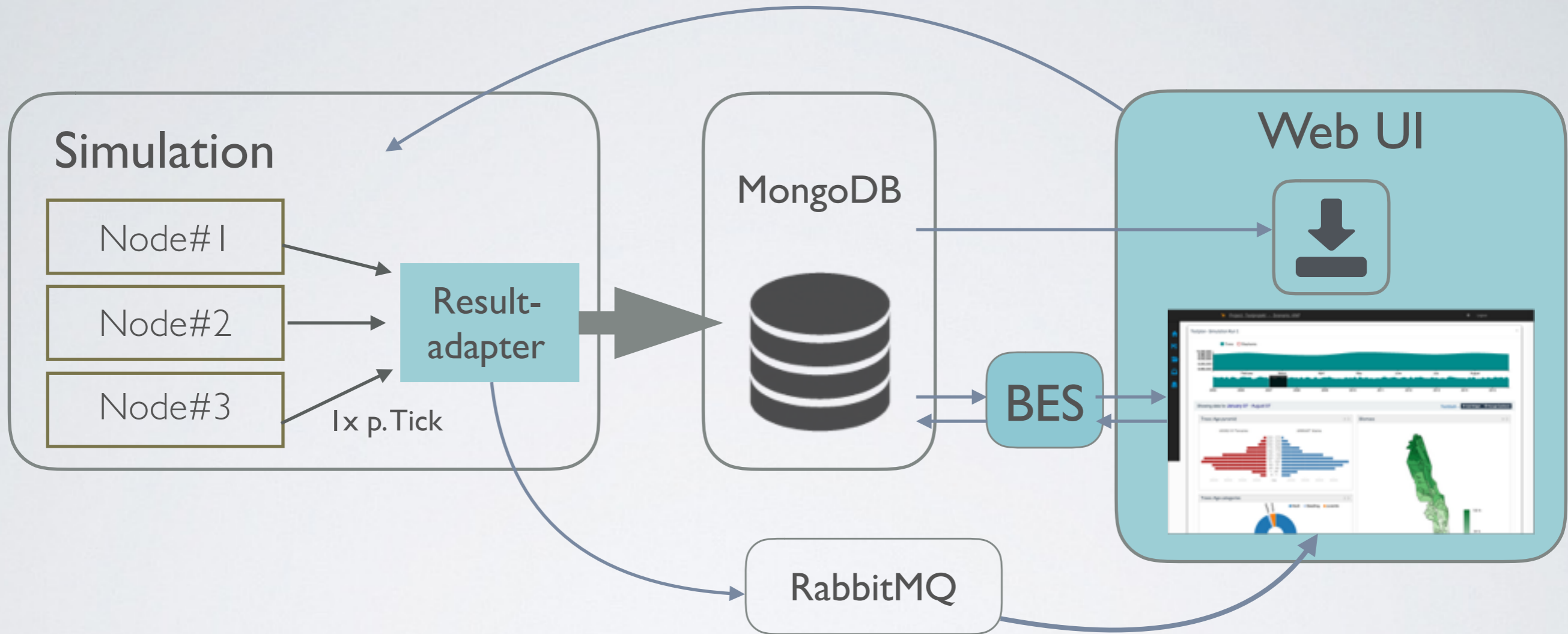
# Visual Analytics

- Erkenntnisgewinn aus extrem großen und komplizierten Datensätzen
- Kombination von
  - Automatischer Datenanalyse
  - Fähigkeiten des Menschen, Muster und Trends visuell zu erfassen
- Mehr als nur die interaktive, visuelle Analyse
  - „Analyse first, show the important, zoom/filter, analyse further, details on demand“ [KeZi06]
  - Multivariate coordinated views mit Linking und Brushing  
=> Dashboard

# Möglichkeiten von Visual Analytics im MAS-Umfeld

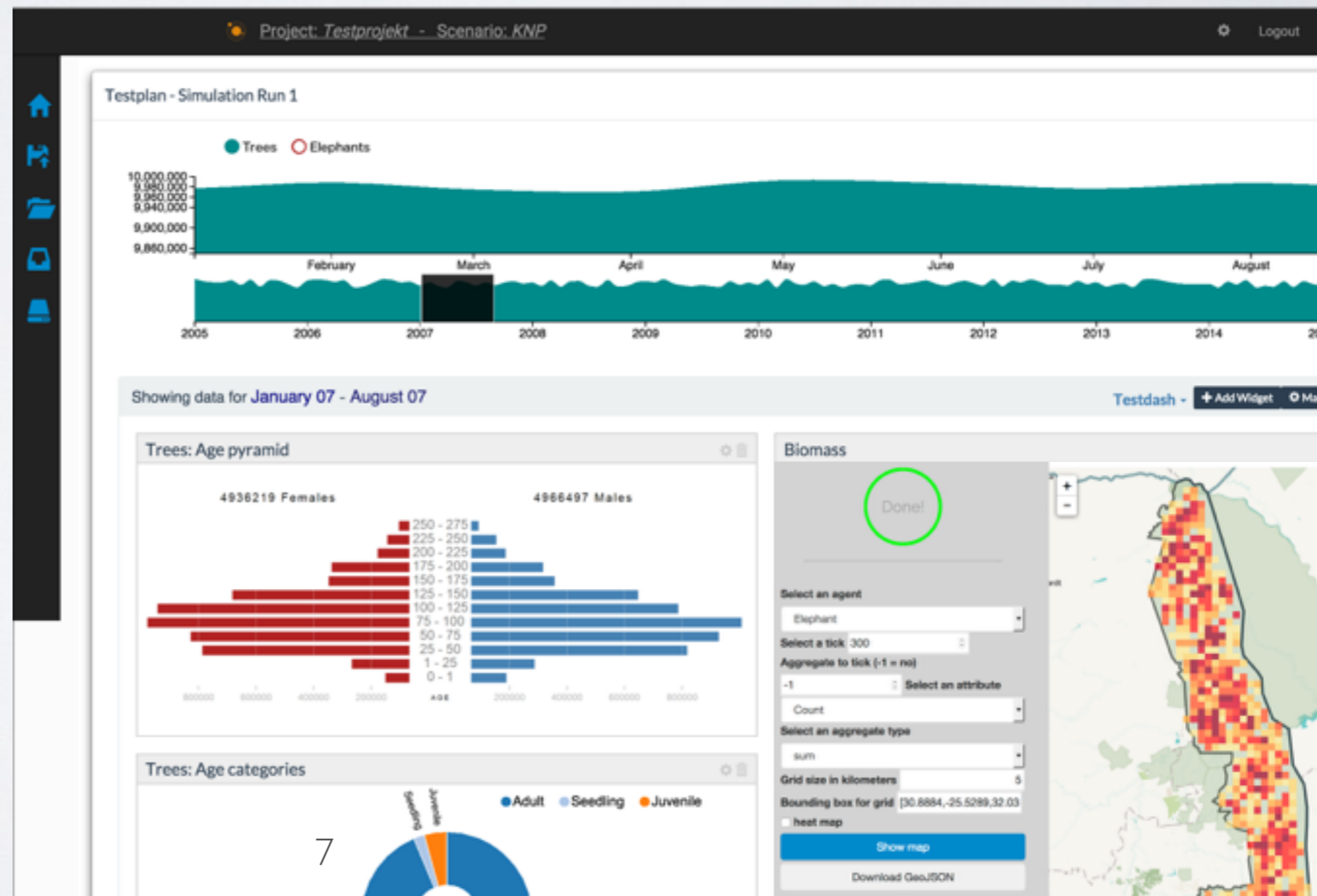
- Interaktive Analyse u. Darstellung der Simulationsergebnisse ermöglichen
- Beobachten von Ereignissen während einer Simulation
- Exploration und Vergleich verschiedener Simulationsergebnisse
- Statische und dynamische Validierung und Kalibrierung der Modelle

# Prototyp mit vorhandenen Gegebenheiten



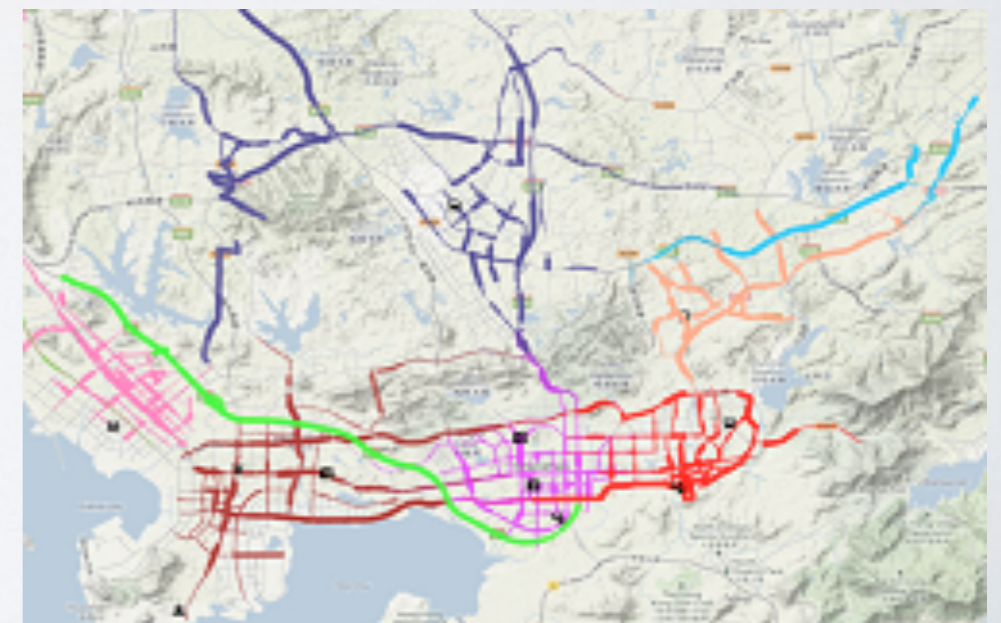
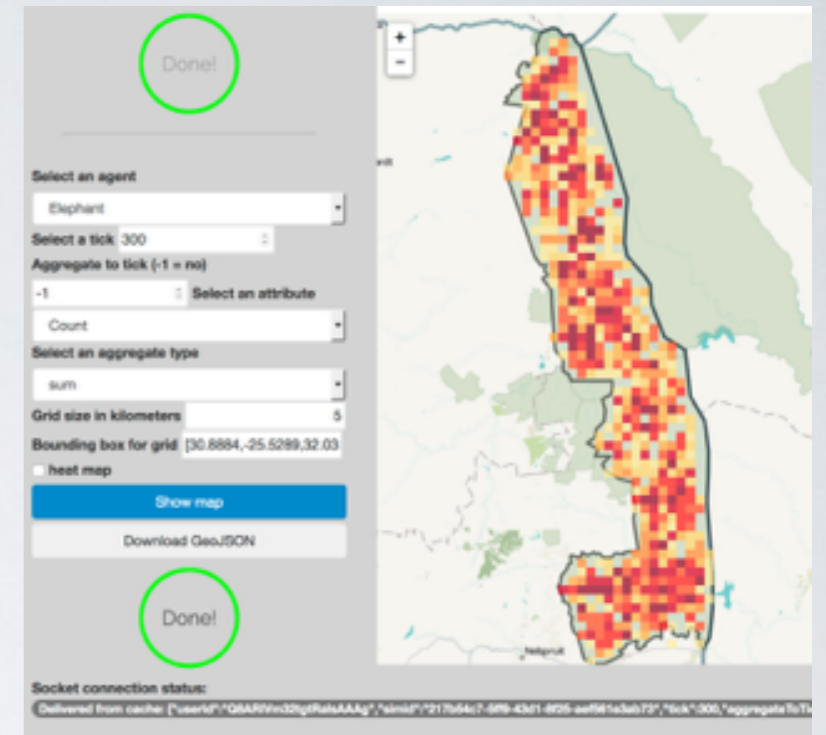
# Prototyp (Ansatz Grundprojekt)

- Datendarstellung mit Linking & Brushing
- Temporale Filterung & Sampling (Records, nicht visuelle Darstellung)
- Pull-Livereload
- Framework-Caching (BES)



# Temporal-geospatiale Darstellung im Prototyp (Ansatz Grundprojekt)

- Spatial Filtering
- Hierarchy binned/regional aggregation density nach [imMens I 4] :
  - Heatmaps
  - Gridmaps
- Multivariate Trajekturenanalyse:
  - Simplification / Flowlines nach [CWL08]
  - Anreicherung durch Attribute nach [CYG I 6]



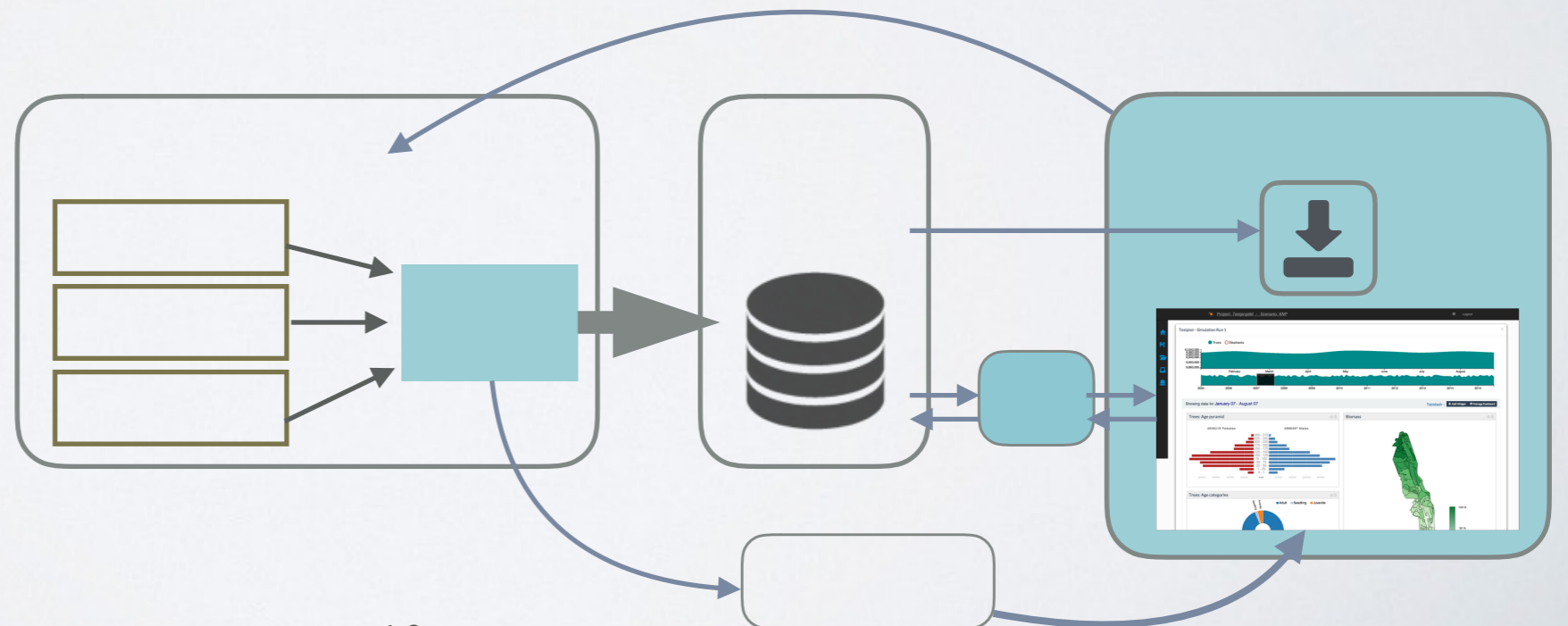


# Hürden bei der Realisierung der VA-Plattform

- Mangelndes Real-time (Post-Processing der Daten, Manuelles Triggern)
- Metriken und Dimensionen nicht freiwählbar ( != generische Agenten)
- Feste Verlinkungen, begrenzte Attributenauswahl
- Tiefgehende Datenanalyse nicht vorhanden

# Hürden bei der Realisierung der VA-Plattform

- Direkter Zugang auf die DB
- Performance der Cluster-DB-Lösung
  - Rausschreiben der Simulationsergebnis (BASE)
  - Hohe Latenz bei Datenabfrage (Index)
  - Speichermenge (Kompression)



# Möglichkeiten zur Auswertung von Simulationsergebnissen

<http://keynotetemplate.com/timeline/#>

## Damals



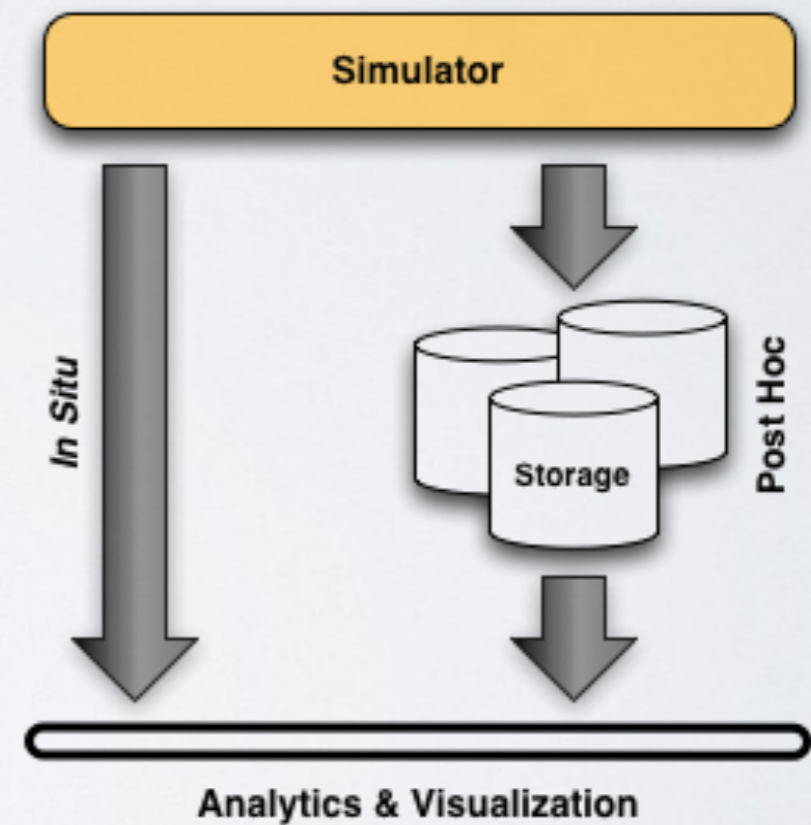
- Result Export
- Local / Dedicated Analysis

## Aktuell



- + In Situ / Batch Analysis
- + APIs / Connectors
- + dedizierte Visualisierungskomponenten

<http://www.paraview.org/in-situ/>



Blockierend & nicht-blockierend  
Ausführung

# Möglichkeiten zur Auswertung von Simulationsergebnissen

<http://keynotetemplate.com/timeline/#>

## Damals

- Result Export
- Local / Dedicated Analysis



## Aktuell

- + In Situ / Batch Analysis
- + APIs / Connectors
- + dedizierte Visualisierungskomponenten



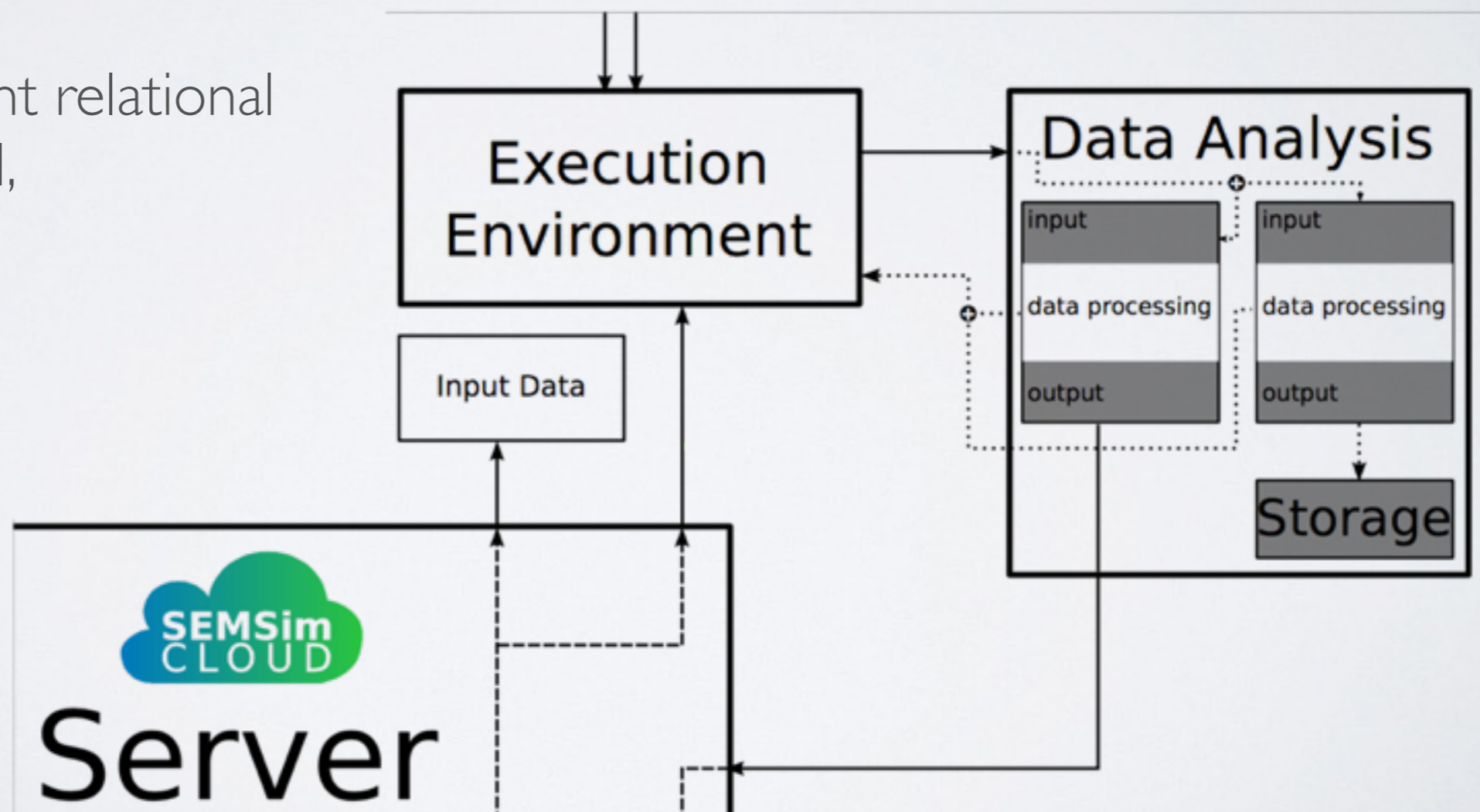
## Forschung

- + Stream Analysis
- + Cloud Analysis
- + Verteilt, skaliert
- + Kollaboration



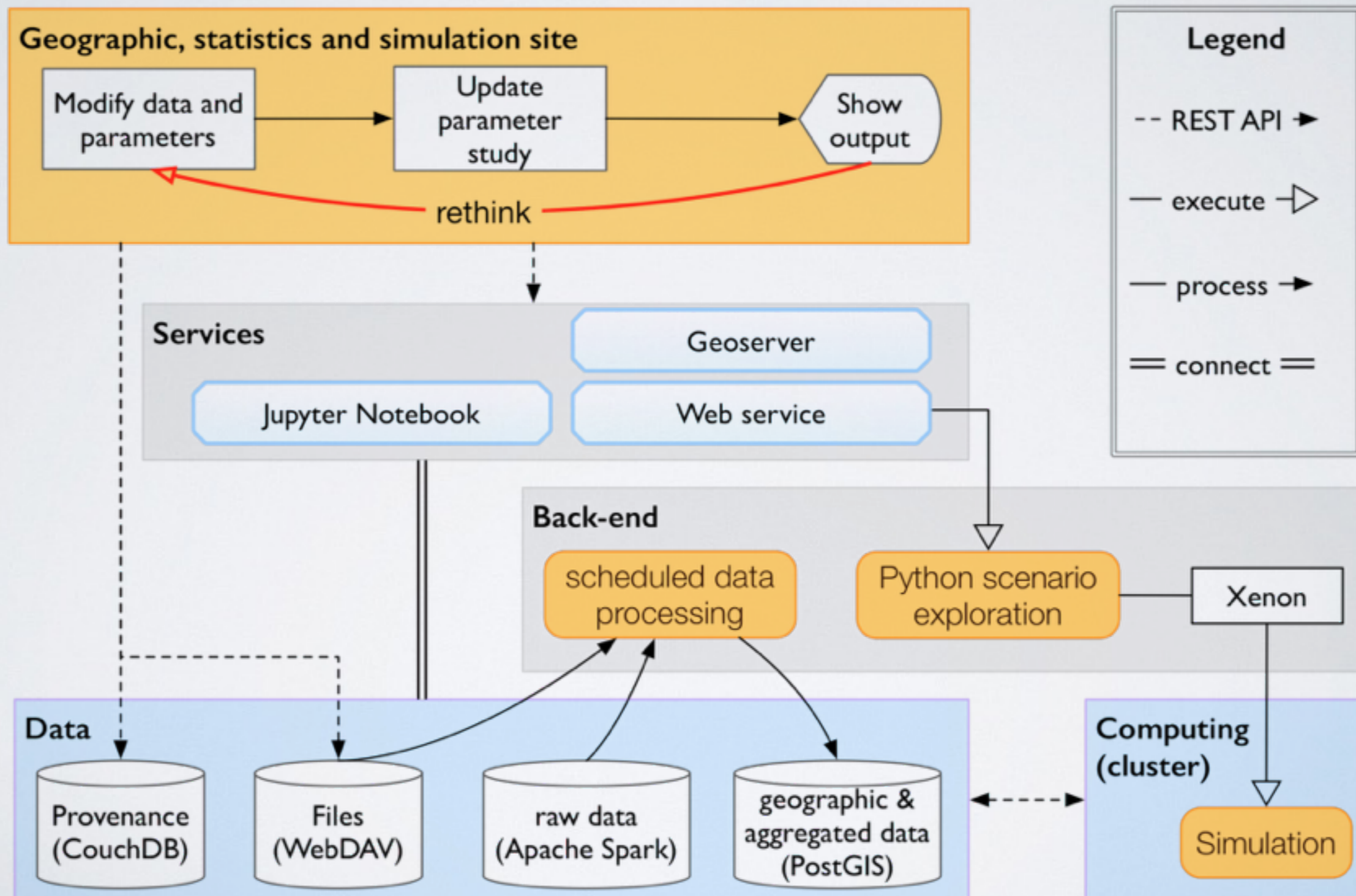
# Cloud-basierte Plattformen/Konzepte für MAS:

- Daniel Zehe, Vaisagh Viswanathan, Wentong Cai, and Alois Knoll: Online Data Extraction for Large-Scale Agent-Based Simulations (2016).
- Streamprocessing-Ansatz
- SimuSQL (Ähnlich wie mJADES [mjad12])
- Time-variant relational data model, Landmarks



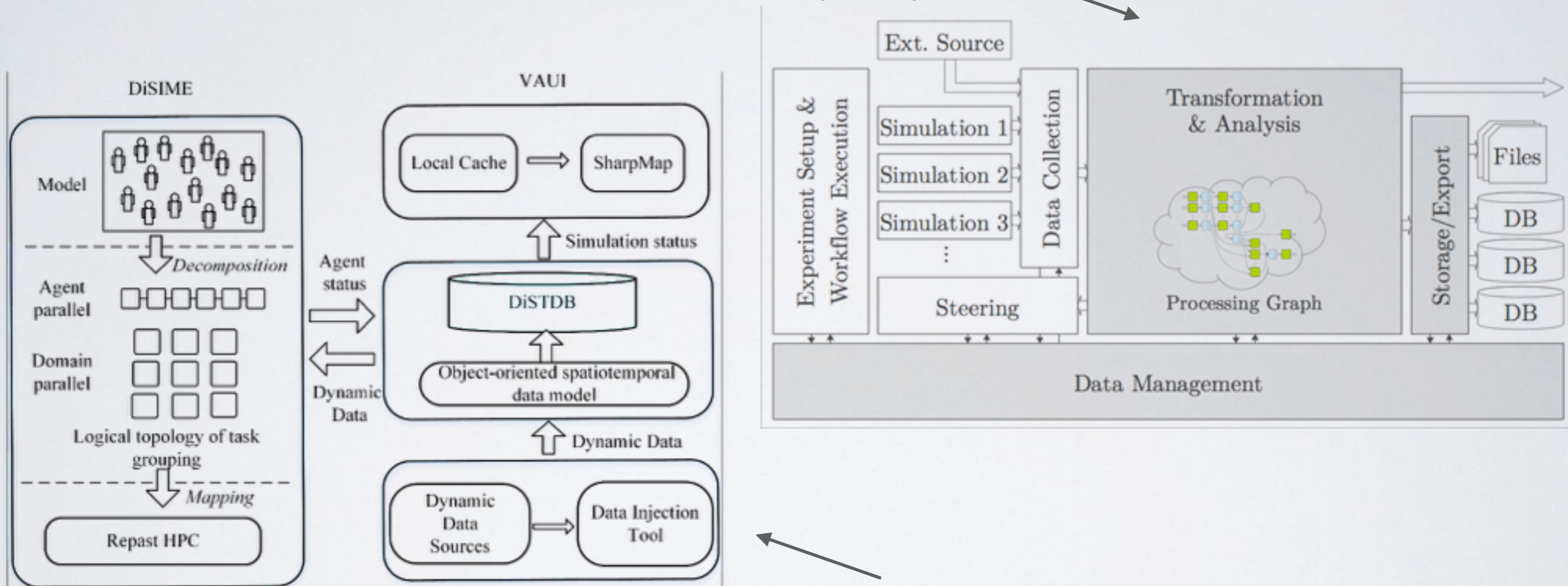
# Cloud-basierte Plattformen/Konzepte für MAS:

- J. Borgdorff, W. van Hage, L.J. Dijkstra, E. Mancini and M.H. Lees: Simulation-supported Urban Movement Analysis (2016)



# Cloud-basierte Plattformen/Konzepte für MAS:

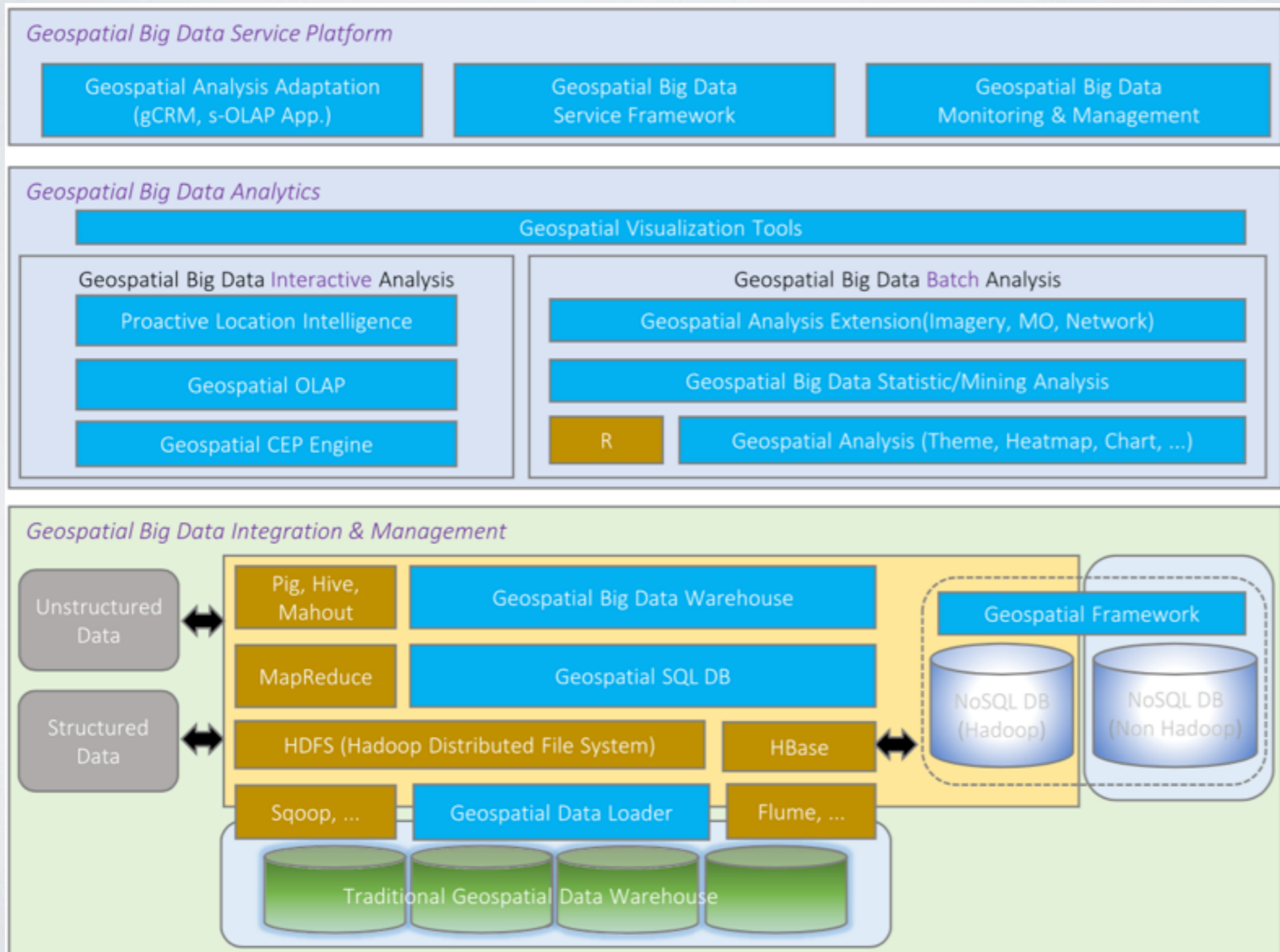
- Johannes Schützel, Holger Meyer, and Adelinde M. Uhrmacher:  
A stream-based architecture for the management and on-line analysis of unbounded amounts of simulation data (2014)



- Zhenqiang Li 1, Xuefeng Guan 1,2,\* , Rui Li 1,2 and Huayi Wu  
4D-SAS: A Distributed Dynamic-Data Driven Simulation and Analysis System for Massive Spatial Agent-Based Modeling (2016)

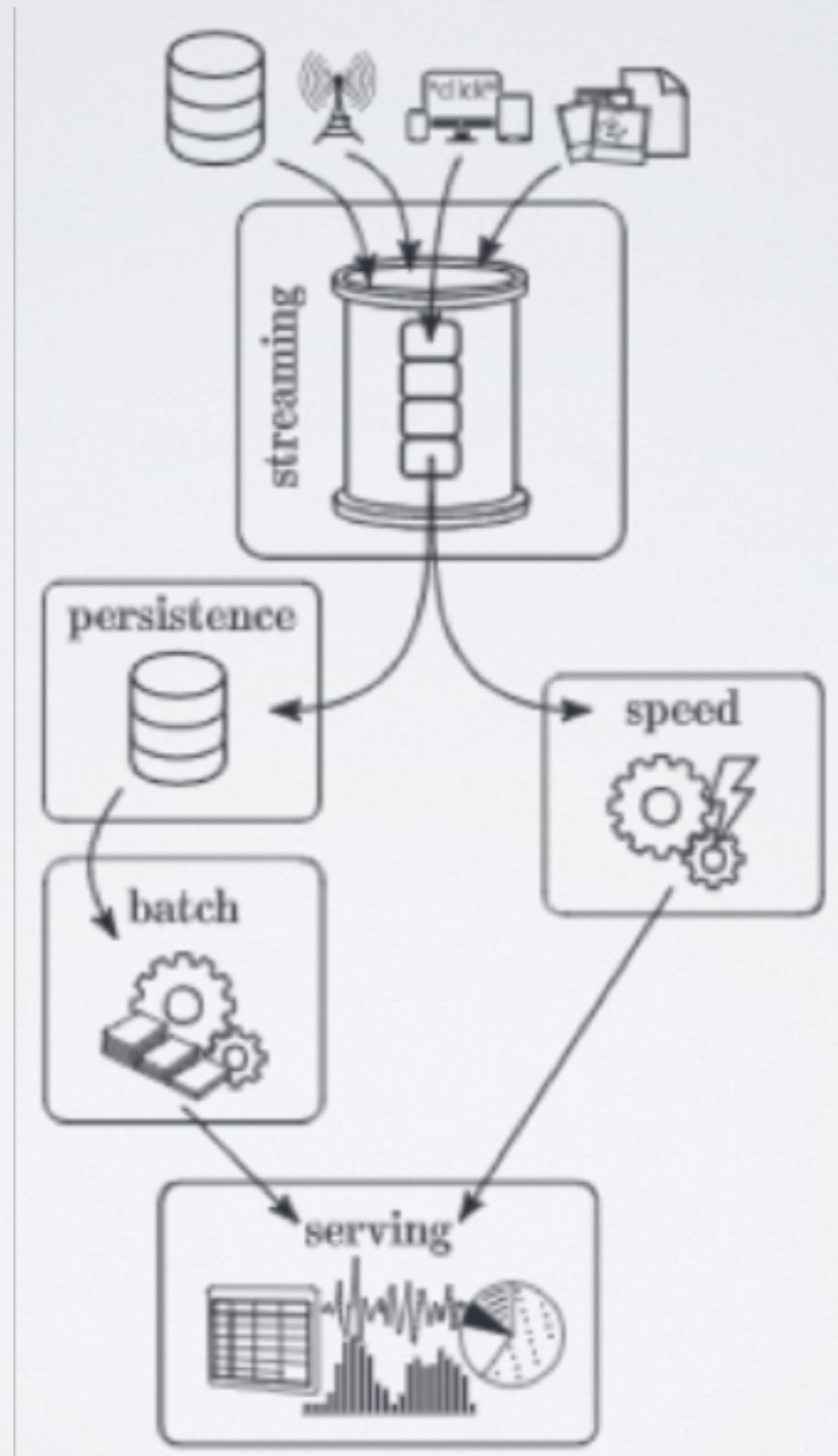
# Integration von GeoVisual Analytics:

- Lee, Jae-Gil and Minseo Kang:  
“Geospatial Big Data: Challenges and Opportunities.”(2015)



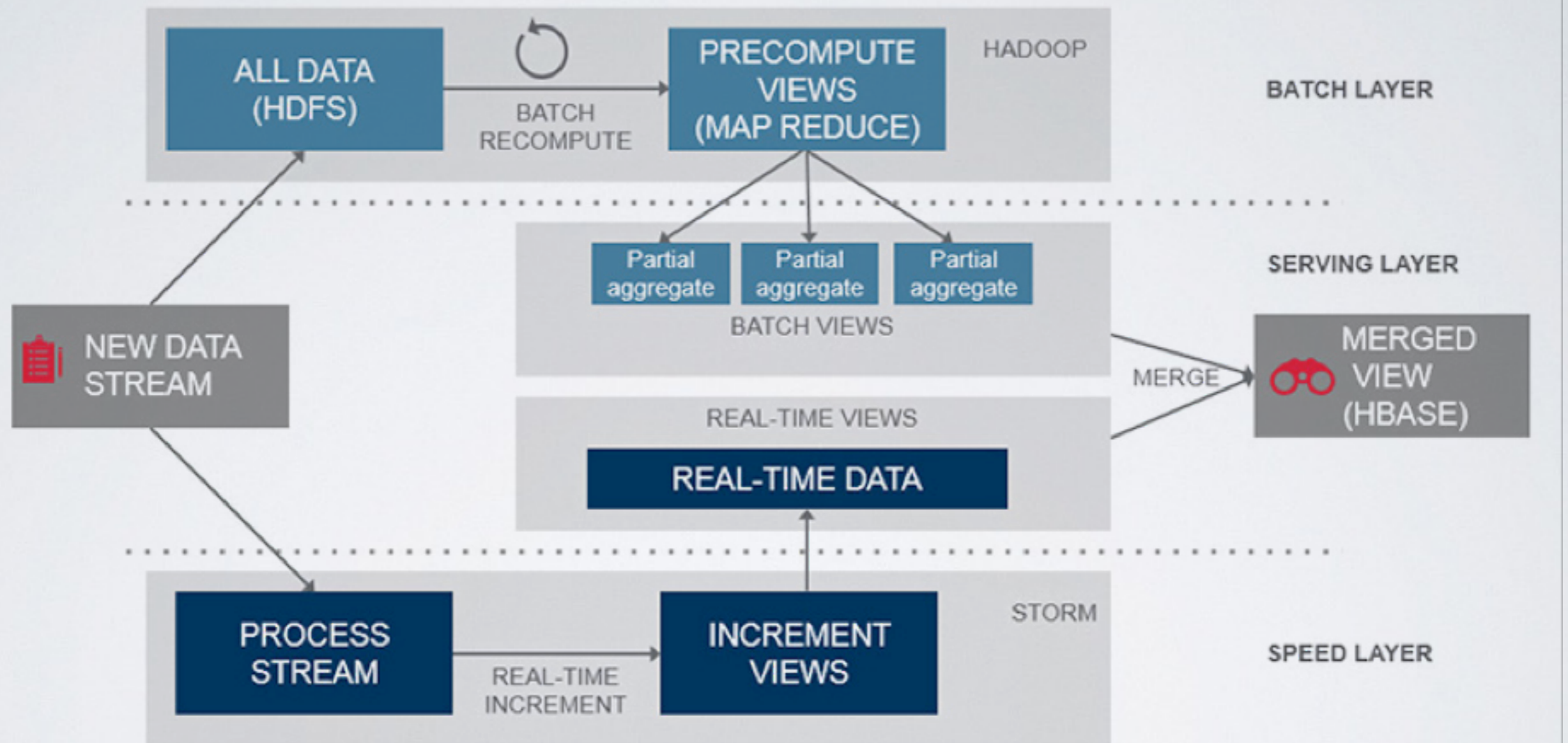


# Big Data: Lambda - Architektur



[WnWI6]

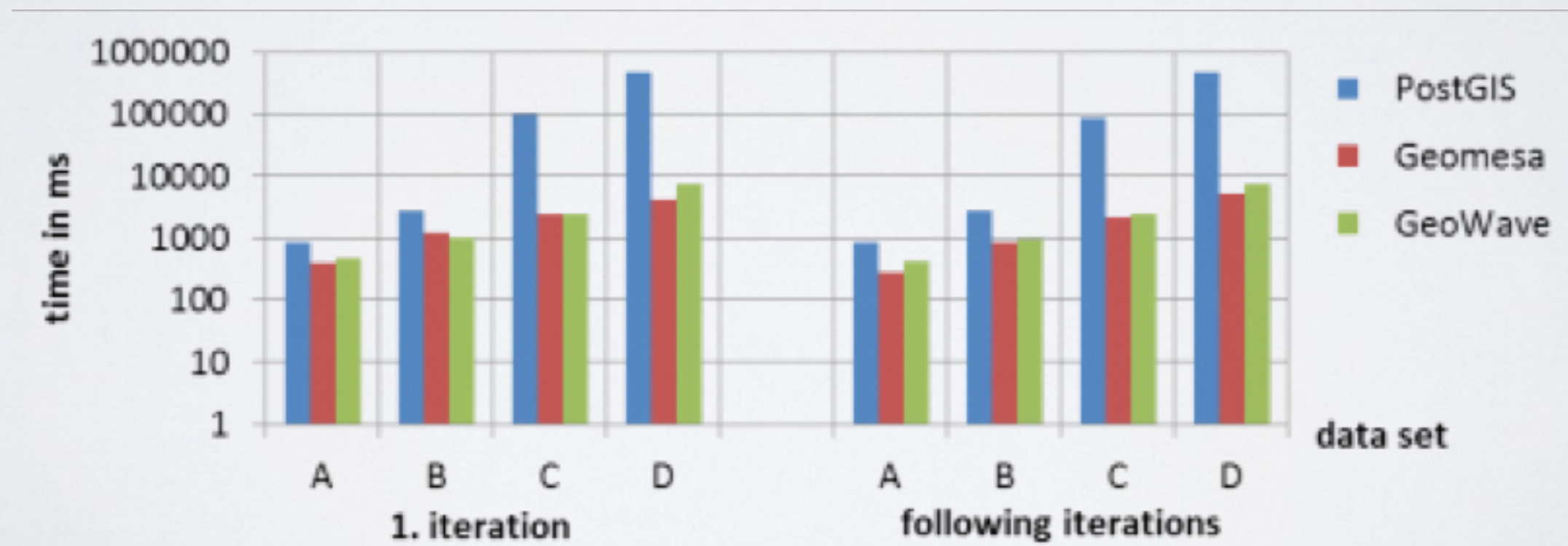
# Big Data: Lambda - Architektur



<http://www.mapr.com/sites/default/files/otherpageimages/lambda-architecture-2-800.jpg>

# Integration von GeoVisual Analytics: Spatiotemporale DB

- GeoMesa: “open-source, distributed, spatio-temporal database built on top of column stores” (Ähnliche: GeoWave, GeoTrellis, GeoJinni)
- Hadoop und Spark für verteilte Berechnungen [Era 15]
- Implementiert Geotools: Geospatiale Methoden
- Multi-dimensionale Indexierungsstrategien

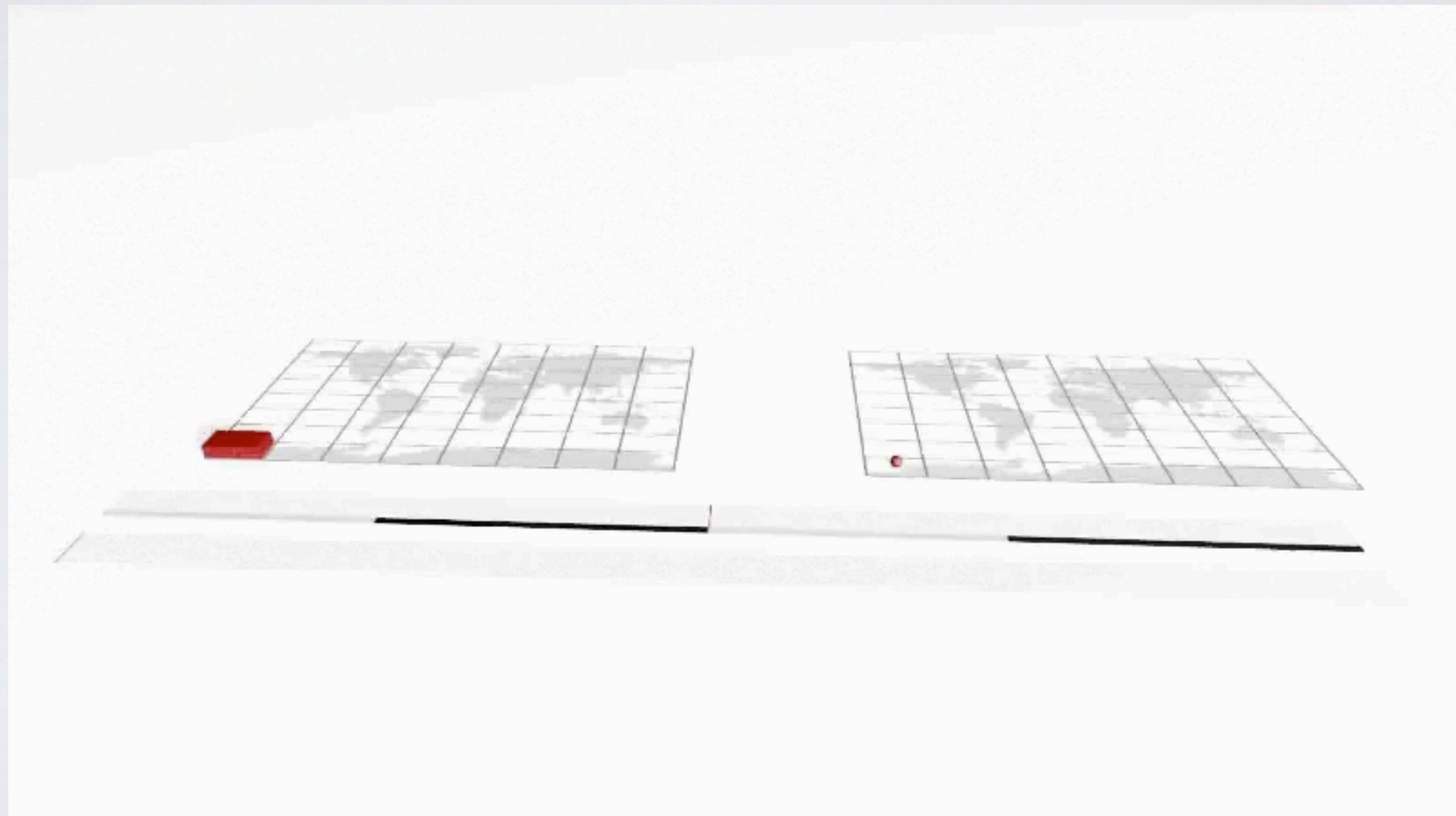


GeoBench - “Query 10: Find all events from 45 to 60 degrees of northern latitude and from 10 to 25 degree of east longitude during the first three weeks of June in 2015.”

# Integration von GeoVisual Analytics:



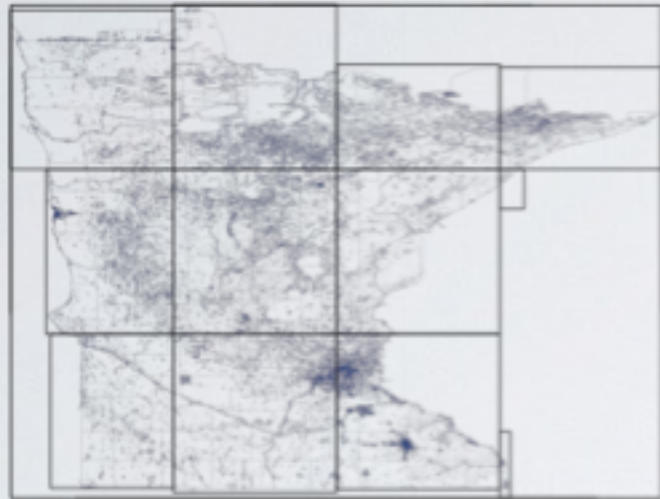
- Verschiedene datenbankgestützte Aggregationen
- Indexierung von temporalen Punkt- oder Polygondaten mit SFC:



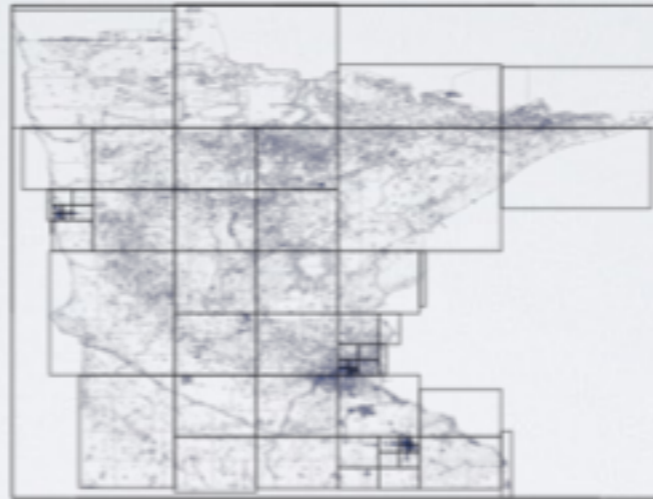
# Integration von GeoVisual Analytics:



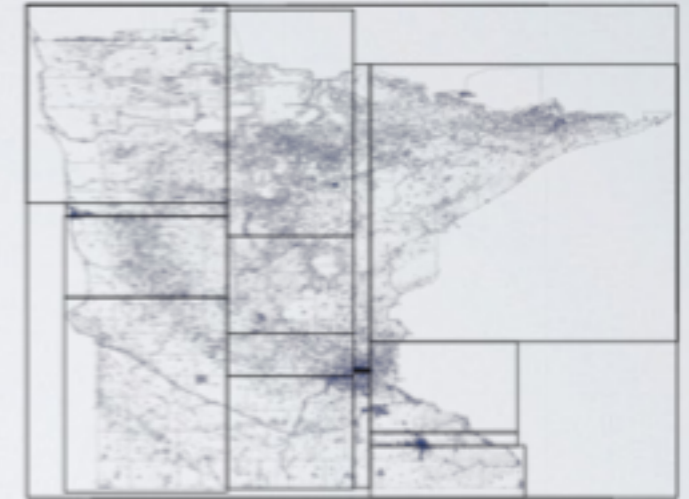
- Partitionierungstechniken (spatial, data, space filling curve)



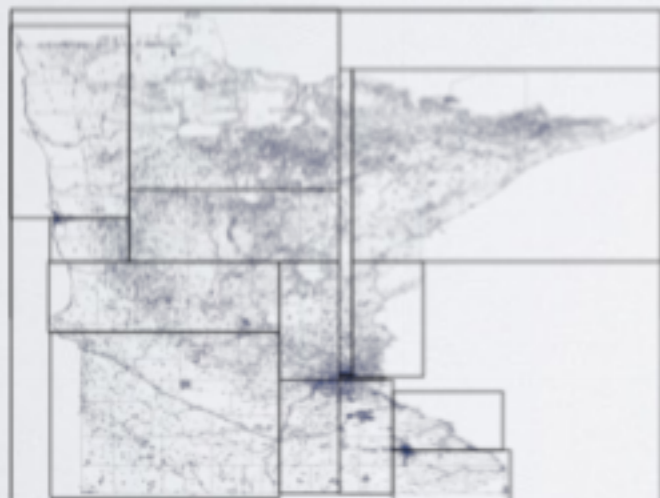
**Grid**



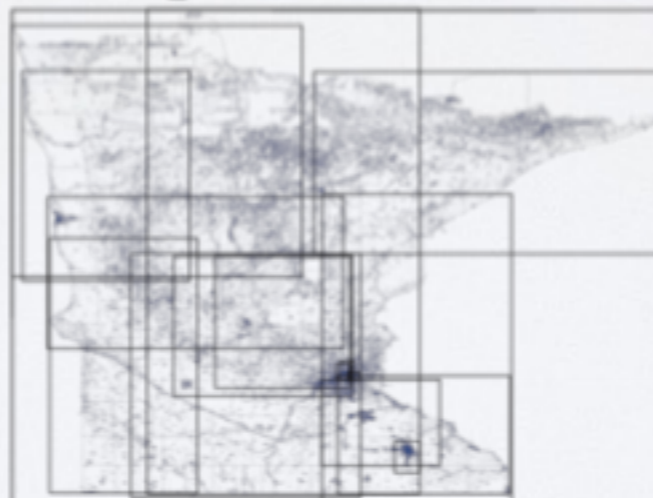
**Quad-tree**



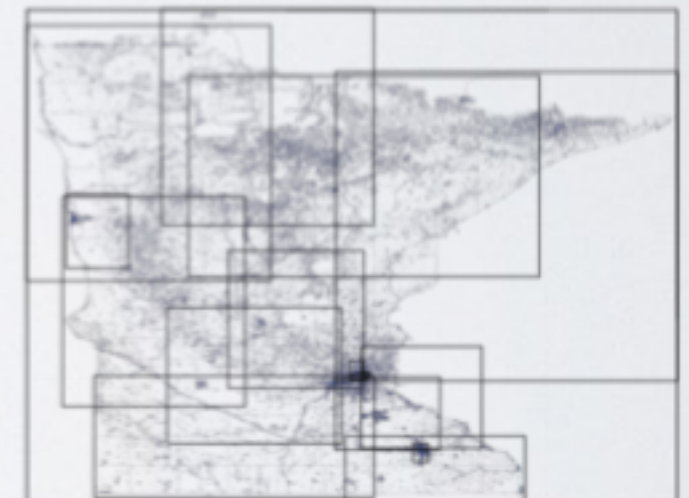
**STR/STR+**



**K-d Tree**



**Z-Curve**

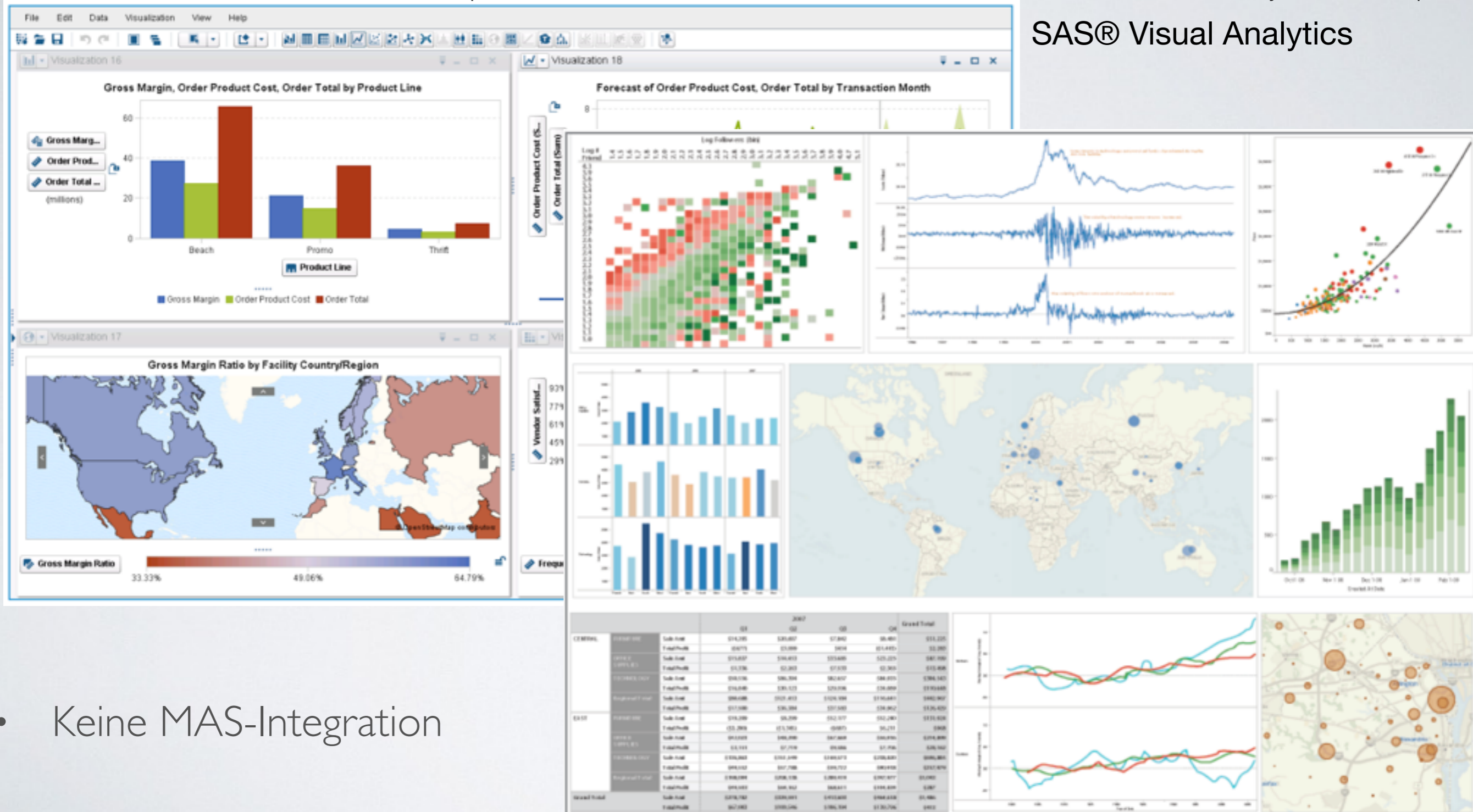


**Hilbert Curve**

# Kommerzielle Anbieter von Visual Analytics-Lösungen

[https://www.sas.com/content/dam/SAS/en\\_us/doc/factsheet/sas-visual-analytics-105682.pdf](https://www.sas.com/content/dam/SAS/en_us/doc/factsheet/sas-visual-analytics-105682.pdf)

SAS® Visual Analytics



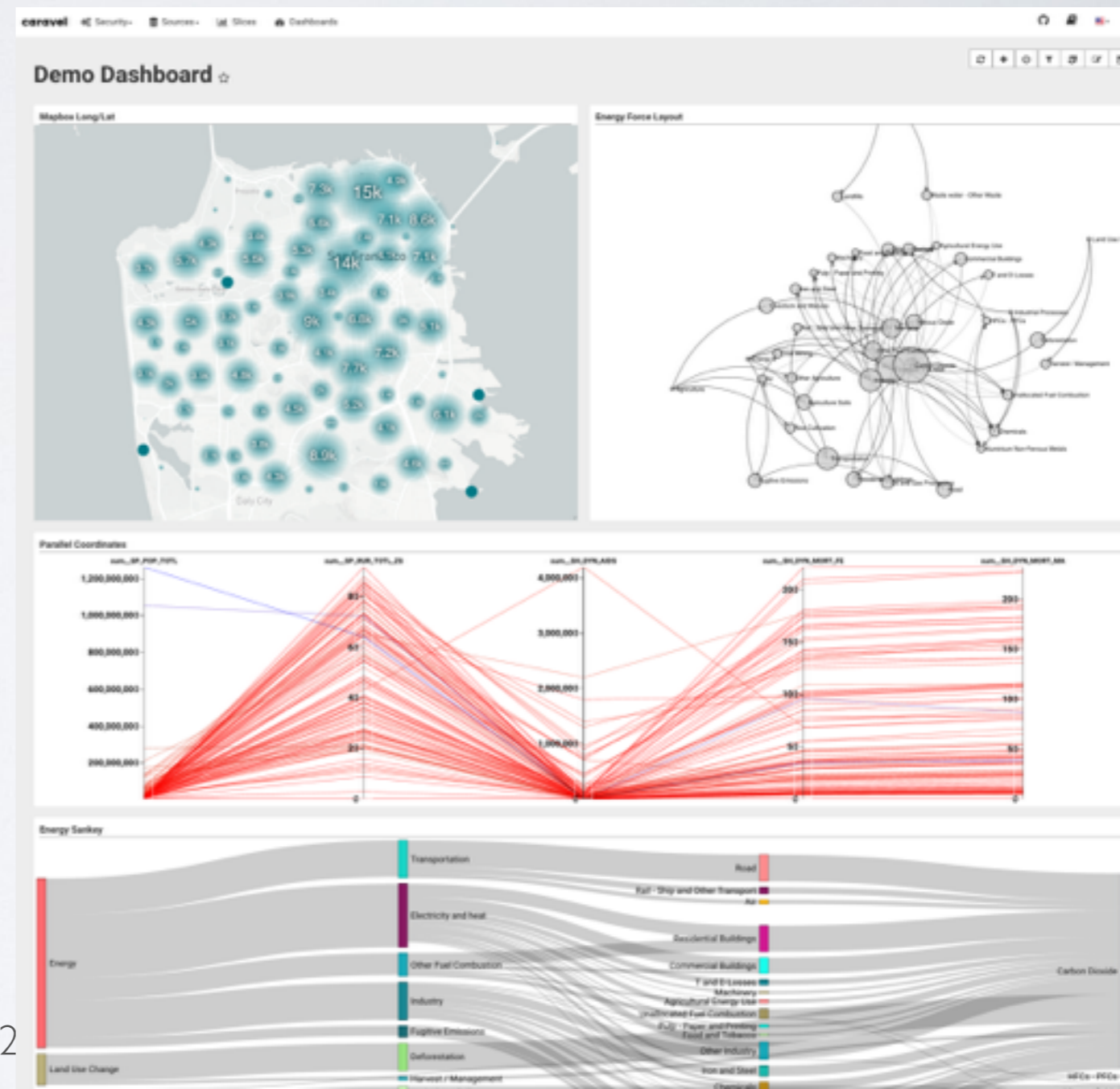
- Keine MAS-Integration

[https://www.tableau.com/sites/default/files/whitepapers/whitepaper\\_selecting-visual-analytics-application.pdf](https://www.tableau.com/sites/default/files/whitepapers/whitepaper_selecting-visual-analytics-application.pdf)

# Dashboard als Visual Analytics-Grundlage

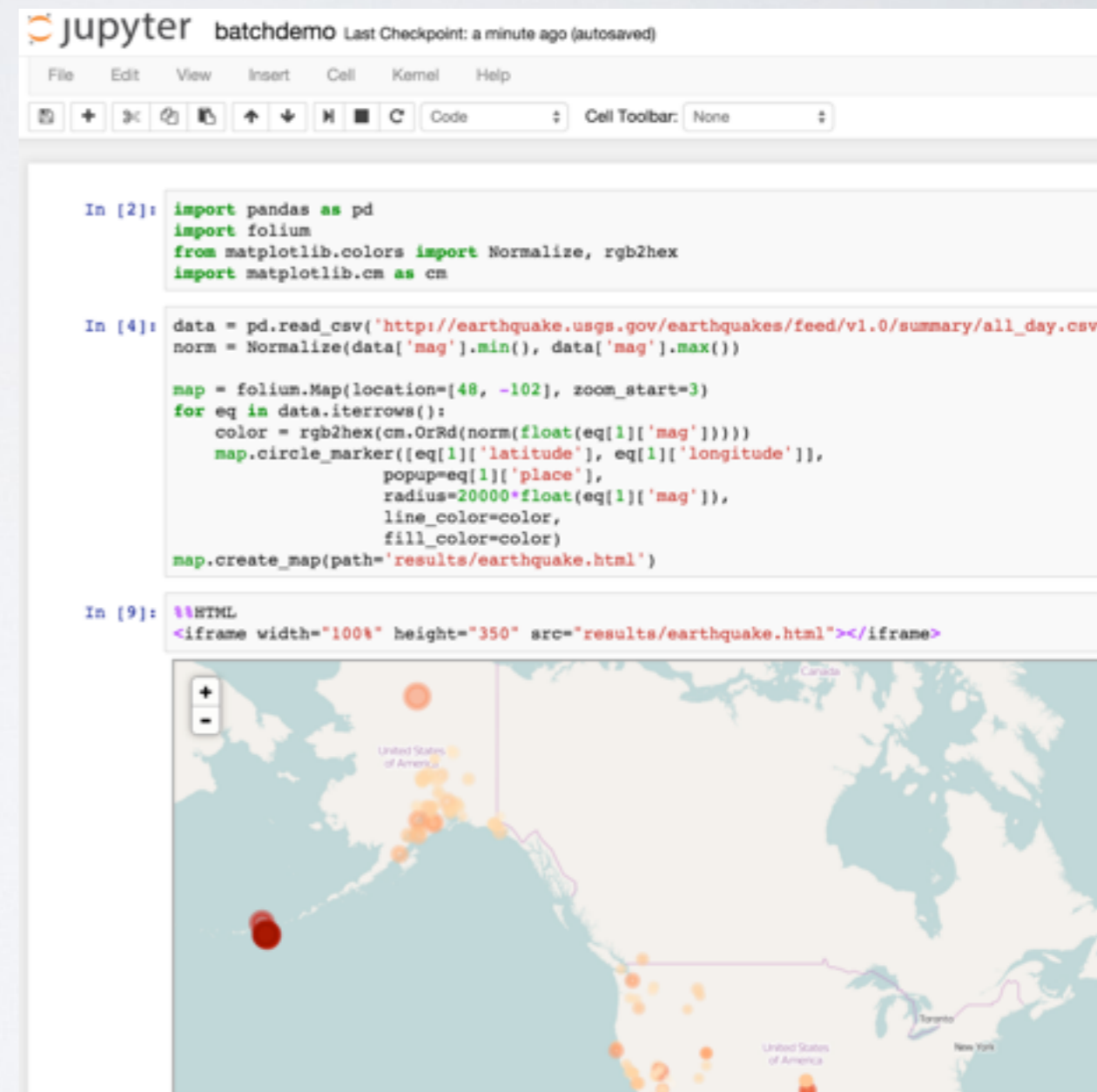
- Opensource-Alternativen:
  - Caravel (Superset) / Re:dash
  - Kibana / Grafana
  - Metabase / Apache Hue
- Datenbankkonnektoren
- Drag'n'Drop Visualisierungen

<http://conferences.oreilly.com/strata/hadoop-big-data-ny/public/schedule/stype/Keynote>



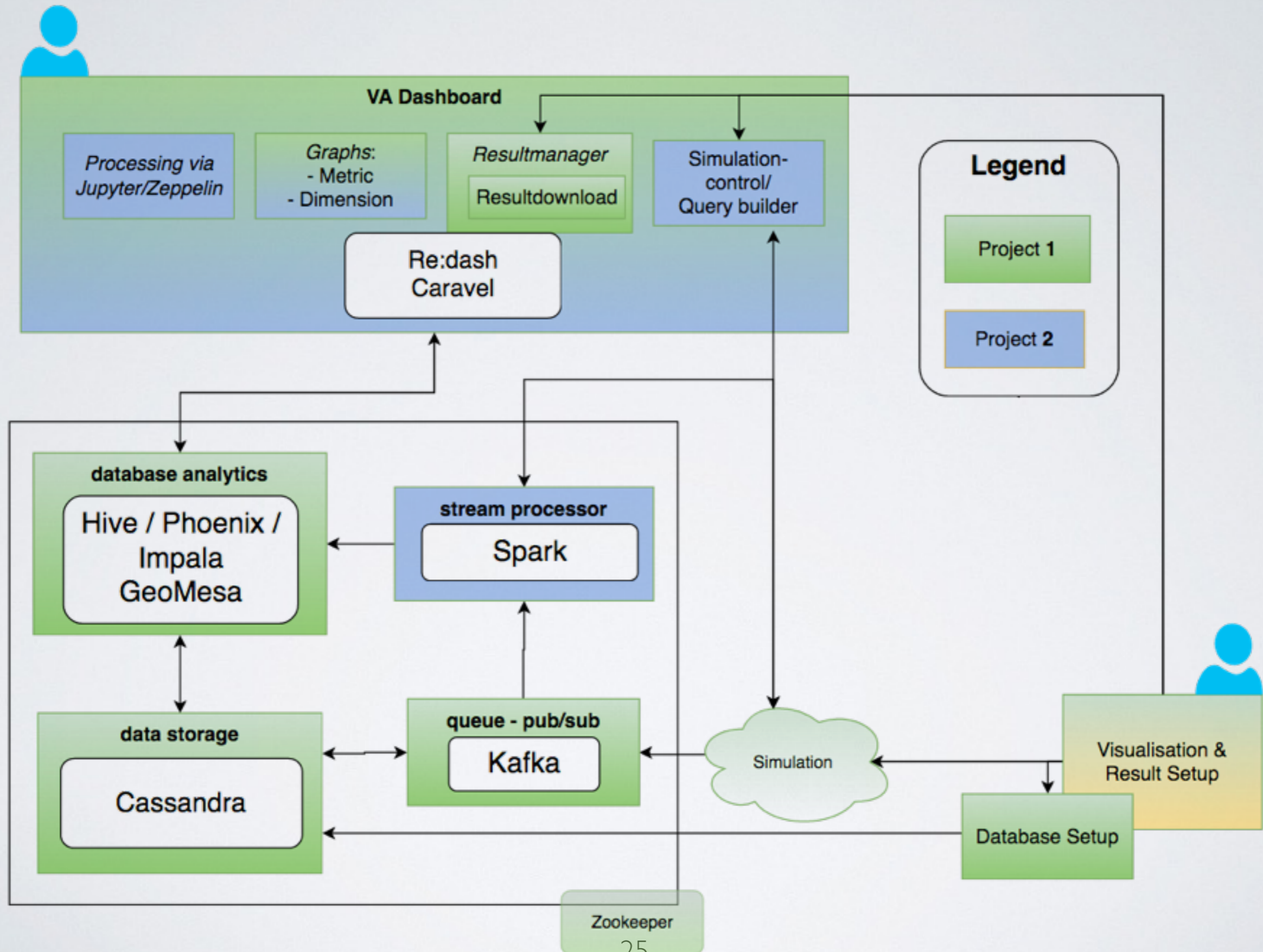
# Integration von benutzerdefinierten Analysen

- Zielgruppe: Domainexperten/Analysten
- Notebooks wie Apache Zeppelin & Jupyter bieten mit Interpretern:
  - Code aus dem Web-Editor serverseitig ausführen
  - Output anzeigen und mit Widgets verknüpfen
  - Einbinden in externe Seiten
- Abfragen auf den Datenbestand mit Presto, Hive o.ä.





# Angestrebte MARS Architektur in Hinsicht auf die VA-Plattform



## Nächste Schritte:

- PO 1 finalisieren:
  - Architektur mit der Simulation verknüpfen - ResultAdapter
- PO 2:
  - Optimierte Visualisierungen des Prototypens integrieren
  - Ausfallsicherheit/Zuverlässigkeit des Systems prüfen/messen (UI?)
  - Bereitstellung der Datenumgebung an die Notebooks
  - Einbinden der Notebooks mit verlinkten Parametern/Metriken in das Dashboard

# Chancen

- Simulationsergebnisse für verschiedene AnwenderInnen (Erfahrung, Hardware, Interesse)
- Einerseits visuelle Exploration der Simulationsdaten und andererseits die Möglichkeit Machine Learning Algorithmen zu benutzen
- Für jede generische Agentensimulationen eine intuitive Analyseoberfläche
- Nötige Performance für interaktive VA-Anwendungen
- Real-time Simulationsanalyse

# Thesis Outline - Forschungswürdige Aspekte

- Spezielle Ausrichtung der Multi-Agenten Modellierung
- “Real-time geospatial and temporal Visual Analytics for MAS”:
  - Ad-hoc Analysen
  - Level of Detail und Datenreduktion
  - Movement Analysis nach [Dem15]
- “Distributed, Parallel Event Recognition in Multi-Agent Systems”
  - Simulationskontrolle

# Risiken

- Orchestrierung - Zusammenspiel/Verlässlichkeit der Komponenten
- Datenaustausch zwischen den einzelnen Systemen
- Metrik für Qualität der Plattform, wann ist was wie gut erreicht
- Erwartung von Tableau usw. werden nie erfüllt werden können
- Spagat zwischen (Big Data-)Datenspeicherung- und GeoVisual Analytics-Problemen
- Simulationsdaten (Format / Zugang / Auswertung)
- Bedeutung/Kontext einer Multi-Agenten Simulationen

# Kooperationen und weiterer wissenschaftlicher Input

- Virginia Modeling, Analysis and Simulation Center (VMASC) - Old Dominion University
- University of Florida
- Netherlands eScience Center: Geo-Visualization
- IEEE Big Data, Summit Forum on Big Data Visualization, Large Data Analysis and Visualization (LDAV)
- FOSS4G
- Visually-supported Computational Movement Analysis (VCMA)

# Literatur

- [KeZi06] D.A. Keim, F. Mansmann, J. Schneidewind, H. Ziegler: „Challenges in visual data analysis“, in: Information Visualization (IV 2006), Invited Paper, July 5-7, London, United Kingdom, IEEE Press, 2006
- [MARS16] C.Huening et al. : „Modeling & Simulation as a Service with the Massive Multi-Agent System MARS“, ADS '16, Proceedings of the 2016 Spring Simulation Multiconference
- [RPS14] Simon J. E. Taylor, et al.: A tutorial on cloud computing for agent-based modeling & simulation with repast. In Proceedings of the 2014 Winter Simulation Conference (WSC '14). IEEE Press, Piscataway, NJ, USA, 192-206.
- [SuV12] V. Sucharitha, S.R. Subash and P. Prakash , Visualization of Big Data: Its Tools and Challenges, International Journal of Applied Engineering Research, 9(18), 2014, pp. 5277-5290.
- [inMems] Zhicheng Liu, Biye Jiang and Jeffrey Heer : imMems: Real-time Visual Querying of Big Data. Computer Graphics Forum (Proc. EuroVis), 32(3), 2013
- [CWA] Lee, Chee Wai, Celso Mendes, and Laxmikant V. Kalé. "Towards scalable performance analysis and visualization through data reduction." Parallel and Distributed Processing, 2008. IPDPS 2008. IEEE International Symposium on. IEEE, 2008.
- [Dem15] U.Demšar et al. Analysis and visualisation of movement: an interdisciplinary review, Movement Ecology 3(1) 2015.

# Literatur

- [CYG16] Yang, Can, and Gyozo Gidófalvi. "Interactive Visual Exploration of Most Likely Movements." Visually-supported Computational Movement Analysis (VCMA 2016), Helsinki, Finland
- [WnW16] Wingerath, Wolfram, et al. "Real-time stream processing for Big Data." *it-Information Technology* 58.4 (2016): 186-194.
- Johannes Schützel, Holger Meyer, and Adelinde M. Uhrmacher. 2014. A stream-based architecture for the management and on-line analysis of unbounded amounts of simulation data. In *Proceedings of the 2nd ACM SIGSIM Conference on Principles of Advanced Discrete Simulation (SIGSIM PADS '14)*. ACM, New York, NY, USA, 83-94
- Daniel Zehe, Vaisagh Viswanathan, Wentong Cai, and Alois Knoll. 2016. Online Data Extraction for Large-Scale Agent-Based Simulations. In *Proceedings of the 2016 annual ACM Conference on SIGSIM Principles of Advanced Discrete Simulation (SIGSIM-PADS '16)*. ACM, New York, NY, USA, 69-78
- J. Borgdorff, W. van Hage, L.J. Dijkstra, E. Mancini and M.H. Lees, Simulation-supported Urban Movement Analysis. *Visually-supported Computational Movement Analysis (VCMA 2016)*, Helsinki, Finland
- Lee, Jae-Gil and Minseo Kang. "Geospatial Big Data: Challenges and Opportunities." *Big Data Research* 2 (2015): 74-81.
- Zhenqiang Li, Xuefeng Guan, Rui Li and Huayi Wu. "4D-SAS: A Distributed Dynamic-Data Driven Simulation and Analysis System for Massive Spatial Agent-Based Modeling." *ISPRS Int. J. Geo-Inf.* 2016, 5, 42
- [mjad12] M. Rak, A. Cuomo and U. Villano, "mJADES: Concurrent Simulation in the Cloud," 2012 Sixth International Conference on Complex, Intelligent, and Software Intensive Systems, Palermo, 2012, pp. 853-860.



Fragen?

Vielen Dank für Eure Aufmerksamkeit

Thank you for your attention!



<http://factoflife.net/upload/images/20160611/funny-cute-elephant.jpg>