

DKG 2015



EBERHARD KARLS
**UNIVERSITÄT
TÜBINGEN**



Theoretische Überlegungen und Beispiele einer „Spatial Data Science“

Peter Mandl

Institut für Geographie und Regionalforschung,
Universität Klagenfurt

Hans-Joachim Rosner

Forschungsbereich Geographie, Universität Tübingen

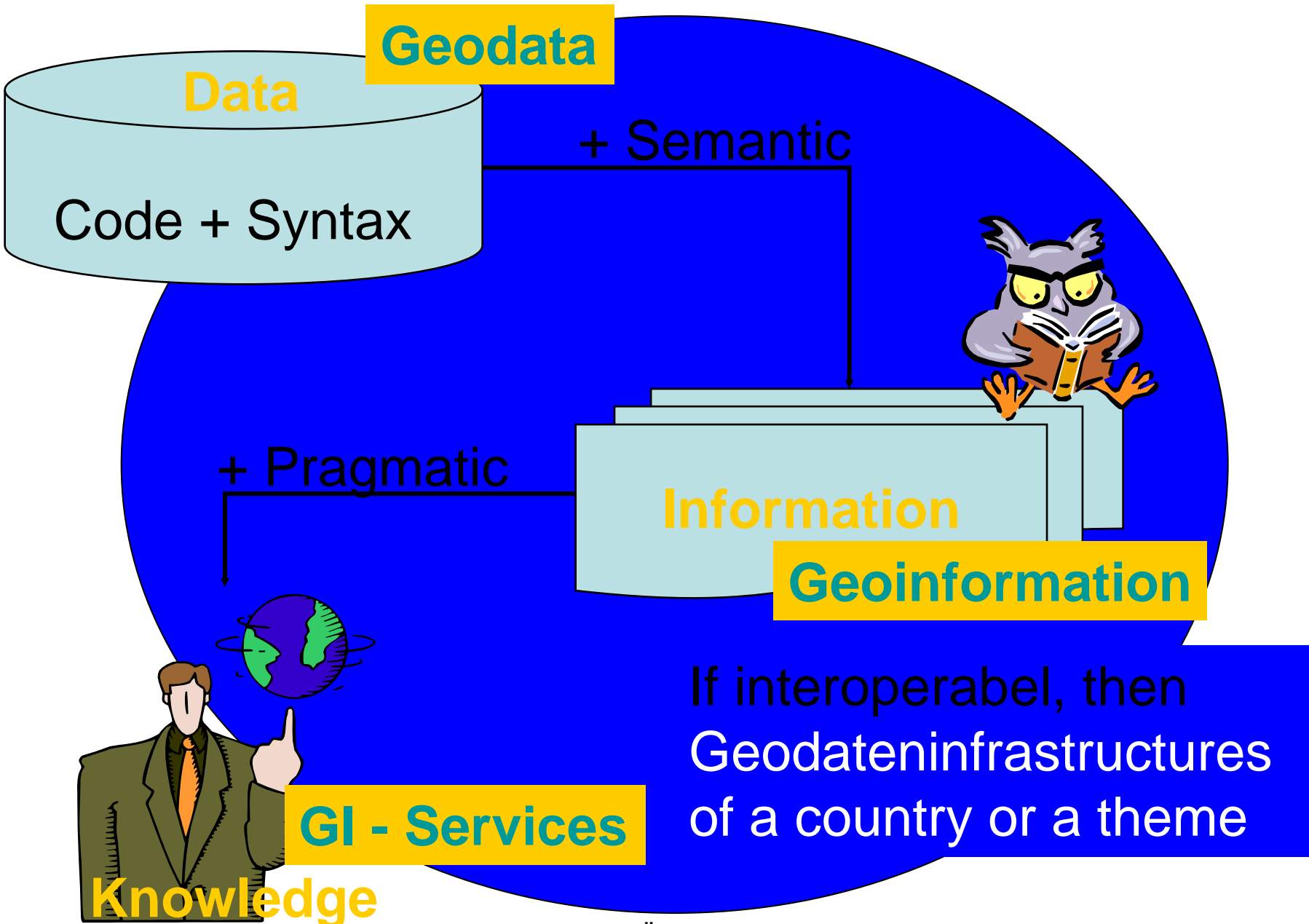
DKG 2015, Berlin, 01.10.2015

VE-FS-14 Vernetzung und Entwicklungspfade



Ablauf

- **Einführung:** Entwicklungspfade zu einer „Spatial Data Science“ (9:00-9:10)
- **Vorstellungsrunde** (9:10-9:15)
- **Vorträge:** *Tina Buchmann und Jürgen P. Kropp* (9:15-9:40) und *Diego Rybski, Bin Zhou* (9:40-10:05)
- **Pause** (10:05-10:20)
- **Vorträge:** *Knut Haase, Mathias Kasper, Matthes Koch, Sven Müller* (10:20-10:45) und *Tim Elrick* (10:45-11:10)
- Abschließende **Diskussion** über die Vision und die Probleme einer „Spatial Data Science“ (11:10-11:30)



Development of Quantitative Geography

- 1970ies: Spatial or Locational Analysis (P. Haggett)
- 1980ies (Ger): Raumwissenschaftlicher Ansatz (D. Bartels)
- Quantitative Geography
- Geographical Systemstheory
- 1990ies: GeoComputation (Stan Openshaw)
- Computational Geography
- 2000s: Geo-Simulation, Evolutionary Approaches, ...

The Hype „Big Data“

BARNES & NOBLE Books for Everyone on Your List: Shop Now







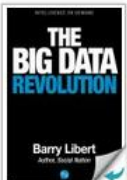





















Shopping Bag (0 items) Spend \$25. Get FREE SHIPPING

All Products Search

Newsstand | Teens | Kids | Toys & Games | Home & Gifts | Movies & TV | Music | Gift Cards | Bargain Books

160 results for big data in All Products.

atches View: 30 Page 1 of 6

 <p>Big Data Bibliography Safari Content Team NOOK Book FREE</p>	 <p>Big Data Marketing: Engage... Lisa Arthur Hardcover \$21.68</p>	 <p>Planning for Big Data Edd Dumortier NOOK Book FREE</p>	 <p>Big Data Now: 2012 Edition Inc. O'Reilly Media NOOK Book FREE</p>	 <p>Big Data Rx Bibliography Safari Content Team NOOK Book FREE</p>	 <p>Big Data: Principles and best... Nathan Marz Paperback \$32.61</p>					
 <p>The Big Data Revolution Barry Libert NOOK Book \$2.99</p>	 <p>The Human Face of Big Data Rick Smolan Hardcover \$36.95</p>	 <p>The Realities of Securing Big Data David Ottenheimer Paperback \$21.65</p>	 <p>Real-Time Big Data Analytics: Emerging Architecture Mike Barlow NOOK Book FREE</p>	 <p>Enterprise Analytics: Optimization... Thomas H. Davenport Hardcover \$40.43 NOOK Book \$28.49</p>	 <p>Too Big to Ignore: The... Phil Simon Hardcover \$33.95</p>	 <p>Big Data: Understanding How... Bill Sommerzo Paperback \$27.60 NOOK Book \$22.99</p>	 <p>Loyalty 3.0: How to... Rajat Panaria Hardcover \$19.34 NOOK Book \$14.99</p>	 <p>Big Data For Dummies Judith Hurwitz Paperback \$20.96</p>	 <p>Microsoft Big Data Solutions Adam Jorgensen Paperback \$36.99</p>	 <p>Big Data Kevin Roebuck Paperback \$39.95</p>
 <p>Big Data Glossary Pete Warden Paperback \$18.01 NOOK Book \$10.49</p>	 <p>Principles of Big Data... Julius J. Berman Paperback \$57.77 NOOK Book \$64.95</p>	 <p>IP, Big Data, and Society... David Houle NOOK Book \$0.99</p>	 <p>Transforming Business: Big... Allison Cerra Paperback \$16.62 NOOK Book \$13.99</p>	 <p>Privacy and Big Data Terence Craig Paperback \$18.01 NOOK Book \$10.49</p>	 <p>Pig In Action: Munging Big Data M. Tim Jones Paperback \$33.29</p>	 <p>Big Data Analytics: Turning... Frank J. Ohlhorst Hardcover \$32.99 NOOK Book \$28.49</p>	 <p>Big Data and Hadoop Kalott Jain NOOK Book \$0.99</p>	 <p>Big Data at Work: Dispelling... Thomas H. Davenport Hardcover \$19.17</p>	 <p>Big Data, Big Analytics... Michael Minelli Hardcover \$32.24</p>	 <p>The Culture of Big Data Mike Barlow NOOK Book FREE</p>

What are BIG DATA? (Definition 2012)

- Very **big amounts of data**, Data Deluge
- **Measured** by sensors etc., **created** by computer programs or by humans
- **structured** (numbers, tables) or **unstructured** (images, videos, texts, websites)

We are interested especially in **BIG Geo-DATA**

<http://www.google.org/flutrends/at/#AT>

The screenshot shows the Google Flu Trends website for Austria. The page title is "Grippe-Trends analysieren - Österreich". The main content area features a line graph titled "Landesweit" showing flu trends from July to June for the 2013-2014 season and previous years. The y-axis ranges from "minimal" to "sehr hoch". Below the graph is a map of Austria with regions colored in shades of green and yellow, indicating flu activity levels. The right sidebar contains sections for "Grippe bekämpfen" (flu prevention tips) and "Animierte Grippetrends in Google Earth".

Grippe-Trends analysieren - Österreich

Google hat festgestellt, dass die Häufigkeit bestimmter Suchbegriffe Anhaltspunkt für die Häufigkeit von Grippefällen sein kann. Für die Google Grippe-Trends werden Daten der Google-Suche gesammelt und ausgewertet. Auf Grundlage der Ergebnisse wird anschließend die Häufigkeit von Grippefällen geschätzt. [Weitere Informationen >](#)

Landesweit ● 2013-2014 ● Vorjahre ▼

sehr hoch
hoch
mittel
niedrig
minimal

Jul Aug Sep Okt Nov Dez Jan Feb Mär Apr Mai Jun

Die Schätzungen wurden auf Grundlage eines Modells erhoben, das mit offiziellen historischen Daten zur Grippe-Häufigkeit abgeglichen und als korrekt befunden wurde. Die Daten sind bis 18. November 2013 aktuell.

Grippe bekämpfen

Grippe wird durch Husten und Niesen übertragen. Mit den folgenden drei einfachen Maßnahmen können Sie das Ansteckungsrisiko mindern:

1. Husten/Niesen Sie in die Armbeuge oder in ein Taschentuch.
2. Waschen Sie sich öfters die Hände.
3. Bleiben Sie zuhause, wenn Sie sich krank fühlen.

Ziehen Sie bei Bedarf einen Arzt zu Rate.

[European Influenza Surveillance Network](#)

Animierte Grippetrends in Google Earth

Laden Sie [hier](#) Daten von Google Grippe-Trends in Google Earth herunter und untersuchen Sie sie. Sie haben noch kein Google Earth? [Laden Sie es hier herunter.](#)

Tabelle einbetten

Verwenden Sie [diesen Einbettungscode](#), wenn Sie die Tabelle in Ihre Website integrieren möchten.

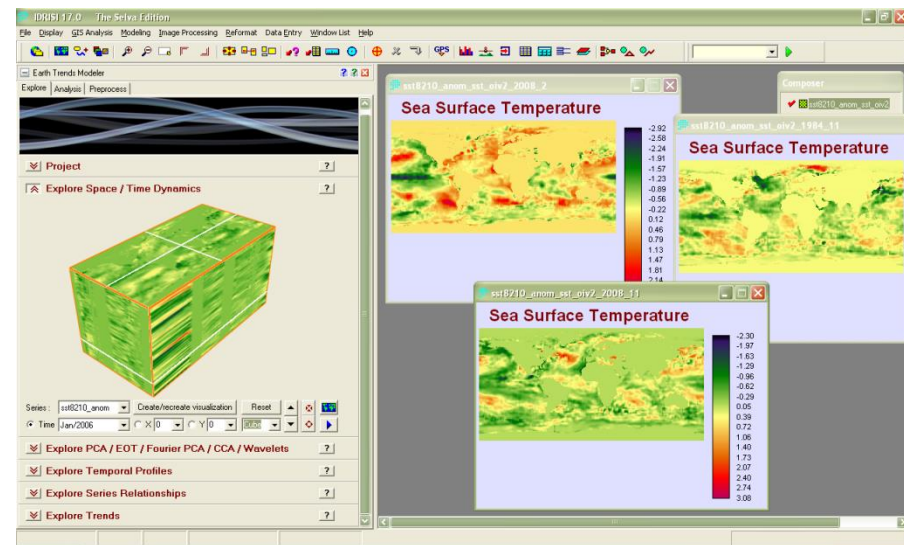
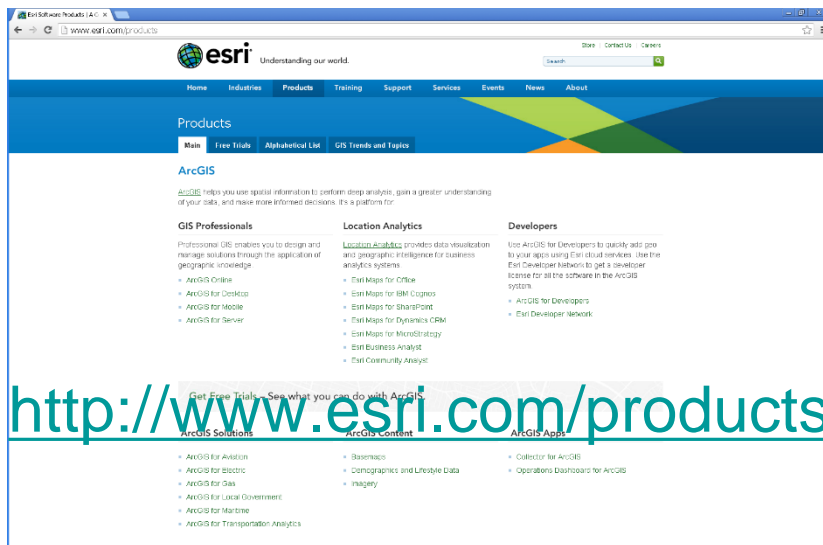
© 2011 Google - [Google.org - Startseite](#) (auf Englisch) - [Nutzungsbedingungen](#) - [Feedback geben](#)

BIG DATA are today (Definition 2015)

- A valuable and ubiquitous **datapool**
- **Often** useable and **combinable**
- Suitable for **complex** queries and analyses
- Mostly similar to a **population** (not a sample)
- **Blurred**
- Highly **up-to-date**
- Usable for **inductive model building**, for **prediction** and **prescription**
- Primarily useful to answer „**what**“ and less „why“ questions

Today: Location Analytics

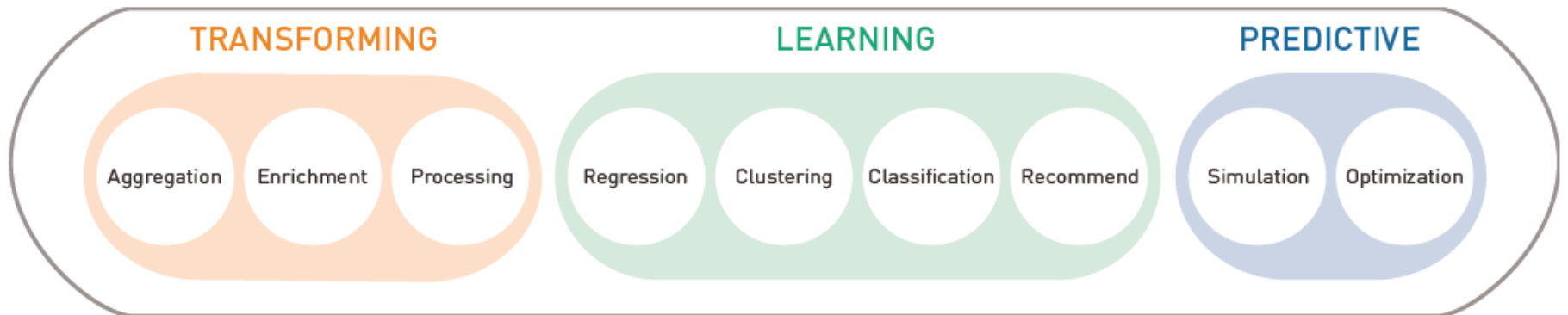
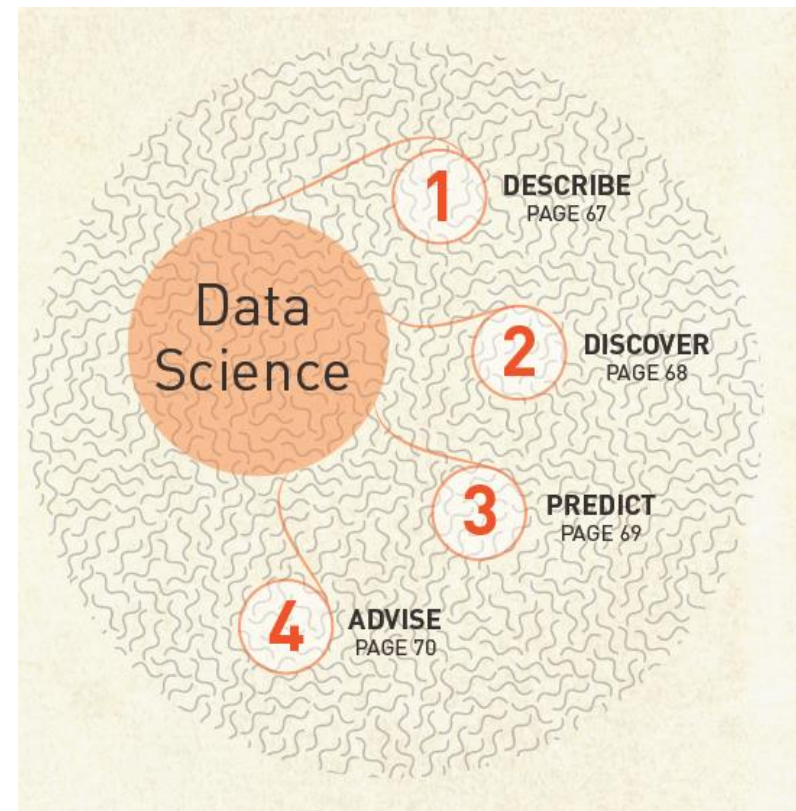
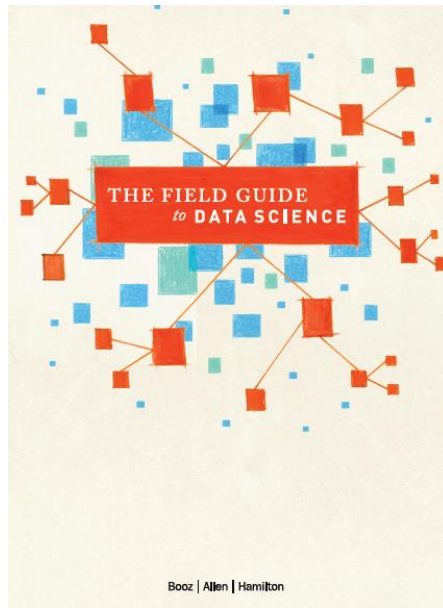
- Explorative spatial data analysis
- Visual Analytics
- Spatial Data Mining
- Multivariate Quantitative Geography (e.g. Crime Analysis, Spatial Econometrics, Spatial Data Analysis)
- Locational analysis
- Space/Time Analysis (e.g. Earth Trends Modeller)



Terms

- Analytics, Forecasting
- Descriptive, predictive, prescriptive analytics and modelling
- Data Warehouse
- Webintelligence, Business Intelligence
- Knowledge discovery
- Data Science

Methods of a „Data Science“



Source: Booz Allen Hamilton

Classes of Analytic Techniques

Cross Industry Standard Process for Data Mining (CRISP-DM)

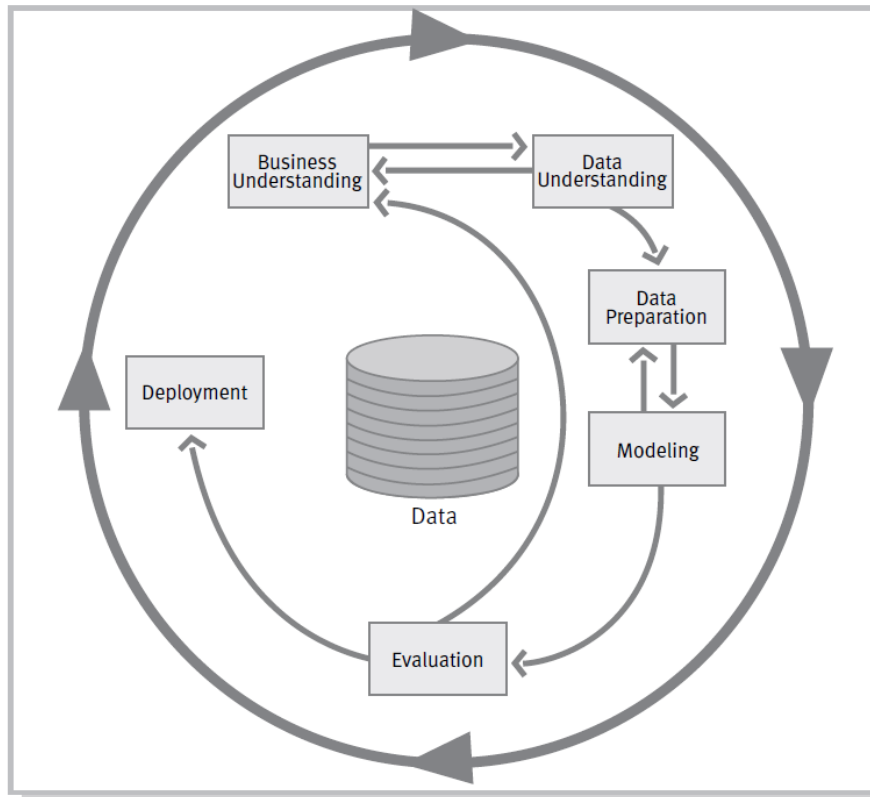


Figure 2: Phases of the CRISP-DM reference model

Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer and Rüdiger Wirth (2000): **CRISP-DM 1.0 Step-by-step data mining guide**. SPSS Inc.

To a Data-Driven Geography?

- In a Spatial Data Science we have to consider or think about
 - Errors
 - Privacy protection
 - Prediction without explanation?
- **Interoperability** of the programs, data, models and methods is to be established.
- The **Big Data** have to be used **multiple** and have to be **combined** and **linked**.
- Many **Best Practice Examples** have to be done and the **models** have to be **distributed** and **offered** via internet **interoperably**.
- Use a **formal language** or **ontology** to describe the new approach of **Spatial Data Science**.

Additional Material

Article:

- Harvey J. Miller & Michael F. Goodchild (2014):
Data-driven geography. In: GeoJournal
<http://link.springer.com/article/10.1007/s10708-014-9602-6>

> Big Data and Geospatial Analysis

<http://www.spatialanalysisonline.com/HTML/index.html?afterword.htm>



Vortrag

Tina Buchmann

Ermittlung von hochaufgelöster Globalstrahlung aus Ertragsdaten von Photovoltaikanlagen – Nutzung von „Big Data“

Zur Ermittlung von Globalstrahlung aus Ertragsdaten von Photovoltaikanlagen ist die Bearbeitung großer Datensätze eine Herausforderung. Das methodische Vorgehen und die Umsetzung werden vorgestellt.



Vortrag

Jürgen P. Kropp, Diego Rybski, Bin Zhou

Assessment of the urban heat island effect in Europe by integrating satellite-based observation and land cover map

Earth observation satellites provide essential information of global coverage and high spatial resolution application domains. Their enormous data volumes generated impose high demands on efficient analytics.



Vortrag

Knut Haase, Mathias Kasper, Matthes Koch, Sven Müller

A Pilgrim Scheduling and Routing Approach to Increase Public Safety during the Hajj

Stampedes (Hajj, 2006 or Duisburg Love Parade, 2010) are horrendous crowd disasters. We present a unique crowd management system to solve the largest pedestrian problem in the world (Hajj, 3 million pilgrims per year).



Vortrag

Tim Elrick

Selbstreferentielles Spatial Data Mining - Auf der Suche nach Indikatoren für Forschungsqualität in der Humangeographie

Forschungsqualität misst sich auch am Grad des internationalen Austausches der wissenschaftlichen Akteure untereinander. Diesem soll anhand von netzwerkanalytischen Betrachtungen nachgespürt werden.



Ablauf

- **Einführung:** Entwicklungspfade zu einer „Spatial Data Science“ (9:00-9:10)
- **Vorstellungsrunde** (9:10-9:15)
- **Vorträge:** *Tina Buchmann und Jürgen P. Kropp* (9:15-9:40) und *Diego Rybski, Bin Zhou* (9:40-10:05)
- **Pause (10:05-10:20)**
- **Vorträge:** *Knut Haase, Mathias Kasper, Matthes Koch, Sven Müller* (10:20-10:45) und *Tim Elrick* (10:45-11:10)
- Abschließende **Diskussion** über die Vision und die Probleme einer „Spatial Data Science“ (11:10-11:30)



Vielen Dank für Ihre Aufmerksamkeit!

Fragen und Vorschläge bitte an:

peter.mandl@aau.at

hans-joachim.rosner@uni-tuebingen.de

Informationen unter:

<http://giscience.aau.at/>

<http://wwwg.uni-klu.ac.at/geo/tqmgeo/>