

DKG 2015



Data, Concepts and Models of a Spatial Data Science by Linking Geography, GI-Science and Data Analytics

Peter Mandl

Institut für Geographie und Regionalforschung,
Universität Klagenfurt

DKG 2015, Berlin, 02.10.2015

RE-FS-14 Theoretische, methodische und
disziplinäre Reflexion

Data, Concepts and Models of

- **Geography:** (big) geo-data, concepts, models, location analytics, spatial data mining, visual analytics, ...
- **GI-Science:** semantic web, linked open data, GIS-modeling, ...
- **Data Analytics:** methods from quantitative geography extended by visual, predictive, prescriptive and explorative aspects, business analytics, data science

(Big) Geo-Data today

- Books, **literature** data bases
- Digital **maps** and web maps, media repositories (images, films, videos)
- Volunteered (VGI) or **citizen** data
- **Statistical** data, geodata-repositories
- **Remote sensing** data, (real time) **sensor data**
- **Communication** data and metadata
- **Personalized data** (self tracking data, shopping cards, social networks, ...)
- **Geodatainfrastructures**, GeoWeb,
- ...

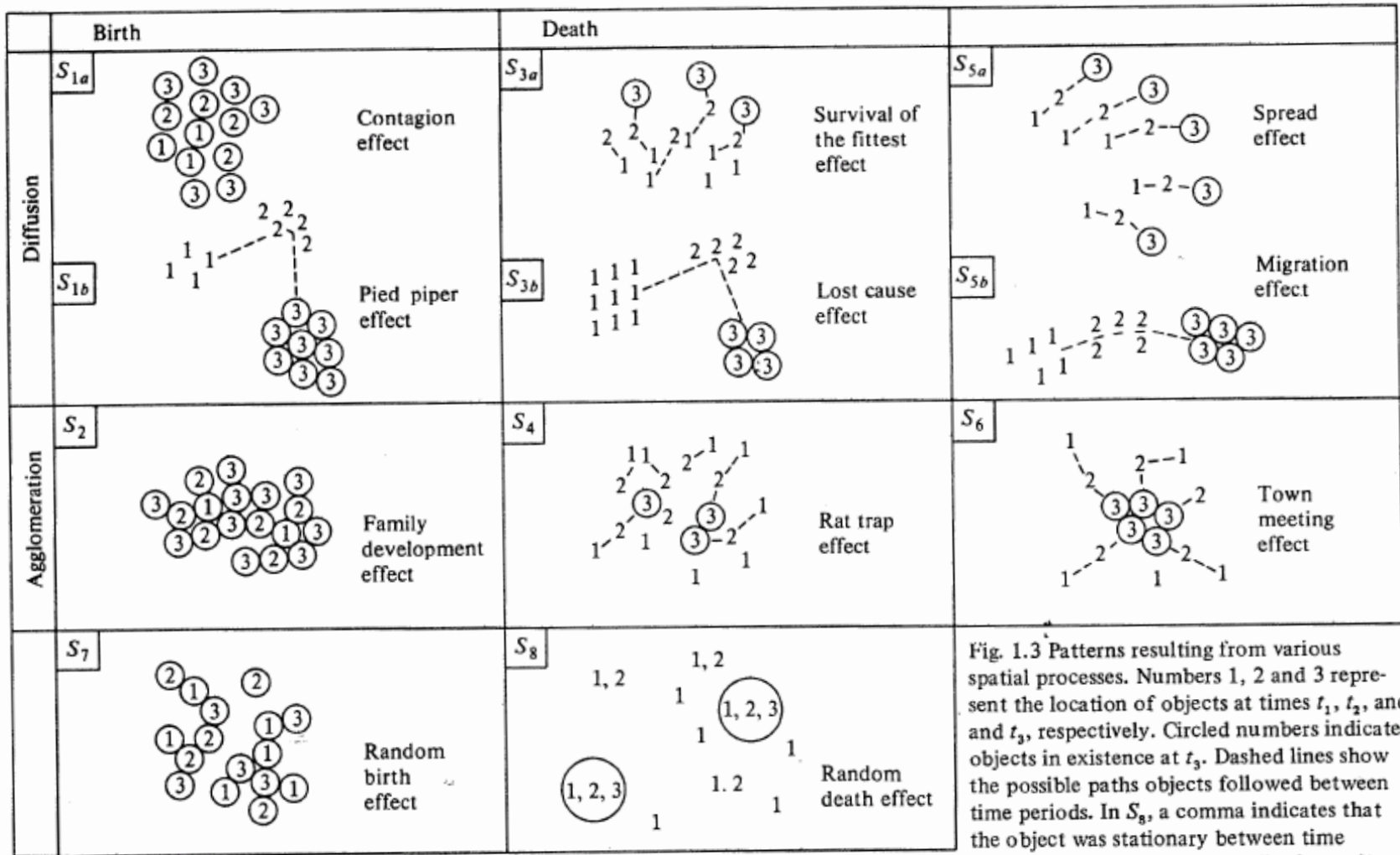
Most having **spatial reference** (locational information, addresses, geotags), **time reference** and **cross-linking**

BIG DATA are today (Definition 2015)

- A valuable and ubiquitous **datapool (structured, unstructured,**
- **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

Modelling categories in (Human-) Geography

- **Description** (visually, statistically, Spatial Data Mining)
- **Classification** (supervised, unsupervised, ...)
- **Location / Allocation** (sites, paths, hinterländer, ...)
- **Assessment** (MADM, MODM, ...)
- **Interaction** (Wilson models)
- **Diffusion** (Markov, differential equations, CA)
- ...

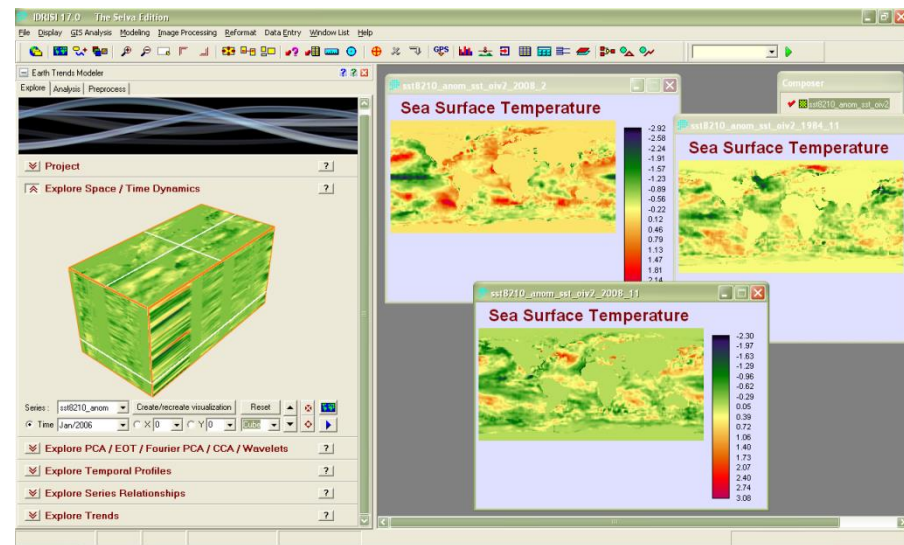
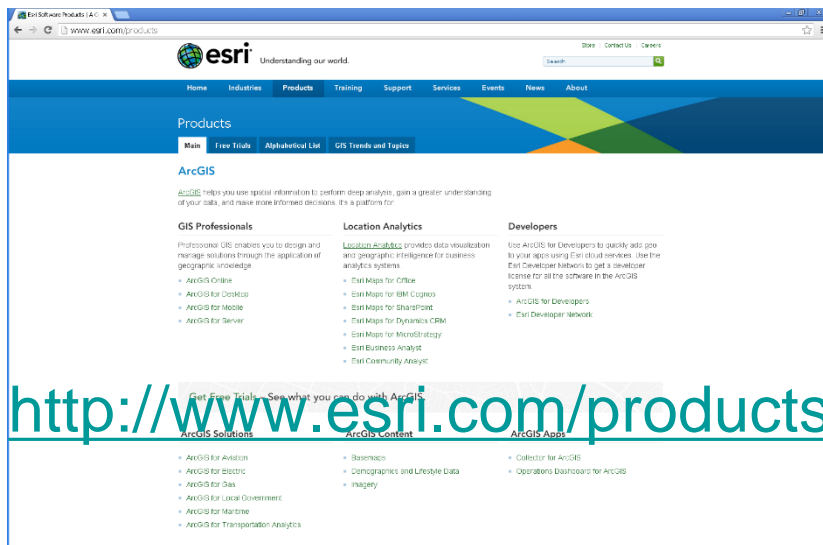


Getis & Boots 1978: Models of Spatial Processes



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)



Data, Concepts and Models of

- **Geography:** (big) geo-data, concepts, models, location analytics, spatial data mining, visual analytics, ...
- **GI-Science:** semantic web, linked open data, GIS-modeling, ...
- **Data Analytics:** methods from quantitative geography extended by visual, predictive, prescriptive and explorative aspects, business analytics, data science

HarmonISA Project

harmonisa.aau.at

The screenshot displays the HarmonISA web application interface. The browser window title is "HarmonISA - Harmonisierter Landnutzungsvierer für Kärnten, Friaul-Julisch Venetien, Slowenien - Mozilla Firefox". The address bar shows the URL "http://harmonisa.uni-klu.ac.at/harmonisa2/index.jsp".

The main content area features a map of Carinthia, Austria, overlaid with a land use classification. The map is color-coded according to the legend on the right. A small inset map in the top left of the main map area shows the location of Carinthia within Austria.

On the left side, there is a "SUBCLASS RELATIONSHIP" panel for the project "Europe - Corine". It shows a hierarchical tree of classes under "Europe - Corine":

- Corine area
 - Agricultural areas
 - Artificial surfaces
 - Artificial non-agricult
 - Green urban area
 - Sport and leisure
 - Industrial, commercia
 - Mine, dump and cons
 - Urban fabric
 - Forests and semi-natural
 - Water bodies
 - Wetlands

On the right side, there is a "Legende" (Legend) panel with buttons for "Kataloge", "Filter", and "Abfrage". The legend text states: "Die Legende enthält nur jene Kategorien die Sie ausgewählt haben." Below this, a list of categories is shown with corresponding color swatches:

- Bebaute Fläche
- Städtisch geprägte Flächen
- Durchgängige städtische Prägung
- Nicht durchgängig städtische Prägung
- Industrie-, Gewerbe- und Verkehrsflächen
- Industrie/Gewerbeflächen
- Straßen/Eisenbahnnetze, funktionell zugeordnete Flächen
- Hafengebiete
- Flughäfen
- Abbauflächen, Deponien, Baustellen
- Abbauflächen
- Deponien, Abraumhalden
- Baustellen
- Künstlich angelegte nicht landwirtschaftlich genutzte Flächen
- Städtische Grünflächen

At the bottom of the interface, the copyright information reads: "© 2005, 2006 Institut für Geographie und Regionalforschung, Universität Klagenfurt" and the version information is "Version 1.7 (Letzte Aktualisierung: 07.03.2006)".

data.ktn.gv.at, data.gv.at

The screenshot shows the 'Open Government Data' page on data.ktn.gv.at. The browser address bar displays 'http://data.ktn.gv.at/data/'. The page features a navigation menu with buttons for 'News', 'Daten', 'Anwendungen', 'Nutzung', and 'Über OGD Kärnten'. Below the menu, a breadcrumb trail reads 'Sie befinden sich hier: Open Government Data > Daten'. A sidebar on the left lists various categories such as 'Arbeit (4)', 'Bevölkerung (4)', 'Bildung und Forschung (5)', 'Finanzen und Rechnungswesen (5)', 'Geographie und Planung (34)', 'Kunst und Kultur (2)', 'Land- und Forstwirtschaft (2)', 'Sport und Freizeit (5)', 'Umwelt (24)', 'Verkehr und Technik (5)', 'Verwaltung und Politik (3)', 'Wirtschaft und Tourismus (3)', and 'Alle Datensätze (90)'. The main content area is titled 'Daten' and includes a search bar and a list of categories.

The screenshot shows the data.gv.at website interface. The browser address bar displays 'https://www.data.gv.at/'. The page features a search bar with the text 'Suchbegriff (z.B. Finanzen, Wahlen)' and a 'Suche starten' button. Below the search bar, there are tabs for 'Daten & Dokumente', 'Apps & News', and 'Katalog durchstöbern'. The main content area is titled 'offene Daten Österreichs – lesbar für Mensch und Maschine'. It includes a sub-heading 'Vielfalt, Transparenz, Offenheit, Demokratie' and a paragraph stating 'data.gv.at bietet einen Katalog offener Datensätze und Dienste aus der öffentlichen Verwaltung, welche auf den Open Data-Prinzipien basieren.' Below this, there is a paragraph about using the data for personal information and commercial purposes, and a link to 'Netiquette'. A diagram shows a computer monitor displaying binary code (0s and 1s) with arrows pointing to a group of people and a smartphone. The text 'Straße XY' is visible above the group of people. At the bottom right, there is a blue button with a star icon and the text 'Daten hinzufügen'.

<http://www.opengeospatial.org/standards>

The screenshot displays the Open Geospatial Consortium (OGC) website. The header features the OGC logo with the tagline "Making location count." and a navigation menu with items: Home, Standards, Programs, Participate, News & Events, About OGC, and Member Login. A search bar is located to the right of the menu.

The main content area is titled "Geospatial and location standards for:" and lists various application domains on the left:

- Aviation
- Built Environment & 3D
- Defense & Intelligence
- Emergency Response & Disaster Management
- Geosciences & Environment
- Government & SDI
- Energy & Utilities
- Law Enforcement / Public Safety
- Mobile Internet & LBS
- Sensor Webs
- University & Research

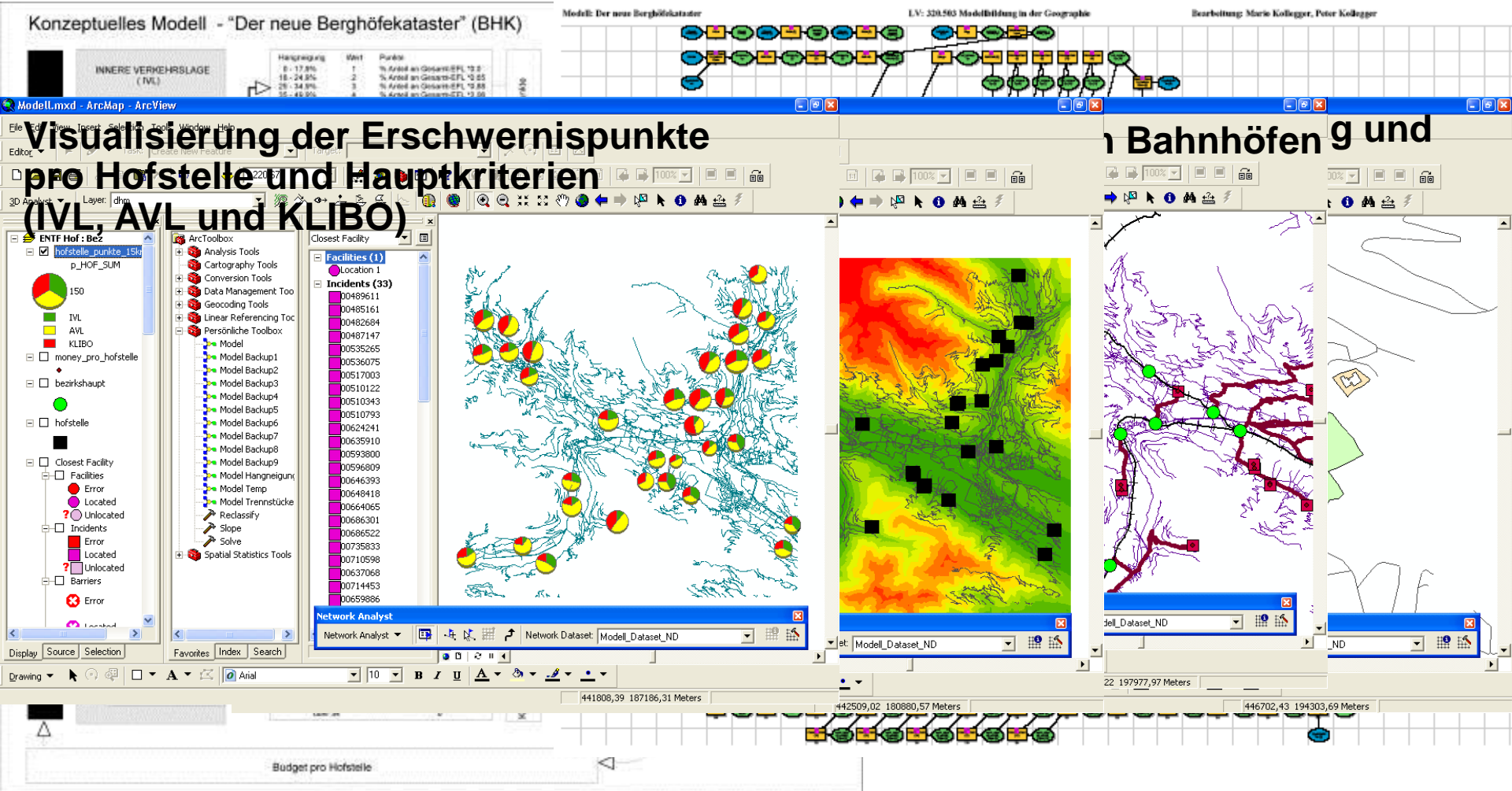
The central part of the page features a complex diagram with "Open" at the top center. It branches into several key areas:

- Open Source**: Includes BIM, Proximity, and Global.
- Open Data**: Includes Time, Planning, and SDI.
- Interoperability**: Includes Information Integration, Geosynchronization, and Data Quality.
- Points of Interest**: Includes Sensor Web, Shared Understanding, and Geoweb.
- Situational Awareness**: Includes Alerts, Visualization, and Real Time.
- Where**: Includes Hydrology and GPS.
- Share**: Includes Map and Information Integration.
- Spatial Policy**: Includes Monitoring and Location.
- Open**: Includes Navigation, Open Source, and GIS.
- Analysis**: Includes Earth Observation and Crowdsourcing.
- Geoweb**: Includes Geosemantics and Metadata.
- Geoweb**: Includes Indoor/Outdoor and Metadata.

At the bottom of the page, there is a footer with the copyright notice: "©1994 - 2015 Open Geospatial Consortium. OpenGIS® and OGC® are registered trademarks of OGC." and a navigation bar with links for Home, Contact Us, Search this Site, Site Map, Join OGC LinkedIn, Follow OGC on Twitter, RSS Feeds, OGC Member Portal, OGC Network, OGC Compliance Testing, and Upcoming Events. The date "Montag, 26. Jänner 2015" is visible in the bottom right corner.

Model „Der neue Berghöfekataster“

Mario Kollegger, Peter Kollegger



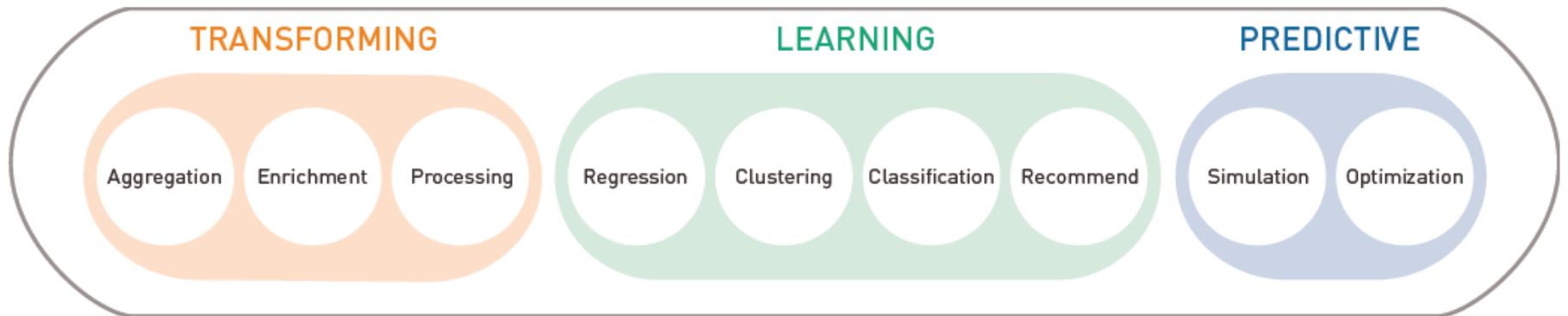
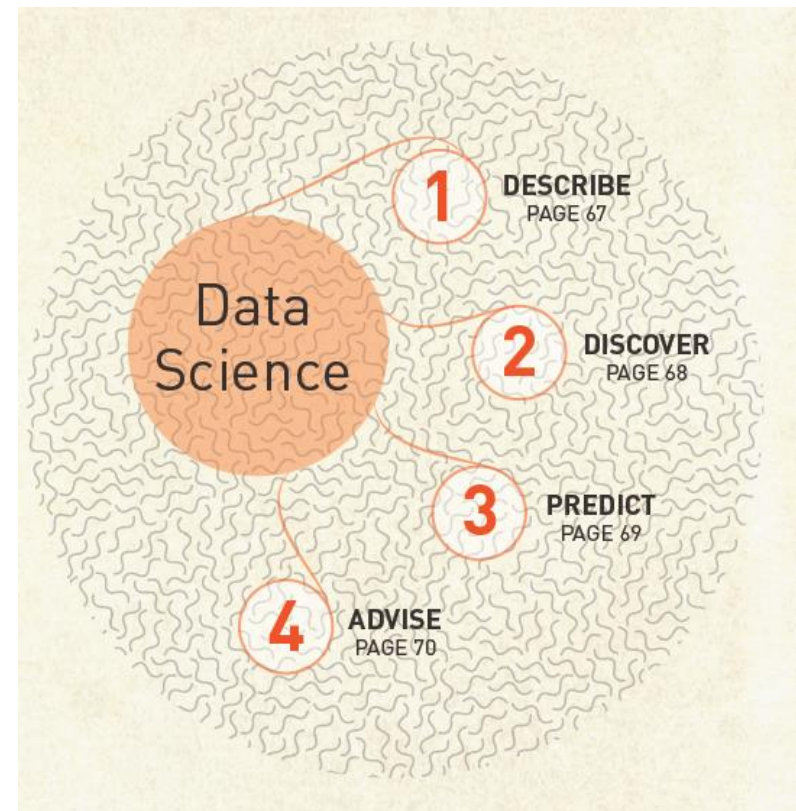
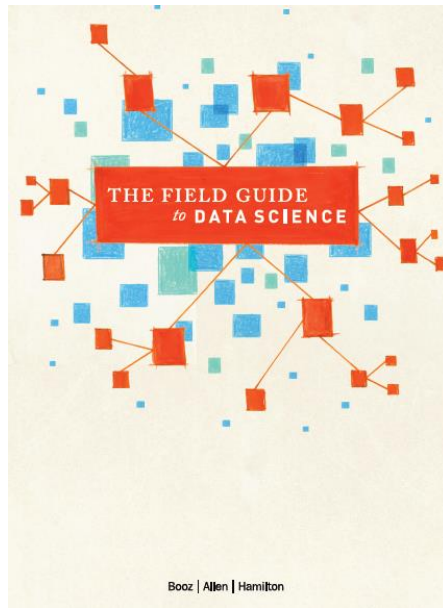
Data, Concepts and Models of

- Geography: (big) geo-data, concepts, models, location analytics, spatial data mining, visual analytics, ...
- GI-Science: semantic web, linked open data, GIS-modeling, ...
- **Data Analytics:** methods from quantitative geography extended by visual, predictive, prescriptive and explorative aspects, business analytics, **data science**

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

Algorithms or Method Name	Description	
Differential Equations	Used to express relationships between functions and their derivatives, for example, change over time.	Different models; themselves tested w system t
Discrete Event Simulation	Simulates a discrete sequence of events where each event occurs at a particular instant in time. The model updates its state only at points in time when events occur.	Discrete analyz in product determin as differ Optimiz gain effici
Discrete Wavelet Transform	Transforms time series data into frequency domain preserving locality information.	Offers vs localizat transform and local
Exponential Smoothing	Used to remove artifacts expected from collection error or outliers.	In comp; where p; exponen decrease
Factor Analysis	Describes variability among correlated variables with the goal of lowering the number of unobserved variables, namely, the factors.	If you su influenc try facto
Fast Fourier Transform	Transforms time series from time to frequency domain efficiently. Can also be used for image improvement by spatial transforms.	Filtering more eff noise ca observin
Format Conversion	Creates a standard representation of data regardless of source format. For example, extracting raw UTF-8 encoded text from binary file formats such as Microsoft Word or PDFs.	There ar package can inter notable
Gaussian Filtering	Acts to remove noise or blur data.	Can be u noise fr
Generalized Linear Models	Expands ordinary linear regression to allow for error distribution that is not normal.	Use if th not follo
Genetic Algorithms	Evolves candidate models over generations by evolutionary inspired operators of mutation and crossover of parameters.	Increase in consic requires Calculat Is strong candidat
Grid Search	Systematic search across discrete parameter values for parameter exploration problems.	A grid at visualize whether

Algorithms or Method Name	Description
Hidden Markov Models	Models sequential data by determining the discrete latent variables, but the observables may be continuous or discrete.
Hierarchical Clustering	Connectivity based clustering approach that sequentially builds bigger (agglomerative) or smaller (divisive) clusters in the data.
K-means and X-means Clustering	Centroid based clustering algorithms, where with K means the number of clusters is set and X means the number of clusters is unknown.
Linear, Non-linear, and Integer Programming	Set of techniques for minimizing or maximizing a function over a constrained set of input parameters.
Markov Chain Monte Carlo (MCMC)	A method of sampling typically used in Bayesian models to estimate the joint distribution of parameters given the data.
Monte Carlo Methods	Set of computational techniques to generate random numbers.
Naïve Bayes	Predicts classes following Bayes Theorem that states the probability of an outcome given a set of features is based on the probability of features given an outcome.
Neural Networks	Learns salient features in data by adjusting weights between nodes through a learning rule.
Outlier Removal	Method for identifying and removing noise or artifacts from data.
Principal Components Analysis	Enables dimensionality reduction by identifying highly correlated dimensions.

Algorithms or Method Name	Description	Tips From the Pros	References and Papers We Love to Read
Regression with Shrinkage (Lasso)	A method of variable selection and prediction combined into a possibly biased linear model.	There are different methods to select the lambda parameter. A typical choice is cross validation with MSE as the metric.	Tibshirani, Robert. "Regression Shrinkage and Selection via the Lasso." <i>Journal of the Royal Statistical Society. Series B (Methodological)</i> 58.1 [1996]: 267-288. Print.
Sensitivity Analysis	Involves testing individual parameters in an analytic or model and observing the magnitude of the effect.	Insensitive model parameters during an optimization are candidates for being set to constants. This reduces the dimensionality of optimization problems and provides an opportunity for speed up.	Saitelli, A., Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. <i>Global Sensitivity Analysis: the Primer</i> . New Jersey: John Wiley & Sons, 2008. Print.
Simulated Annealing	Named after a controlled cooling process in metallurgy, and by analogy using a changing temperature or annealing schedule to vary algorithmic convergence.	The standard annealing function allows for initial wide exploration of the parameter space followed by a narrower search. Depending on the search priority the annealing function can be modified to allow for longer explorative search at a high temperature.	Bertsimas, Dimitris, and John Tsitsiklis. "Simulated Annealing." <i>Statistical Science</i> . 8.1 [1993]: 10-15. Print.
Stepwise Regression	A method of variable selection and prediction. Akaike's information criterion AIC is used as the metric for selection. The resulting predictive model is based upon ordinary least squares, or a general linear model with parameter estimation via maximum likelihood.	Caution must be used when considering Stepwise Regression, as over fitting often occurs. To mitigate over fitting try to limit the number of free variables used.	Hocking, R. R. "The Analysis and Selection of Variables in Linear Regression." <i>Biometrics</i> . 32.1 [March 1976]: 1-49. Print.
Stochastic Gradient Descent	General-purpose optimization for learning of neural networks, support vector machines, and logistic regression models.	Applied in cases where the objective function is not completely differentiable when using sub-gradients.	Witten, Ian H., Elise Frank, and Mark A. Hall. <i>Data Mining: Practical Machine Learning Tools and Techniques</i> . Massachusetts: Morgan Kaufmann, 2011. Print.
Support Vector Machines	Projection of feature vectors using a kernel function into a space where classes are more separable.	Try multiple kernels and use k-fold cross validation to validate the choice of the best one.	Hsu, Chih-Wei, Chih-Chung Chang, and Chih-Jen Lin. "A Practical Guide to Support Vector Classification." <i>National Taiwan University</i> . 2003. Print.
Term Frequency Inverse Document Frequency	A statistic that measures the relative importance of a term from a corpus.	Typically used in text mining. Assuming a corpus of news articles, a term that is very frequent such as "the" will likely appear many times in many documents, having a low value. A term that is infrequent such as a person's last name that appears in a single article will have a higher TF-IDF score.	Ingersoll, Grant S., Thomas S. Morton, and Andrew L. Farris. <i>Taming Text: How to Find, Organize, and Manipulate It</i> . New Jersey: Manning, 2013. Print.
Topic Modeling (Latent Dirichlet Allocation)	Identifies latent topics in text by examining word co-occurrence.	Employ part-of-speech tagging to eliminate words other than nouns and verbs. Use raw term counts instead of TF-IDF weighted terms.	Blei, David M., Andrew Y. Ng, and Michael I. Jordan. "Latent Dirichlet Allocation." <i>Journal of Machine Learning Research</i> . 3 [March 2003]: 993-1022. Print.
Tree Based Methods	Models structured as graph trees where branches indicate decisions.	Can be used to systematize a process or act as a classifier.	James, G., D. Witten, T. Hastie, and R. Tibshirani. <i>Tree-Based Methods. In An Introduction to Statistical Learning</i> . New York: Springer, 2013. Print.
Wrapper Methods	Feature set reduction method that utilizes performance of a set of features on a model, as a measure of the feature set's performance. Can help identify combinations of features in models that achieve high performance.	Utilize k-fold cross validation to control over fitting.	John, G. H., R. Kohavi, and K. Pfleger. "Irrelevant Features and the Subset Selection Problem." Proceedings of ICML-94, 11th International Conference on Machine Learning. New Brunswick, New Jersey. 1994. 121-129. 59. Conference Presentation.



Algorithms or Method Name	Description
Agent Based Simulation	Simulates the interaction of autonomous agents.
Collaborative Filtering	Also known as 'Recommendation' or 'Filtering'. It uses the behavior of similar items to suggest items to users based on their past behavior.
Coordinate Transformation	Provides a different perspective on data by changing its coordinate system.
Design of Experiments	Applies control experiments to effects on systems caused by changes in parameters.

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

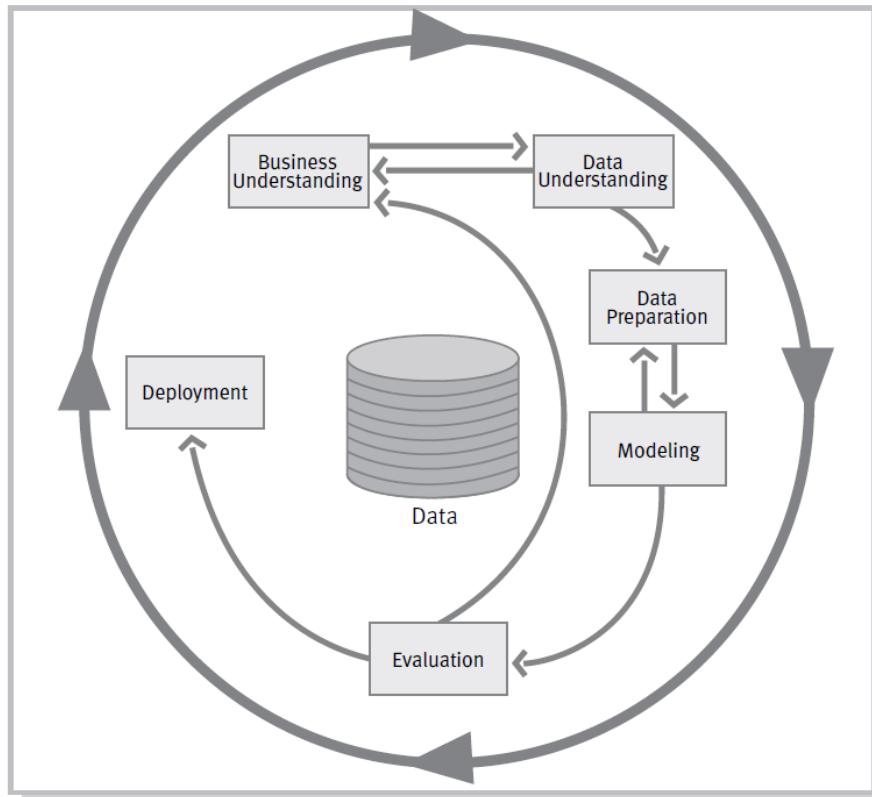


Figure 2: Phases of the CRISP-DM reference model

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.

Development of a Spatial Data Science

In a Spatial Data Science

spatio-temporal problems using established and new tools from the mentioned linked research fields are solved

What is to do?

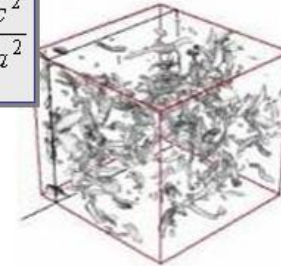
- **Use New** Science Paradigms
 - 4th scientific paradigm of an e-science (data exploration approach)
 - http://research.microsoft.com/en-us/um/people/gray/talks/NRC-CSTB_eScience.ppt
- **Design Suitable** Concepts and Frameworks
 - <http://www.esri.com/news/releases/12-3qtr/carl-steinitz-explains-geodesign-process-in-new-esri-press-book.html>
- **Use** Analytics Tools and Methods
 - rapidminer.com/
- **Integrate** Visualization
 - CommonGIS
 - <http://www.iais.fraunhofer.de/1871.html>
- **Apply** More Spatio-Temporal Modeling
 - IDRISI or TerrSet
 - www.clarklabs.org/

Use New Science Paradigms

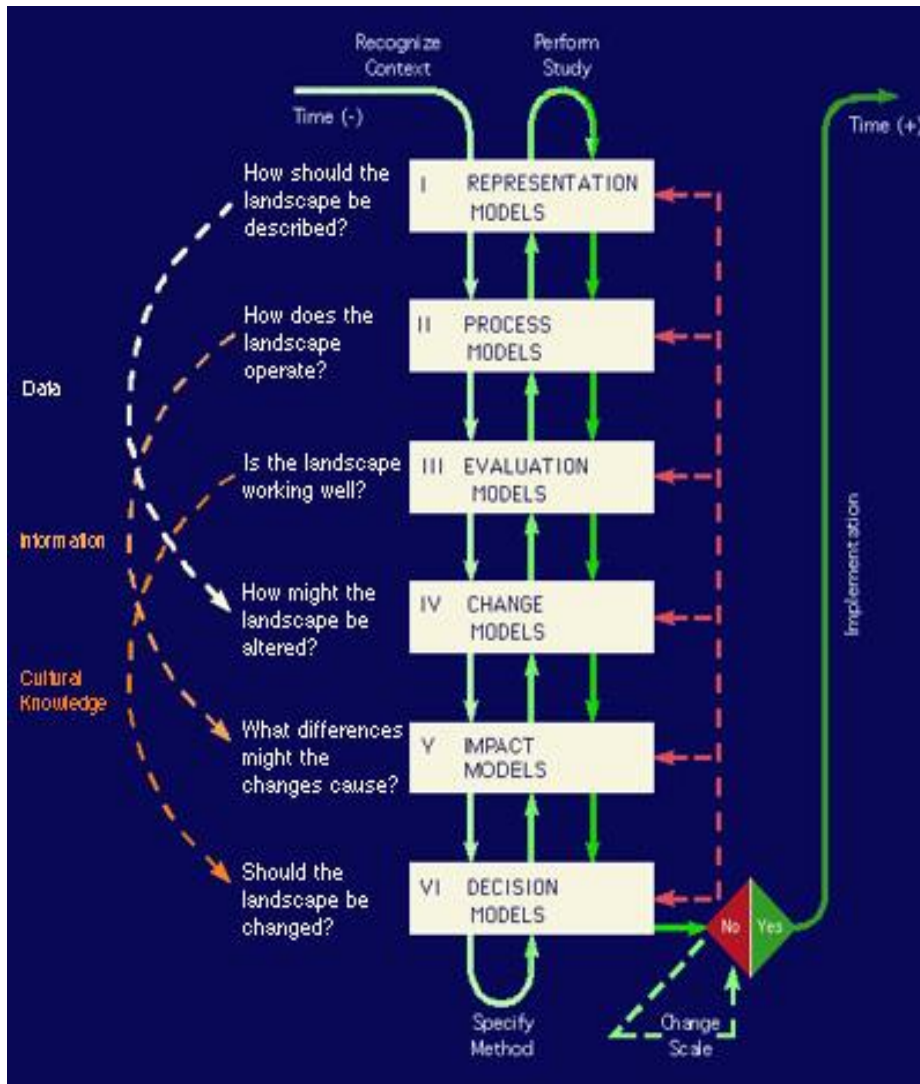
- Thousand years ago:
science was **empirical**
describing natural phenomena
- Last few hundred years:
theoretical branch
using models, generalizations
- Last few decades:
a **computational** branch
simulating complex phenomena
- Today:
data exploration (eScience)
unify theory, experiment, and simulation
 - Data captured by instruments
Or generated by simulator
 - Processed by software
 - Information/Knowledge stored in computer
 - Scientist analyzes database / files
using data management and statistics



$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K \frac{c^2}{a^2}$$



Design Suitable Concepts and Frameworks



<http://www.esri.com/news/arcwatch/0412/a-conversation-with-carl-steinitz.html>

<http://www.esri.com/news/releases/12-3qtr/carl-steinitz-explains-geodesign-process-in-new-esri-press-book.html>

Use Analytics Tools and Methods

rapidminer.com/

The screenshot shows the main RapidMiner interface. The central workspace displays a workflow diagram with the following operators: Training, Set Role, Linear Regression, Aggregate, Scoring, Filter Examples, and Apply Model. The left sidebar shows the 'Operators' list and 'Repositories' section. The right sidebar shows the 'Parameters' and 'Context' panels.

The screenshot shows the 'Result Overview' window for a 'LinearRegression (Linear Regression)' process. The data table displays 18 rows of results with columns: Row No., prediction(...), insulation, Temperature, Num_Occu..., Avg_Age, and Home_Size.

Row No.	prediction(...)	insulation	Temperature	Num_Occu...	Avg_Age	Home_Size
1	251.321	5	69	10	70.100	7
2	216.028	5	80	1	66.700	1
3	226.087	4	89	9	67.800	7
4	209.529	7	81	9	52.400	6
5	164.669	4	58	8	22.900	7
6	180.512	4	58	6	37.400	3
7	221.188	6	51	2	51.600	3
8	164.001	2	73	5	37.400	4
9	264.712	9	39	1	56.900	7
10	221.364	8	84	5	64.500	2
11	221.328	10	74	6	58.300	1
12	262.580	5	49	6	68.600	6
13	214.082	8	45	2	33.900	8
14	212.392	3	49	4	49.700	4
15	253.199	9	66	6	66.200	5
16	275.043	9	57	10	70.100	7
17	190.837	9	66	10	32.900	6
18	234.624	4	47	3	55.200	6

The bottom window shows the 'Log' window with the following entries:

- LinearRegression[1] (Linear Regression)
- Scoring[1] (Read CSV)
- Filter Examples[1] (Filter Examples)
- Filter Examples[2][1] (Filter Examples)
- Apply Model[1] (Apply Model)

Other Geo Tools

„Geo“-Software, which is based on Big Data approaches - like Spatial Data Mining or Geospatial Visual Analytics:

<http://www.csiss.org/clearinghouse/select-tools.php>

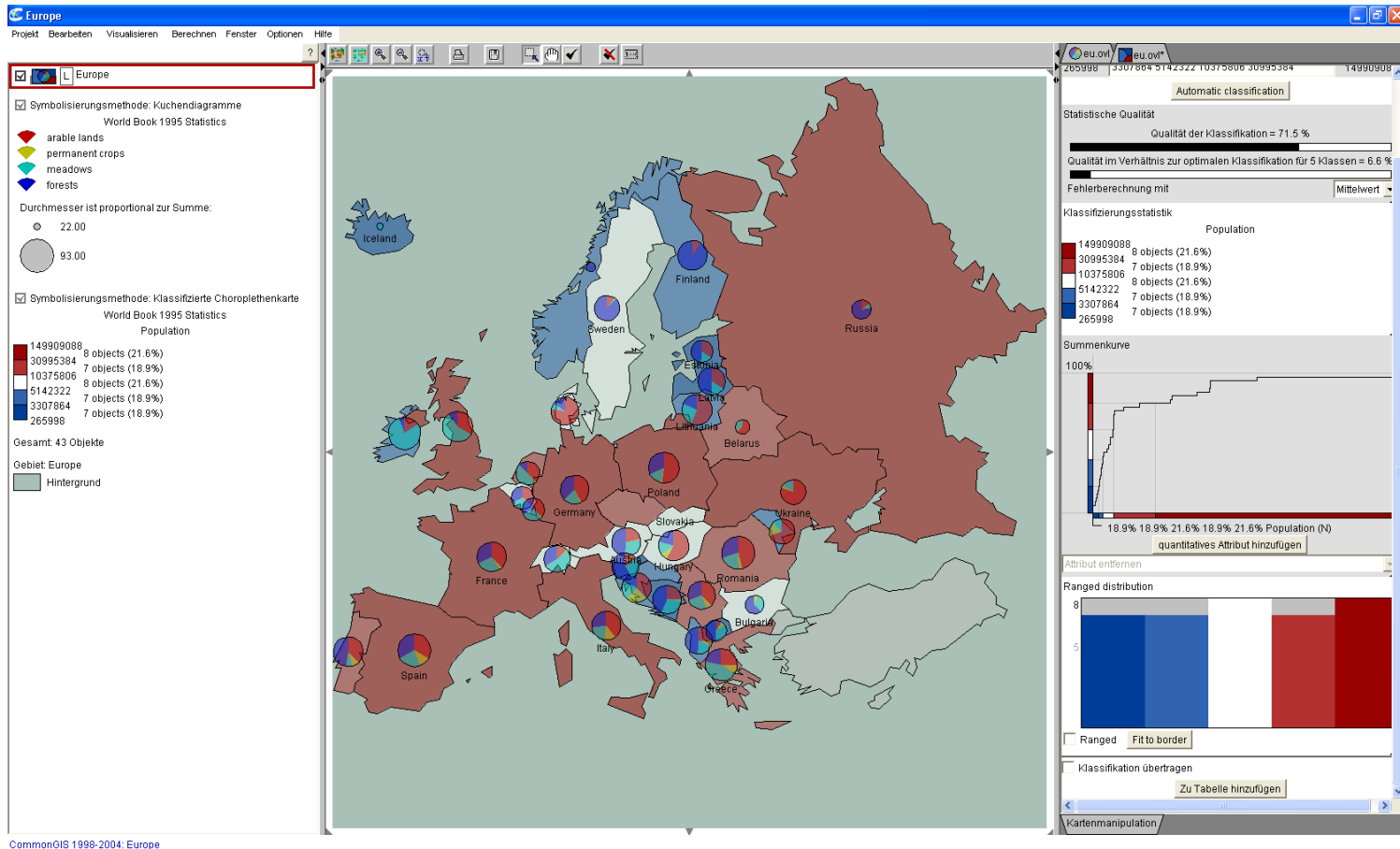
Special example: Common GIS

<http://www.iais.fraunhofer.de/1865.html>

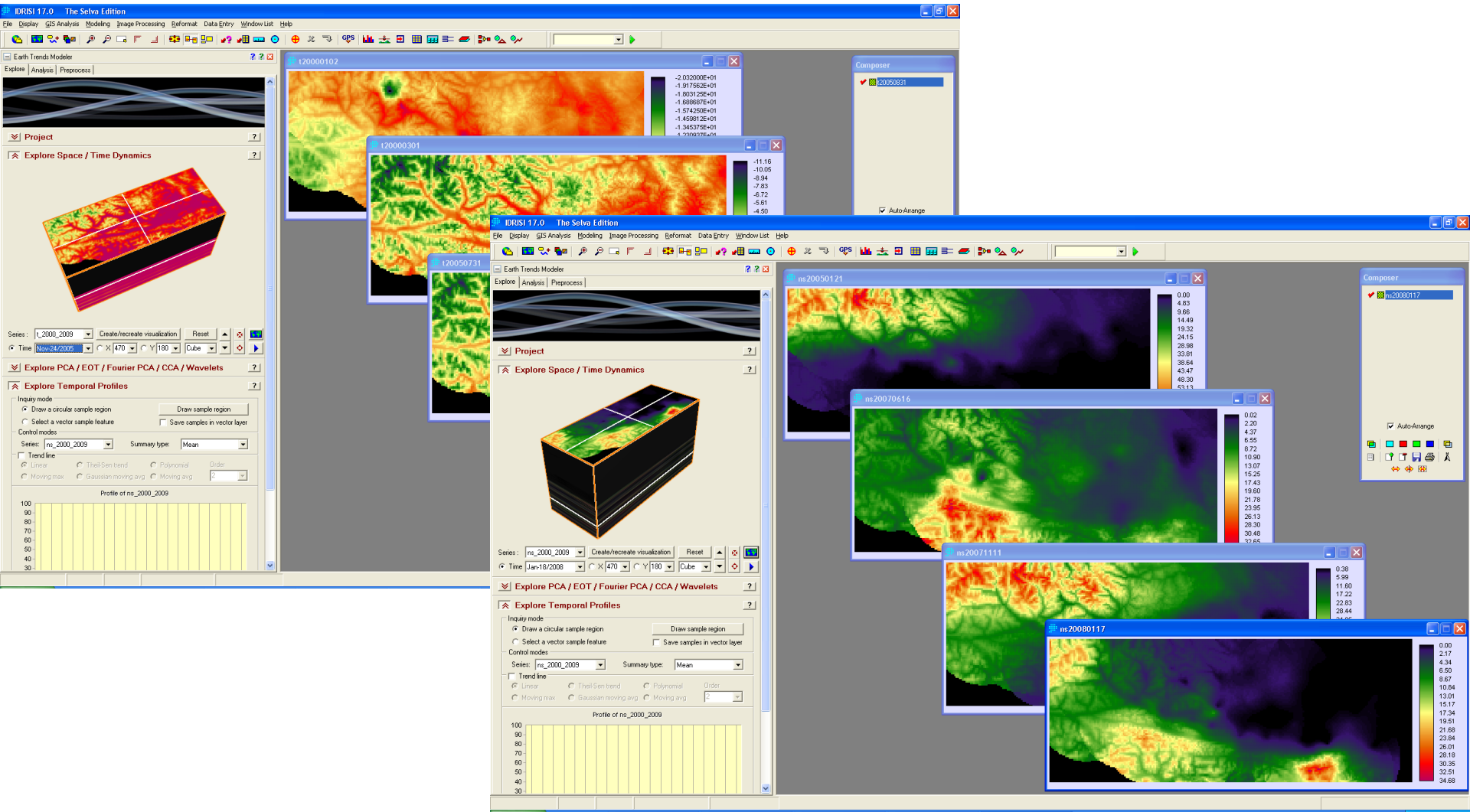
(Educational Version is free of charge)

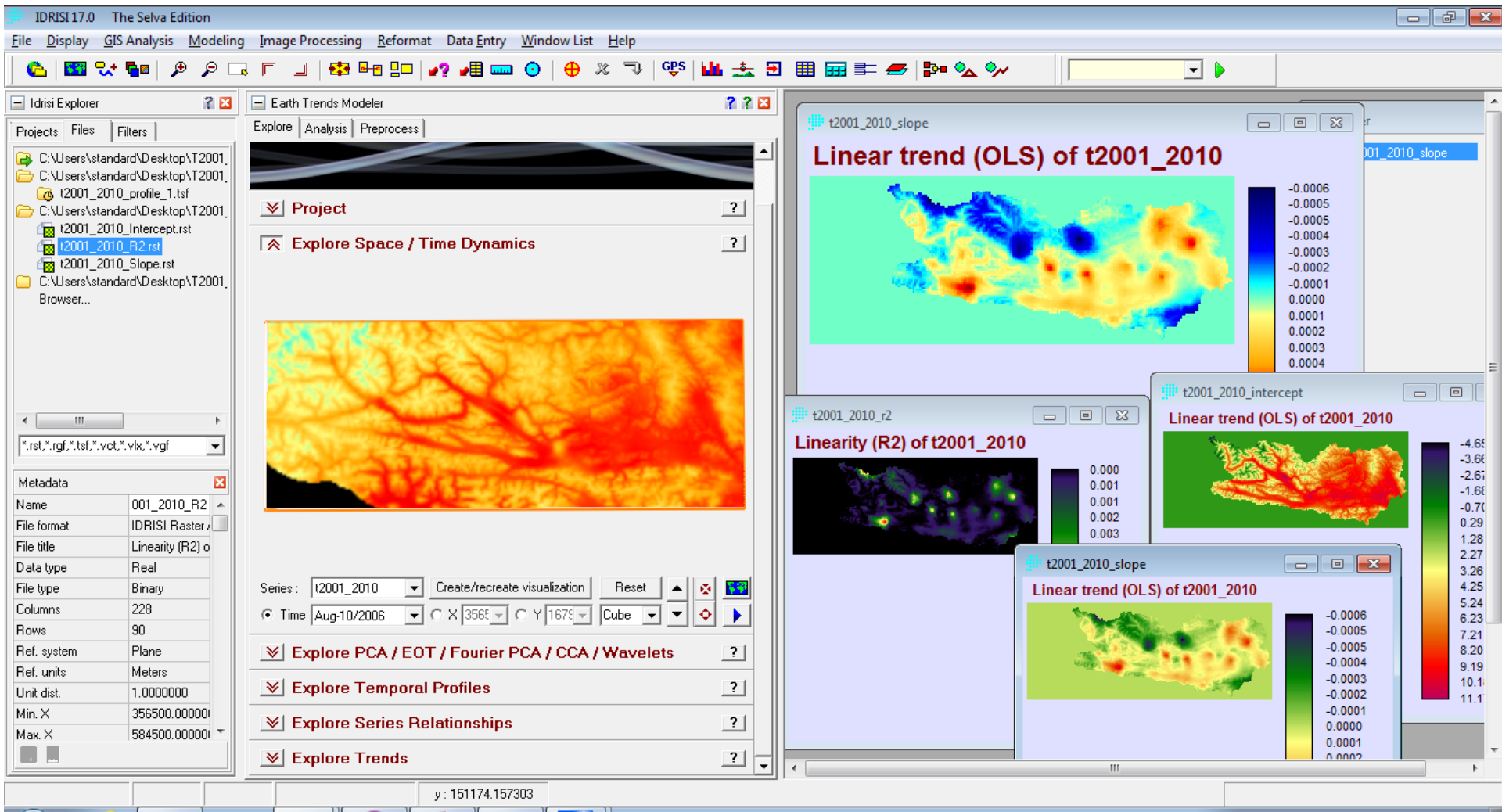
Integrate Visualization CommonGIS

<http://www.iais.fraunhofer.de/1871.html>



Apply More Spatio-Temporal Modeling IDRISI or TerrSet www.clarklabs.org/





FlyOnTime.us

Find a Route

From: (city or airport) To: (city or airport; optional)

Example route: [LAX to SFO](#) is 61% on-time and 3 min. early on average

Find An Airline/Flight

Airline: Flight #: (optional)

Security Lines

Search wait time statistics for [security lines](#).

You can also contribute by notifying us when you get on line and then past security via Twitter or [from your mobile phone](#).

Site News

March 12, 2011. Mentioned in [The New York Times](#).
March 14, 2010. Mentioned on [National Public Radio](#).
July 21, 2009. Mentioned in [The Washington Post](#).
June 24, 2009. Mentioned in [The Politico](#).

See how FlyOnTime.us can [save your tax dollars!](#)

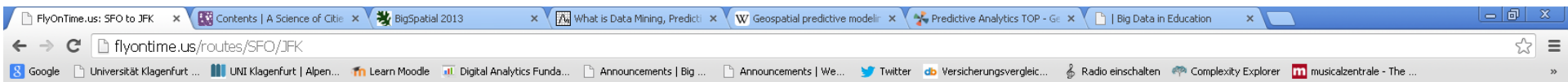
Data on this site is derived from:

- The [Bureau of Transportation Statistics](#) via [data.gov](#)
- The [Federal Aviation Administration](#)
- The [National Oceanic and Atmospheric Administration](#)

[People Like You](#)

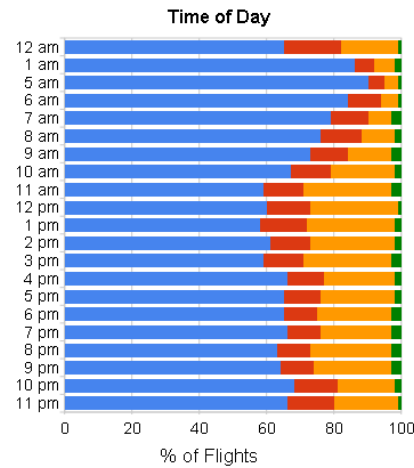
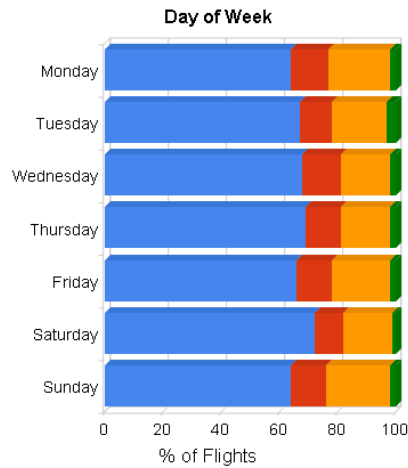
[Terms of Use](#)

<http://flyontime.us/routes/SFO/JFK>



Best Days and Times to Fly from SFO

Based on all flights originating at San Francisco, CA: San Francisco International.



■ On Time ■ >20 min. Delay
■ 5-20 min. Delay ■ Cancelled/Diverted

■ On Time ■ >20 min. Delay
■ 5-20 min. Delay ■ Cancelled/Diverted

Holiday Delays at SFO

Based on all flights originating at San Francisco, CA: San Francisco International.

Holiday	Average Arrival	Be Prepared For	Cancelled
<i>Most Days</i>	<i>11 min. early</i>	<i>20 min. late</i>	3%
Day Before Memorial Day (326 flights)	13 min. early	on time	1%
Memorial Day (387 flights)	6 min. early	26 min. late	1%
Day After Memorial Day (391 flights)	6 min. late	71 min. late	2%
Labor Day (403 flights)	12 min. early	1 min. late	0%
Wednesday Before Thanksgiving Day (396 flights)	6 min. early	12 min. late	0%
Thanksgiving Day (285 flights)	11 min. early	7 min. late	0%
Friday After Thanksgiving Day (317 flights)	12 min. early	3 min. late	0%
Saturday After Thanksgiving Day (371 flights)	1 min. late	43 min. late	1%

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

The screenshot shows the Google Flu Trends interface for Austria. The browser address bar displays www.google.org/flutrends/at/#AT. The page title is "Grippe-Trends" and the language is set to "Deutsch".

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 ▾

The line graph shows flu activity levels from July to June. The y-axis ranges from "minimal" to "sehr hoch". A color-coded bar at the bottom indicates the months from Jul to Jun. The 2013-2014 season (dark blue line) shows a significant peak in late January and early February, reaching a level between "hoch" and "sehr hoch".

Map of Austria: The map shows regional flu activity. The northern and eastern regions are colored in shades of orange and yellow, indicating higher activity, while the southern and western regions are in shades of green, indicating lower activity.

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

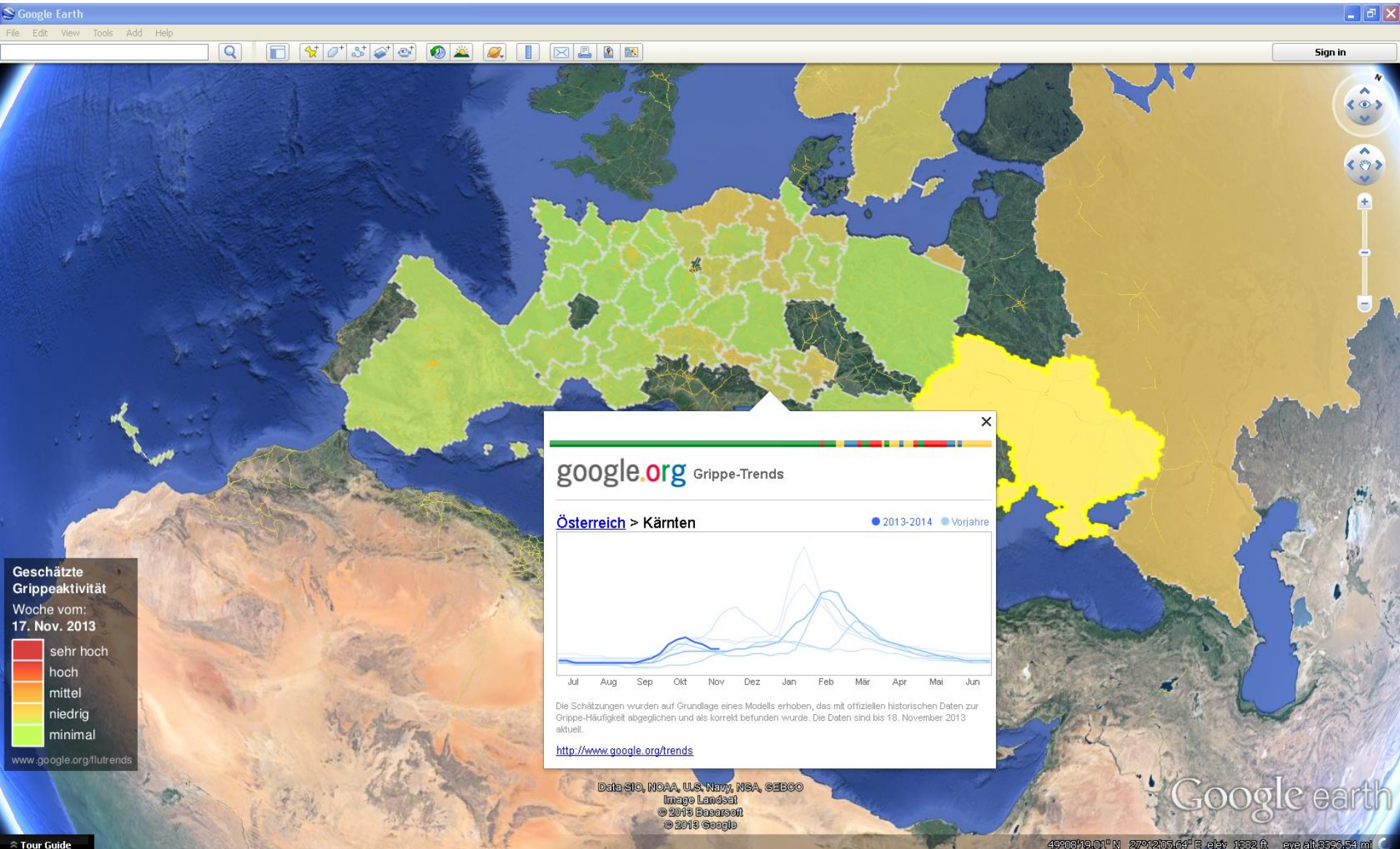
Laden Sie [hier](#) Daten von Google Grippe-Trends in Google Earth herunter und untersuchen Sie sie. Sie haben noch kein Google Earth? [Laden Sie es hier herunter.](#)

Tabelle einbetten

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

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.

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



What is to do?

- **Use New** Science Paradigms
 - 4th scientific paradigm of an e-science (data exploration approach)
 - http://research.microsoft.com/en-us/um/people/gray/talks/NRC-CSTB_eScience.ppt
- **Design Suitable** Concepts and Frameworks
 - <http://www.esri.com/news/releases/12-3qtr/carl-steinitz-explains-geodesign-process-in-new-esri-press-book.html>
- **Use** Analytics Tools and Methods
 - rapidminer.com/
- **Integrate** Visualization
 - CommonGIS
 - <http://www.iais.fraunhofer.de/1871.html>
- **Apply** More Spatio-Temporal Modeling
 - IDRISI or TerrSet
 - www.clarklabs.org/

Additional Material

Slides:

- Shashi Shekhar (2014): Spatial Data Science: Challenges & Opportunities

http://www.nist.gov/itl/iad/upload/Shashi-Shekhar-14-3-CCC_NIST_WEB.pdf

> Maggi Kelly (2014): Spatial Data Science: the 21st Century Mapping Toolkit

[https://www.youtube.com/watch?v=qWQp11CE4S](https://www.youtube.com/watch?v=qWQp11CE4S0)

[0](https://www.youtube.com/watch?v=qWQp11CE4S0)

To a Data-Driven Geography or Spatial Data Science?

- In a **Spatial Data Science** we have **to consider**
 - Errors / Calibration / Privacy protection / Prediction without explanation?
- **Granularity** of the data (**resolution**) and **Accuracy**
- **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**.

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



**Thank you very much for
your attention!**

Questions and Suggestions please to:

peter.mandl@aau.at

Informations at:

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