

Wissenschaftliche Umweltdaten in Geodateninfrastrukturen

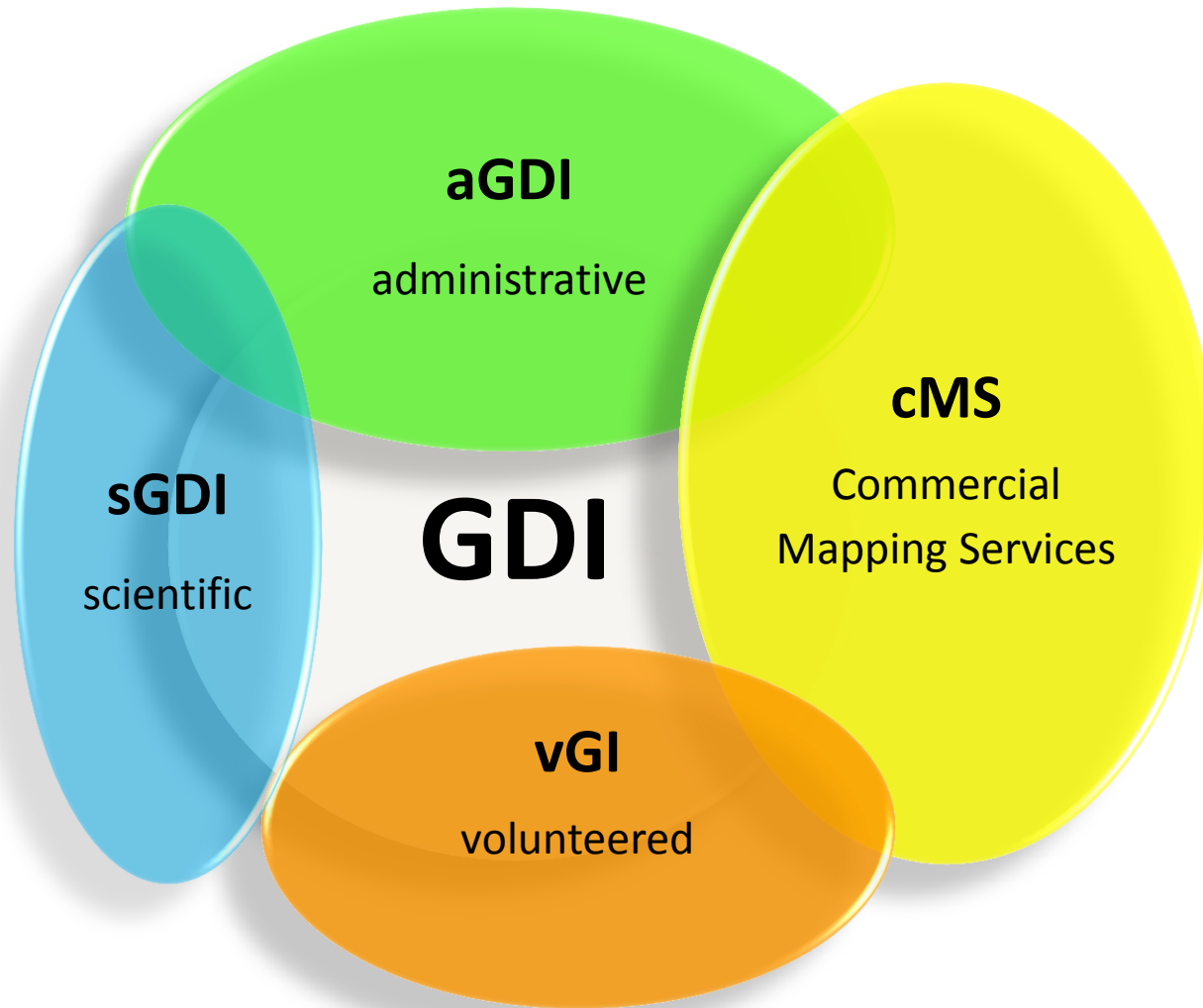
Dr.-Ing. Stephan Mäs

Professur für Geoinformationssysteme
Technische Universität Dresden

Inhalt

- wissenschaftliche Dateninfrastrukturen
- GLUES Projekthintergrund
 - Rahmenbedingungen & Ziele
- Besondere Aspekte einer GDI für wissenschaftliche Daten
 - Eindeutige Identifikation von Datensätzen
 - Metadaten
 - Beschreibung der Entstehungsgeschichte
 - Lizenzierung, Haftung
 - Privatsphäre
 - Organisation der GDI
 - Beispiele von Diensten & Tools zur Verarbeitung und Visualisierung von Daten bzw. Metadaten

Arten von Geodateninfrastrukturen



Wissenschaftliche Dateninfrastrukturen

- Forschungsinitiativen
 - US National Science Foundation initiative on scientific **cyberinfrastructures**
 - Esfri: European roadmap for research infrastructures 2007, **e-Infrastructures**
- Viele existierende Beispiele:
 - Social Network for Bioinformatic Workflows: <http://www.myexperiment.org/>
 - iPlant Collaborative cyberinfrastructure <http://www.iplantcollaborative.org/>
 - GCMD - Global Change Master Directory: <http://gcmd.nasa.gov/index.html>
- Zentrale Funktionalitäten:
 - Recherche in Katalogen,
 - Zugriff auf Daten und Visualisierungen
 - Rechenkapazitäten
 - Interaktion und Kollaboration zwischen Wissenschaftlern

Warum sollen wissenschaftliche Daten ausgetauscht werden?

- verbesserte Dokumentation, Transparenz, Vergleichbarkeit und Nachhaltigkeit der Forschungsarbeiten
- die Stimulans zur Wiederverwendung wissenschaftlicher Daten und damit der Kollaboration zwischen Wissenschaftlern
- die Unterstützung von datenintensiver multidisziplinärer Forschung
- die erhöhte Rentabilität öffentlich geförderter Forschungsinitiativen
- die Bereitstellung der Daten und Forschungsergebnisse für die Öffentlichkeit und für politische Entscheidungsträger

Scientific Geodata Infrastructures

The screenshot displays the IPCC-Data.org interface. At the top, it features the logos of the Intergovernmental Panel on Climate Change (IPCC), WMO, and UNEP. Below the header, there are navigation links for IPCC, WG1, WG2, WG3, and TGICA. The main content area is divided into three sections:

- Map:** A global map showing temperature anomalies. A color scale at the bottom indicates temperatures from 220 K (blue) to 320 K (red). The map shows a clear warming trend, particularly in the mid and high latitudes.
- Selection Description:**
 - IPCC-DDC 4th Assessment Report
 - Scenario 20C3M
 - 30 year mean climatologies
 - BCCR-BCM2.0
 - time: Jan 1901-1930
- Domain:**
 - Latitude: 90 N to -90 S
 - Longitude: -180 W to 180 E
 - Buttons: Clear selection
 - Time: Jan 1901-1930
- Plot type:**
 - Radio buttons for Colour (selected) and Format
 - Buttons: PNG, Make Figure

At the bottom, there is a table with three columns: Dataset, Field, and Selection Notes.

Dataset	Field	Selection Notes
<ul style="list-style-type: none"> IPCC-DDC CRU TS-2.1 observation climatologies IPCC-DDC 4th Assessment Report <ul style="list-style-type: none"> Scenario 1PT02X Scenario 1PT04X Scenario 20C3M <ul style="list-style-type: none"> 30 year mean climatologies <ul style="list-style-type: none"> BCCR-BCM2.0 CCSM3 CGCM3.1(T47) CNRM-CM3 CSIRO-MK3.0 ECHAM5/MPI-OM 	<ul style="list-style-type: none"> Daily Maximum Temperature Daily Minimum Temperature Precipitation Pressure Shortwave Flux Specific Humidity Temperature Wind Eastward Wind Northward 	<ul style="list-style-type: none"> IPCC-DDC 4th Assessment Report Scenario 20C3M <ul style="list-style-type: none"> Experiments run with greenhouse gasses increasing as observed through the 20th century. 30 year mean climatologies BCCR-BCM2.0 <ul style="list-style-type: none"> more info...

Intergovernmental Panel on Climate Change

<http://www.ipcc-data.org/maps/>

Integration von GIS-Funktionalitäten in Lösungen

- Fördermaßnahme des BMBF “Nachhaltiges Landmanagement” Modul A
 - Systemverständnis und Wechselwirkungen zwischen Landmanagement und Klimaschutz
 - 12 Regionalprojekte entwickeln Beispiellösungen für global verteilte Untersuchungsgebiete mit unterschiedlichen Schwerpunkten



Koordinationsprojekt GLUES



GLUES GDI for Research on Sustainable Land Management

(Global Assessment of Land Use Dynamics, Greenhouse Gas Emissions and Ecosystem Services)

Unsere Aufgabe:

Aufbau der GLUES GDI für den Austausch wissenschaftlicher Daten

<http://geoportal.glues.geo.tu-dresden.de>

<http://modul-a.nachhaltiges-landmanagement.de/en/scientific-coordination-glues>

Aspekte einer GDI für wissenschaftliche (Geo-)Daten

- Eindeutige Identifikation von Datensätzen
- Metadaten
- Beschreibung der Entstehungsgeschichte
- Lizenzierung, Haftung
- Privatsphäre
- Organisatorische Fragen der GDI
- Nachhaltigkeit der Komponenten

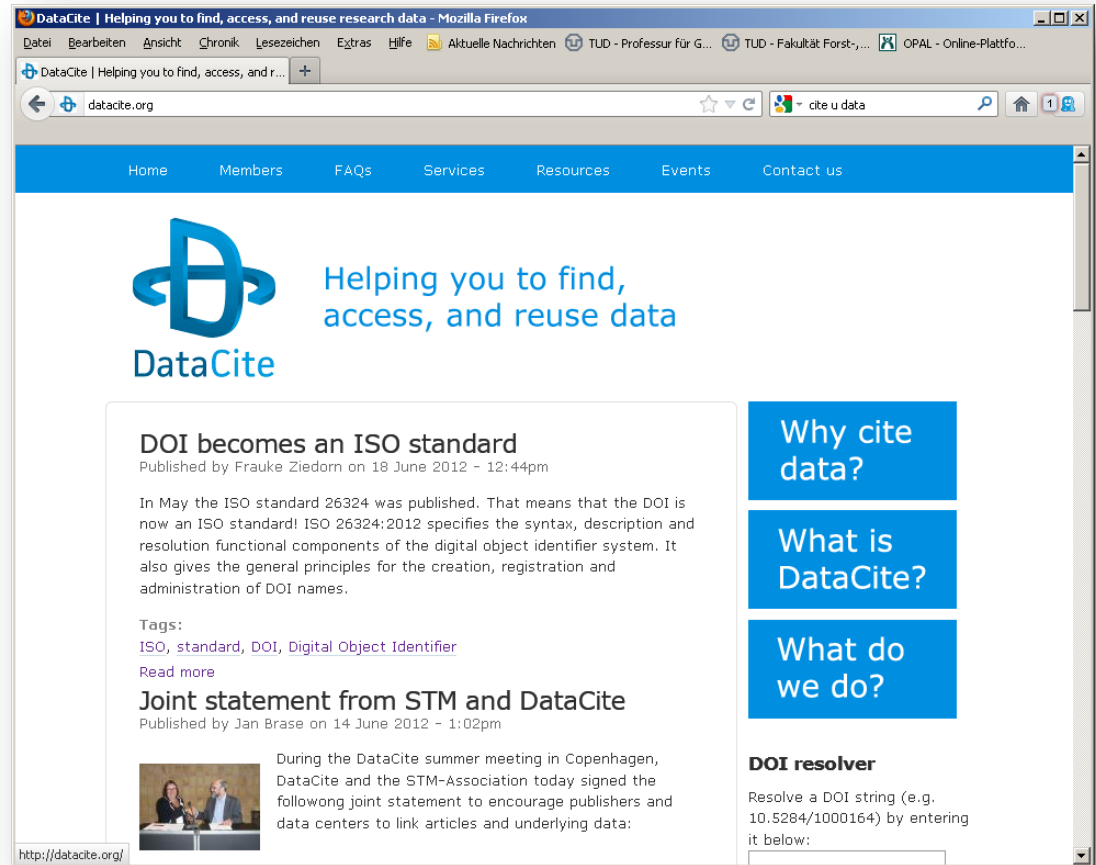
Referenzierung wissenschaftlicher (Geo-)Daten

Digital Object Identifier

- Eindeutige Identifikation
- Persistenz
- Möglichkeiten zum Auffinden und Zitieren
- Zentrale Registrierung

Offene Fragen:

- Anerkennung als Publikation durch entsprechende „Quotation Indexes“



<http://datacite.org/>

Metadaten für wissenschaftliche (Geo-)Daten

- Referenz-Vokabular (Thesauri) für Grundlegende wissenschaftliche Begriffe wie z.B. Modell, Szenario, Indikator, Treiber...
- Standardisierte Verweise auf Publikationen, Szenarien...(in derzeitigen Standards nicht vorgesehen)
- Automatisierte Erfassung

GLUES CATALOG

Detailsansicht
← Zurück zur Ergebnisliste

Datensatz:
PIK limited WCRP CMIP3 multi-model dataset CCSM3 SRESA1B Cloud cover

This dataset is part of the "PIK limited WCRP CMIP3 multi-model dataset". It combines climate anomalies from original model output of model CCSM3 for emissions scenario SRESA1B with a base climatology from CRU TS3.0. These are monthly high-resolution (0.5x0.5 degree) grids of total cloud fraction. Unit: Fraction. Temporal resolution: monthly.

Beschreibung | Kategorisierung | Zugriff | Vertrieb | Qualität | Bearbeitungsschritt | Metadaten

Beschreibung

Titel:	PIK limited WCRP CMIP3 multi-model dataset CCSM3 SRESA1B Cloud cover	
Kurztitel:	-	
Kurzbeschreibung:	This dataset is part of the "PIK limited WCRP CMIP3 multi-model dataset". It combines climate anomalies from original model output of model CCSM3 for emissions scenario SRESA1B with a base climatology from CRU TS3.0. These are monthly high-resolution (0.5x0.5 degree) grids of total cloud fraction. Unit: Fraction. Temporal resolution: monthly.	
Datum:	Erzeugung:	2012-08-08
	Publikation:	-
	Revision:	-
Kontakt:	Rolle:	Ansprechpartner
	Name:	Sebastian Ostberg
	Organisation:	Potsdam Institute for Climate Impact Research (PIK)
	Position:	-
	Telefonnummer:	-
	Faxnummer:	-
	Adresse:	P.O. Box 60 12 03
	Ort:	14412 Potsdam
	Administrative Einteilung:	-
	Land:	-
	Email Adresse:	ostberg@pik-potsdam.de
Bezugssystem:	WGS 84, lat/long (Geographische Koordinaten)	
Geografisches Begrenzungsrechteck (Gebietstyp: deckt ab):	Westliche Länge:	-180.0
	Östliche Länge:	180.0
	Nördliche Breite:	90.0
	Südliche Breite:	-90.0
Geografischer Identifikator:	-	
Zeitliche Beschränkung:	Anfang:	2006-01-01
	Ende:	2099-12-31
Räumliche Darstellungsart:	Raster	
Sprache:	Eng	
Zeichensatz:	utf8	
Identifikator:	ccsm3sresa1bcloudcover (Namensraum: urn:glues:pik:metadata:dataset)	

Entstehung und Verwendung der Daten

ISO 19115 Metadatenelement Lineage:

- Beschreibung der Entstehungsgeschichte des Datensatzes
 - für wissenschaftliche Modelloutputs:
 - Beschreibung der Inputdaten
 - Beschreibung des Modells / Modelldurchlaufs
- Interaktive Visualisierung der Entstehung des Geodatensatzes und der erfolgten Nutzungen, ...

Beispiel:

<http://geoportal.glues.geo.tu-dresden.de:8080/MetaViz/detail.jsp?id=glues:lm:metadata:dataset:promet>

Entstehung und Verwendung der Daten

METAVIZ

VIEWING DETAILS FOR DATASET PROMET [SHOW ALL CATALOG DATA](#)

DATASET
ECHAM5 A1B SCENARIO 1971-1990; 1990-2040; 2071-2100...
Ocean Circulation & Climate Dynamics Marine Meteo...

DATASET
HWSD - HARMONIZED WORLD SOIL DATABASE
International Institute for Applied Systems Analys...

DATASET
USGS SRTM30
NASA

DATASET
GLOBCOVER 2009
ESA

DATASET
GLUES REGIONS
TU Dresden

DATASET
GTAP 7 DATA BASE
Center for Global Trade Analysis, Purdue University.

MODEL
PROMET
Department of Geography, LMU Munich

DATASET
PROMET
Department fuer Geographie, LMU Muenchen

MODEL

MODEL

DATASET

DATASET

DETAILS

PROMET

Keywords: PROMET, Department fuer Geographie, LMU Muenchen, Crops, Barley, Maize, Millet, Rice, Rye, Wheat, Groundnut, Oilpalm, Rapeseed, Soy, Sunflower, Cassava, Potatoe, Sugarcane, Sorghum, potential yield, L...

Extent: WE: 180.0/-180.0; NS: 90.0/-90.0
Time: Time not set.

Children:
[PROMET Potential Yield Potential Yield for 2021-2040](#)
[PROMET Potential Yield Potential Yield for 1961-1990](#)
[PROMET Potential Yield Potential Yield for 2071-2100](#)
Parent: Parent not set.

Links: [METADATA](#)

[LINEAGE INFORMATION](#)

Integration von GIS-Funktionalitäten in Lösungen

Entstehung und Verwendung der Daten

METAVIZ

VIEWING DETAILS FOR DATASET PROMET [SHOW ALL CATALOG DATA](#)

DETAILS

PROMET

Keywords: PROMET, Department fuer Geographie, LMU Muenchen, Crops, Barley, Maize, Millet, Rice, Rye, Wheat, Groundnut, Oilpalm, Rapeseed, Soy, Sunflower, Cassava, Potatoe, Sugarcane, Sorghum, potential yield, L...

Extent: WE: 180.0/-180.0; NS: 90.0/-90.0
Time: Time not set.

Children:
[PROMET Potential Yield Potential Yield for 2021-2040](#)
[PROMET Potential Yield Potential Yield for 1961-1990](#)
[PROMET Potential Yield Potential Yield for 2071-2100](#)
Parent: Parent not set.

Links: [METADATA](#)
[LINEAGE INFORMATION](#)

Diagram:

- Left sidebar: 6 DATASET icons.
- Center: MODEL icon.
- Right sidebar: DATASET PROMET (Department fuer Geographie, LMU Muenchen).
- Bottom: MODEL CAPRI (Institute for Food and Resource Economics, Bonn Un...) and MODEL DART (Kiel institute for the World Economy).
- Far right: DATASET CAPRI (Institute for Food and Resource Economics, Bonn Un...) and DATASET DART (Kiel institute for the World Economy).

Integration von
GIS-Funktionalitäten
in Lösungen

Entstehung und Verwendung der Daten

METAVIZ

VIEWING DETAILS FOR DATASET DART [SHOW ALL CATALOG DATA](#)

DATASET
GTAP 7 DATA BASE
 Center for Global Trade Analysis, Purdue University.

DATASET
WORLD POPULATION PROSPECTS
 UN, Department of Economic and Social Affairs

DATASET
PROMET
 Department fuer Geographie, LMU Muenchen

MODEL
DART
 Kiel institute for the World Economy

DATASET
DART
 Kiel institute for the World Economy

DETAILS

DART

Keywords: DART, Kiel institute for the World Economy

Extent: WE:180.0/-180.0; NS:90.0/-90.0
2004-01-01 - 2030-12-31

Children: 24 children [click]
Parent: Parent not set.

Links: [METADATA](#)
[LINEAGE INFORMATION](#)

LINEAGE

PROCESS DESCRIPTION

RATIONALE not set

DESCRIPTION
 The DART model is a multi-region, multi-sector, recursive dynamic CGE model of the world economy and was developed at the Kiel Institute for the World Economy to analyse international climate policies in the late 1990s. It has been extended to capture land use change and the impact of biofuels. The model is based on microeconomic theory: in each of these regions, the economy is modelled as competitive economy with flexible prices and market clearing. Agents represented in the model are consumers, who maximise utility, producers, who maximise profits and regional governments. The regions are connected via bilateral trade

PUBLICATION

1. PUBLICATION

TITLE OF PUBLICATION Klepper, Gernot, Peterson, Sonja and Springer, Katrin (2003) DART97: A Description of the Multi-regional, Multi-sectoral Trade Model for the Analysis of Climate Policies, Kiel Working Papers, Kiel Institute for the World Economy.

PUBLISHING DATE 2003-02-01

CITATION <http://econpapers.repec.org/paper/kielkieliw1149.htm>

2. PUBLICATION

Integration von GIS-Funktionalitäten in Lösungen

Entstehung und Verwendung der Daten

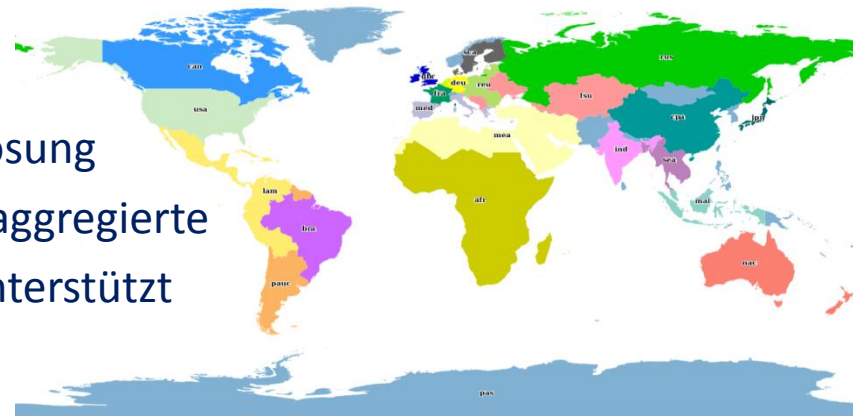
- z.B. auch für die Beschreibung der Herkunft INSPIRE konformer Datensätze sinnvoll

Probleme

- Identifikation der Metadatenätze
- ISO 19115 ist eher für Fernerkundungsdaten konzipiert
→ Modellbeschreibung ist schwierig
- Aufwändige Datenerfassung
- Verlinkung zwischen Modellen und Daten teilweise nur in eine Richtung navigierbar
- Hierarchische Strukturierung von Datensätzen und Subdatensätzen bisher in Katalogen kaum visualisiert und schlecht navigierbar
- CSW Schnittstelle bei großen Datenmengen zu langsam

Genauigkeit und Skalierung

- Genauigkeitsbeschreibung wissenschaftlicher Modelloutputs
 - Sind kaum verfügbar
 - ISO 19113 Elemente eher ungeeignet
 - Sinnvoller wären
 - Beschreibungen zu Abhängigkeiten der Qualität von In- und Outputdaten
 - Informationen welche Parameter den größten Einfluss auf die Qualität haben
 - Bedingungen unter denen Parameter als genau angenommen werden können
- Auflösung in ISO 19115
 - keine Beschreibung der zeitlichen Auflösung
 - Verlinkung von Referenzregionen (z.B. aggregierte Länder) bei statistischen Daten nicht unterstützt



Lizensierung wissenschaftlicher (Geo-)Daten

The screenshot shows the Creative Commons Attribution 3.0 Unported (CC BY 3.0) license page. The header features the Creative Commons logo and the text "Attribution 3.0 Unported (CC BY 3.0)". Below the header, there is a summary of the license and a "Disclaimer" link. The main content is divided into three sections: "You are free:", "Under the following conditions:", and "With the understanding that:". The "You are free:" section lists three freedoms: "to Share", "to Remix", and "to make commercial use of the work". The "Under the following conditions:" section lists the "Attribution" condition, which requires users to attribute the work to the author. The "With the understanding that:" section lists the "Waiver" and "Public Domain" conditions.

<http://creativecommons.org>

<http://creativecommons.org/science>

The screenshot shows the Open Data Commons website. The header features the text "Open Data Commons" and "Legal tools for Open Data". Below the header, there is a navigation menu with links for "Home", "Licenses", "Norms", "FAQ", "News", and "About". The main content area features a large heading "Open Data Commons is the home of a set of legal tools to help you provide and use Open Data". Below this heading, there are three bullet points: "Licenses and Dedications", "2-minute Guide to Making Your Data Open", and "Find Out More About the Project". At the bottom, there is a paragraph of text that reads: "If you're wondering about things like: why open data matters? or why do I need this legal stuff, can't I just post my data online? we suggest you check out the FAQ. If you want to know what we mean by open data visit the Open Definition which".

<http://opendatacommons.org/>

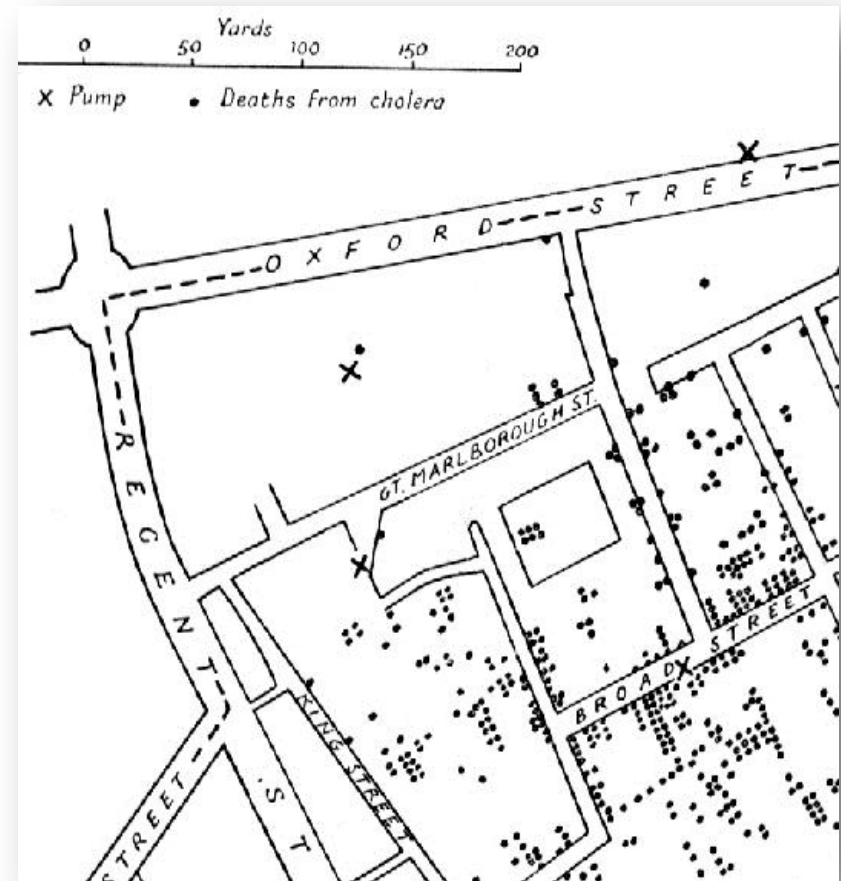


Persönlichkeits- und Datenschutz



www.eo2heaven.org

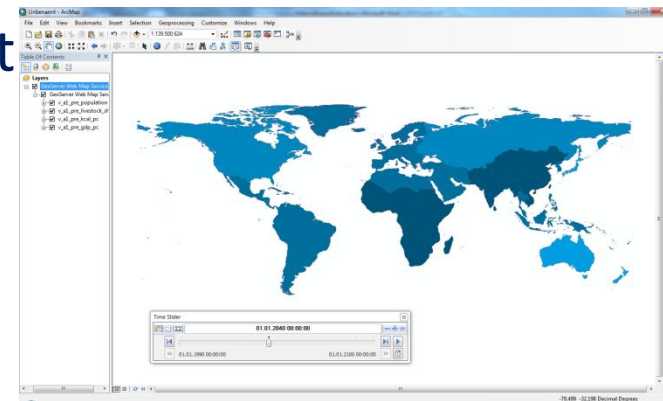
- aktuell ggf. etwas übersensibilisierte Diskussion
- Forschungsbedarf:
 - Formalisierung und Automatisierung der entsprechenden Datenschutzbestimmungen zur Anwendung in der Geoprozessierung (Ansätze aus Lizenzmodellierung)



Cholera-Tote in London
(John Snow 1854)

Herausforderungen für wissenschaftliche GDI

- Übersichtlichkeit bei der Suche in großen Datenmengen
<http://geoportal.glues.geo.tu-dresden.de/geoportal/Applications/geofacet/index.html>
- Umgang mit Zeit und Multidimensionalität
 - zeitvariante Geodaten sind inzwischen in Ansätzen technologisch beherrschbar
- Bereitstellung und Integration von Umweltmodellen / Modellkomponenten
- Integration von Echtzeitdaten



Organisatorische Fragen einer wissenschaftlichen GDI

- Motivation zur Publikation von Daten und Software
 - Anerkennung der Bereitstellung von Geodaten & Geosoftware als Publikation (,ISI', Quotation Index)
 - Passende Lizenzmodelle und Nutzungsmodelle
- Wissenschaftliche Infrastrukturleistungen dieser Art brauchen Nachhaltigkeit
 - entsprechende DFG, BMBF und EU Initiativen helfen, sind jedoch immer noch eher Programmforschung und weniger Langfrist-Förderung...
 - Disziplinübergreifende Organisationsstruktur
 - Einhaltung von Service Level Agreements

Fazit

- GDI bringen klare Vorteile für den Austausch wissenschaftlicher Daten
 - es gibt noch eine Vielzahl von Herausforderungen und Problemen, die schwerwiegendsten sind wohl:
 - organisatorischer Art
 - Überzeugungsarbeit, dass sich der anfängliche Mehraufwand lohnt
- Nutzerfreundlichkeit ist wichtigste Grundvoraussetzung

<http://geoportal.glues.geo.tu-dresden.de>