

Hydroinformatik I - WiSe 2020/2021

HyBHW-S1-01-10: BigData

Prof. Dr.-Ing. habil. Olaf Kolditz

¹Helmholtz Centre for Environmental Research – UFZ, Leipzig

²Technische Universität Dresden – TUD, Dresden

³Center for Advanced Water Research – CAWR

⁴TUBAF-UFZ Center for Environmental Geosciences – C-EGS, Freiberg / Leipzig

Dresden, 15.01.2021

Big Data

Big Data - Übersicht

1. Definition
 2. Forschungsbedarf
 3. Beispiele aus der Umwelt
 4. Wasser 4.0
-

5. pwd: Lehre2021!

Big Data Definition(en)

Big Data

Definition: Wikipedia

Der aus dem englischen Sprachraum stammende Begriff Big Data (von englisch big ‚groß‘ und data ‚Daten‘) bezeichnet Datenmengen, welche

- ▶ zu groß,
- ▶ zu komplex,
- ▶ zu schnelllebig oder
- ▶ zu schwach strukturiert

sind, um sie mit manuellen und herkömmlichen Methoden der Datenverarbeitung auszuwerten.[1] Im deutschsprachigen Raum ist der traditionellere Begriff Massendaten gebräuchlich. „Big Data“ wird häufig als Sammelbegriff für digitale Technologien verwendet, die in technischer Hinsicht für eine neue Ära digitaler Kommunikation und Verarbeitung und in sozialer Hinsicht für einen gesellschaftlichen Umbruch verantwortlich gemacht werden.[2] Er steht dabei grundsätzlich für große digitale Datenmengen, aber auch für deren Analyse, Nutzung, Sammlung, Verwertung und Vermarktung.[3]

Source: https://de.wikipedia.org/wiki/Big_Data

Big Data

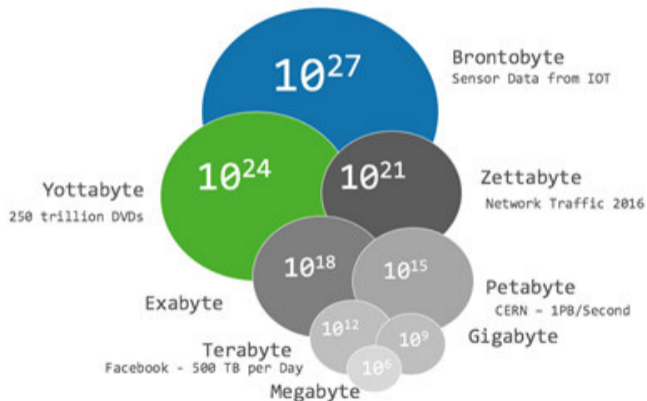
Definition: Wikipedia

In der Definition von Big Data bezieht sich das „Big“ auf die drei Dimensionen volume (Umfang, Datenvolumen), velocity (Geschwindigkeit, mit der die Datenmengen generiert und transferiert werden) sowie variety (Bandbreite der Datentypen und -quellen).[4] Erweitert wird diese Definition um die zwei V's value und validity, welche für einen unternehmerischen Mehrwert und die Sicherstellung der Datenqualität stehen.[5] Der Begriff „Big Data“ unterliegt als Schlagwort einem kontinuierlichen Wandel; so wird mit ihm ergänzend auch oft der Komplex der Technologien beschrieben, die zum Sammeln und Auswerten dieser Datenmengen verwendet werden.[6][7] Die gesammelten Daten können dabei aus verschiedensten Quellen stammen (Auswahl):

Source: https://de.wikipedia.org/wiki/Big_Data

Big Data

Definition



Source: <http://api.ning.com>

Big Data

Definition

Big Data (Definition)

- **V**olume (Datenmenge -> Fernerkundung)
 - **V**ariety (Heterogenität -> multivariante Daten ...)
 - **V**elocity (Geschwindigkeit -> in-situ Visualisierung ...)
 - **V**eracity (Wahrheitsgehalt -> Unsicherheiten ...)
 - **V**isualization (in-situ, VISLAB 2.0)
- **V**alue

Data Management

Q

Visual Analytics

Big Data

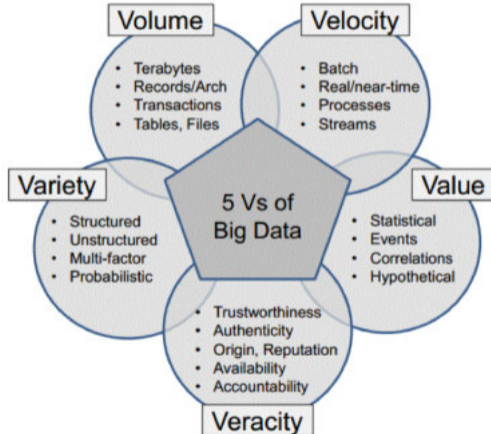
Definition



Source <https://www.i-scoop.eu/big-data-action-value-context/>

Big Data

Definition



10/31

Big Data Forschung

Big Data

Forschung

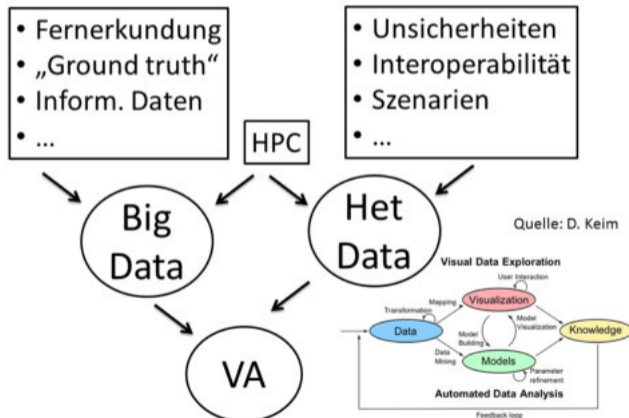
Forschungsbedarf (aus Sicht der Visualisierung)

- Challenge: **Big Data**, z.B. aus Fernerkundung (Hyperspektraldaten), hochaufgelöste Geophysik und Simulationsergebnisse (High-Performance-Computing)
- Methodik: **Visual Analytics** (neue Erkenntnisse gewinnen aus großen und heterogenen Datenmengen)
- Methodik: Interaktive **Workflows**: automatisierte Datenanalyse, Qualitätsmanagement (Validierung), visuelle Datenexploration, Mustererkennung, in-situ Visualisierung
- Sichtbarkeit: EUROVIS Workshop „EnvirVis“ (EES TI)

Big Data

Forschung

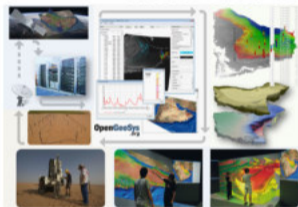
Visual Analytics (VA)



Big Data

Forschung

Interaktive Workflows



- **In-situ Visualisierung**
(Echtzeitvisualisierung
modellierter Daten,
VISLAB 2.0)

Data
Visualization

Data
Integration

Data
Simulation

Big Data

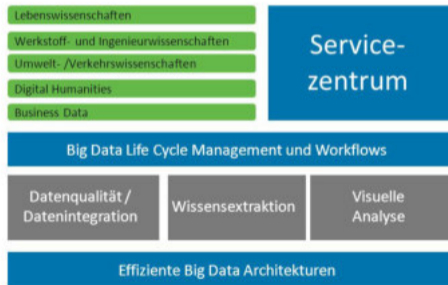
Forschung: Umweltinformationssysteme (www.ufz.de/vislab)



Big Data

Forschung: ScaDS >> ScaDS.AI

 GROBSTRUKTUR DES ZENTRUMS



www.scds.de

Source www.scds.de

Big Data

Beispiele aus der Umwelt

Big Data

Behörden



The screenshot shows the website of the Umwelt Bundesamt (UBA). The top navigation bar includes the UBA logo and several menu items: 'Das UBA', 'Themen', 'Presse', 'Publikationen' (highlighted in blue), 'Tipps', and 'Daten'. Below the navigation bar, the breadcrumb trail reads 'Publikationen > Umweltinformationssysteme'. The main content area features a grid of categories: 'Nachhaltigkeit | Strategien | Internationales'. The selected article is titled 'Umweltinformationssysteme' with the subtitle 'Big Data – Open Data – Data Variety'. To the left of the article is a thumbnail image of the publication cover, dated '5/8/2015'. The article text discusses the challenges of handling large data volumes in environmental information systems and the need for clear conceptual alignment and appropriate tools.

Umwelt Bundesamt

Das UBA Themen Presse **Publikationen** Tipps Daten

Publikationen > Umweltinformationssysteme

Nachhaltigkeit | Strategien | Internationales

Umweltinformationssysteme

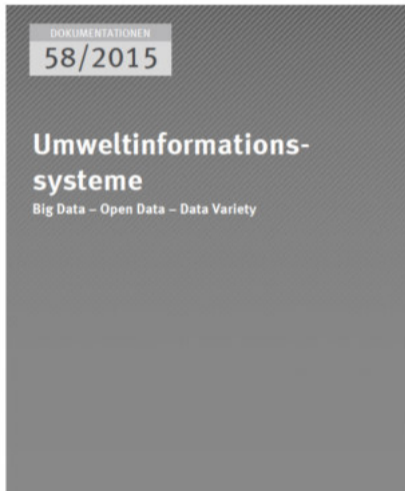
Big Data – Open Data – Data Variety

Mit dem Zuwachs an Daten und Informationen in Datenbanken und Informationssystemen steht die angewandte Umweltinformatik ambitionierten Aufgaben gegenüber. Die Aufbereitung dieser großen Datenmengen für verschiedene Zielgruppen erfordert eine klare konzeptionelle Ausrichtung der Anwendungsentwicklungen. Die Nutzung von Methoden und Werkzeugen der Informatik ist ein Weg, Produkte und Dienste aus diesen Daten zu generieren. Die Bereitstellung dieser Produkte für die verschiedenen Nutzergruppen wie der wissenschaftlichen Community der Modellierer, der Fachnutzer in Umweltbehörden oder eine App auf mobilen Endgeräten für die Öffentlichkeit spiegeln die breite Vielfalt von Datenangeboten wider.

Source: UBA

Big Data

Behörden



Themen (Workshop 2015)

- ▶ Umwelt-Sensordaten (SOS Web Services)
- ▶ Metadaten
- ▶ Geodateninfrastrukturen
- ▶ Cloud Computing für die Kalibrierung von Hochwassersimulationen
- ▶ Geovisualisierung
- ▶ Crowdsourcing
- ▶ ...

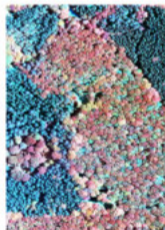
Big Data

Forschung: Fernerkundung

Erfassung und Auswertung von Hyperspektraldaten

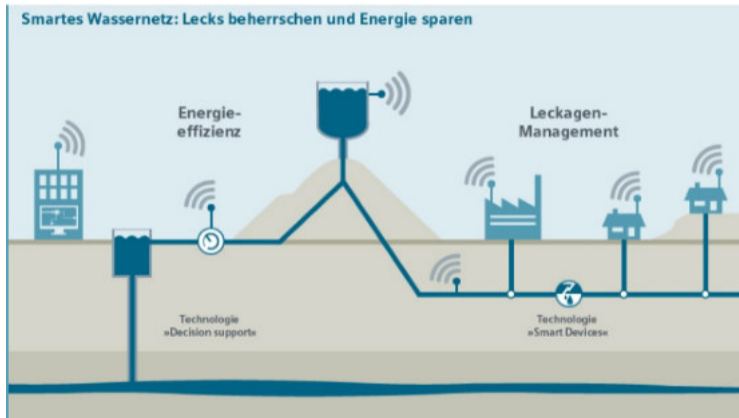
Früherkennung von Waldschäden

- Vorbereitung schneller, sicherer und umweltschonender Maßnahmen zur Früherkennung und Prävention von Kalamitätsfällen im Forstbereich
- Hochdimensionale funktionale und hochaufgelöste 3D-Rasterdaten aus Befliegungen von Waldgebieten
- Eine Vielzahl von Einflussfaktoren überlagert das hochdimensionale Signal
- Kalibrierung der Spektraldaten durch Methoden des maschinellen Lernens
- Algorithmen zur Separierung einzelner Baumkronen und Ableitung deren hyperspektraler Signatur
- Aussagen zur Baumart, Vitalität sowie zu biotischen und abiotischen Stressfaktoren



Big Data

Industrie: Von Big to Smart Data



Source: Siemens

<https://www.siemens.com/innovation/de/home/pictures-of-the-future/digitalisierung-und-software/von-big-data-zu-smart-data-projekt-icewater.html>

Big Data Wasser 4.0

Big Data

Industrie 4.0

Industrie 4.0 ist ein Begriff, der auf die Forschungsunion der deutschen Bundesregierung und ein gleichnamiges Projekt in der Hightech-Strategie der Bundesregierung zurückgeht; zudem bezeichnet er ebenfalls eine Forschungsplattform.[1][2][3] Die industrielle Produktion soll mit moderner Informations- und Kommunikationstechnik verzahnt werden.[4] Technische Grundlage hierfür sind intelligente und digital vernetzte Systeme. Mit ihrer Hilfe soll eine weitestgehend selbstorganisierte Produktion möglich werden: Menschen, Maschinen, Anlagen, Logistik und Produkte kommunizieren und kooperieren in der Industrie 4.0 direkt miteinander.[4] Durch die Vernetzung soll es möglich werden, nicht mehr nur einen Produktionsschritt, sondern eine ganze Wertschöpfungskette zu optimieren. Das Netz soll zudem alle Phasen des Lebenszyklus des Produktes einschließen – von der Idee eines Produkts über die Entwicklung, Fertigung, Nutzung und Wartung bis hin zum Recycling.[4]

Source: Wikipedia

Big Data

Industrie 4.0

Mit der Bezeichnung "Industrie 4.0" soll das Ziel zum Ausdruck gebracht werden, eine vierte industrielle Revolution einzuleiten.

1. Die erste industrielle Revolution bestand in der Mechanisierung mit Wasser- und Dampfkraft, darauf folgte
2. die zweite industrielle Revolution: Massenfertigung mit Hilfe von Fließbändern und elektrischer Energie,
3. daran anschließend die dritte industrielle Revolution oder digitale Revolution mit Einsatz von Elektronik und IT (v. a. die speicherprogrammierbare Steuerung) zur Automatisierung der Produktion.

Mit dem Ausdruck „4.0“ wird Bezug genommen auf die bei Software-Produkten übliche Versionsbezeichnung, die bei größeren Änderungen von einer neuen Version spricht, die erste Ziffer der Versionsnummer um Eins erhöht und gleichzeitig die zweite Ziffer auf Null zurücksetzt.

Source: Wikipedia

Big Data

Wasser 4.0

GWP-Arbeitskreis Wasser 4.0



Big Data

Wasser 4.0

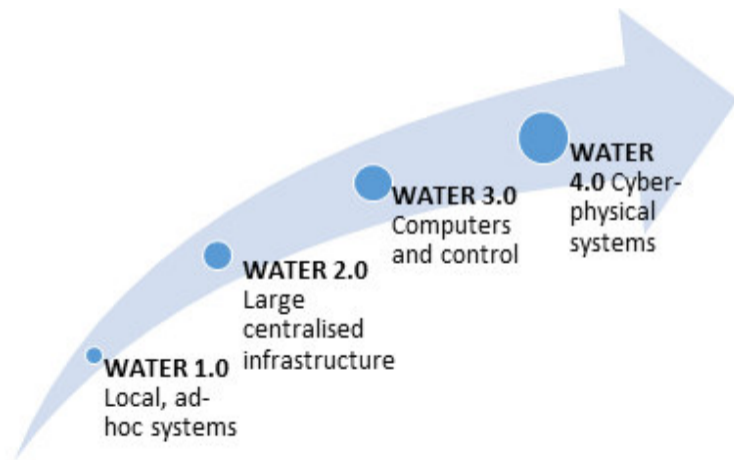
Wasser 4.0 stellt die Digitalisierung und Automatisierung in den Mittelpunkt einer Strategie für eine ressourceneffiziente, flexible und wettbewerbsfähige Wasserwirtschaft.

Source: German Water Partnership

<http://www.germanwaterpartnership.de/de/arbeitskreise/wasser-40/index.htm>

Big Data

Wasser 4.0

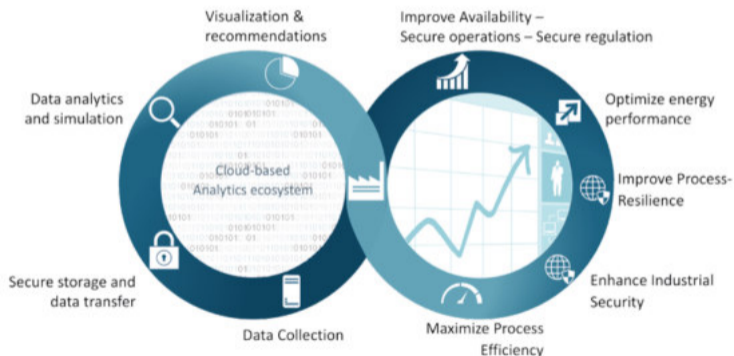


Big Data

Wasser 4.0

From Data...

...to Value



Big Data

Wasser 4.0

1st Industrial Revolution

Mechanization of work, powered by water and steam



First mechanical loom, 1784

2nd Industrial Revolution

Specialized mass production of goods, powered by electricity



First production line, slaughterhouse Cincinnati, 1870

3rd Industrial Revolution

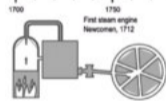
Electronics and information technology facilitate automated production; ICT gives rise to computerization



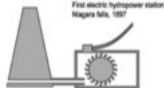
First programmable logic controller (PLC, Modicon 084, 1969)

4th Industrial Revolution

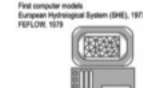
Intelligent devices in intelligent global networks provide permanent availability and analysis of data and information; Merging of physical and virtual worlds into Cyber-Physical Systems (CPS); Internet of Things (and services)



First steam engine
Newcomen, 1712



First electric hydropower station,
Niagara falls, 1897



First computer models
European Hydrological System (SHE), 1977
FEFLOW, 1979



1st Revolution in Water Management

Utilisation of steel to handle high water pressure (steam boilers, hydraulic steelwork)

2nd Revolution in Water Management

Pumps and turbines use and generate electricity

3rd Revolution in Water Management

IT capable of physical-numerical modelling of water systems; Integration of field sensors into IT systems

4th Revolution in Water Management

Interfacing of real and virtual water systems (CPS); Real-time and forecasting models reduce risks and costs; Distribution and collection concepts include Internet-based networking through to the end user (Smart sensing)

Big Data

Wasser 4.0

Engineering 5 (2019) 828–832



Contents lists available at [ScienceDirect](#)

Engineering

journal homepage: www.elsevier.com/locate/eng



Views & Comments

Environmental Information Systems: Paving the Path for Digitally Facilitated Water Management (Water 4.0)



Olaf Kolditz^{a,b,e}, Karsten Rink^a, Erik Nixdorf^a, Thomas Fischer^a, Lars Bilke^a, Dmitri Naumov^a, Zhenliang Liao^{c,e}, Tianxiang Yue^{d,e}

^a Department of Environmental Informatics, Helmholtz Center for Environmental Research (UFZ), Leipzig 04318, Germany

^b Applied Environmental Systems Analysis, Technische Universität Dresden, Dresden 01069, Germany

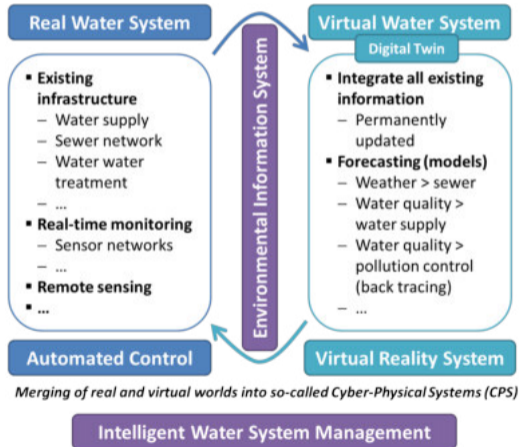
^c UN Environment–Tongji Institute of Environment for Sustainable Development & College of Environmental Science and Engineering, Tongji University, Shanghai 200092, China

^d Department for Ecological and Environmental Informatics, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 1000101, China

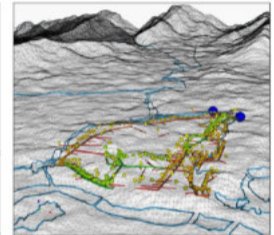
^e Sino-German Research Center for Environmental Information Science (RCEIS), Leipzig 04318, Germany

Big Data

Wasser 4.0 (www.ufz.de/vislab)



(a)



(b)

