

OCTOBER 23 · FRIDAY

09:00 – 09:45	Anmeldung	_Erich-Brost-Haus
10:00 – 10:30	R Begrüßung <i>Referenten: Julia Stein, Metin Tolan, Holger Wormer</i>	Saal (Erich-Brost-Haus, TU Dortmund)
10:30 – 11:00	K How to use science to make better data journalism projects [E] <i>Referenten: Lena Groeger</i> How to use science to make better data journalism projects Discussions of data journalism often focus the creator: how one should parse and analyze data, how one should visualize or write about it, how one should pick the best tools for the job. This talk will instead focus on the user: how they interpret a piece of data journalism in the wild, how they understand the meaning behind our words and visuals (or fail to). But we don't have to rely on assumptions about how users think and behave – science has a lot to say on the matter. New findings in psychology, neuroscience, and behavioral science can give us a better understanding of how the mind works when it goes about interpreting a piece of data journalism, and can help us make better, clearer, and more effective projects. Let's see how. Video: https://www.youtube.com/watch?v=WF5-M0zliTo&list=PLr8RTcS5UiniN1S9Ohws3P5NyUasnz9ix&index=1	Saal (Erich-Brost-Haus, TU Dortmund)
11:00 – 12:00	P Fruchtbare Schnittstelle oder vermintes Gelände? Wie Datenjournalismus und Wissenschaften voneinander profitieren <i>Moderatoren: Holger Wormer</i> <i>Referenten: Stefan Hornbostel, Katharina Morik, Ralf Spiller, Volker Stollorz</i> Auftaktdiskussion der Tagung Fruchtbare Schnittstelle oder vermintes Gelände? Wie Datenjournalismus und Wissenschaften voneinander profitieren	Saal (Erich-Brost-Haus, TU Dortmund)
12:00 – 13:00	G Lunch	_Erich-Brost-Haus
13:00 – 14:00	K Unsicherheiten auf Karten visualisieren Klimakarten und Klimawandel als Datenerkenntnis <i>Referenten: Jochen Schiewe, Birgit Schneider</i> Wie lassen sich unsichere Informationen in Medienkarten darstellen? Eigentlich ist es allen Beteiligten – Erstellern und Nutzern von Medienkarten – bewusst, dass die dort dargestellten Informationen gewisse Unsicherheiten in sich bergen – seien es nicht korrekte Positionen von Unglücksorten, ungenau Verläufe von Orkanen, geschätzte Zahlen von Flüchtlingen, etc. Doch nur äußerst selten werden solche Unsicherheiten auch in Karten kommuniziert. Basierend auf diversen Forschungs- und Entwicklungsarbeiten in diesem Themengebiet geht dieser Beitrag darauf ein, wann eine Visualisierung von Unsicherheiten überhaupt sinnvoll ist und welche Möglichkeiten dafür zur Verfügung stehen. Prof. Dr. Jochen Schiewe (HafenCity Universität Hamburg) Die Verdattung der Atmosphäre. Klima und Klimawandel als Datenerkenntnis seit 1800 Zahlreiche wissenschaftliche Bilder vom Klima bestehen Bilder aus Daten. Mittels Methoden der Visualisierung wird das Klima zu einem Forschungsobjekt und damit zu einer Erkenntnis gemacht. Der Zusammenhang von Daten und Visualisierungsmethoden stand bereits am Beginn der modernen Klimaforschung um 1800 als Alexander von Humboldt erstmals Messwerte in Form einer Klimazonenkarte darstellte. Die Visualisierung von Daten war aber auch für die Erkenntnis des anthropogenen Klimawandels sowie für dessen Kommunikation bedeutend. Denn auch gegenwärtig werden die viele der wichtigsten Beweise des Klimawandels in Form von Datenbildern angeführt – weshalb es andererseits auch diese Visualisierungen sind, die von Klimawandelleugnern gezielt in Zweifel gezogen werden. In meinem Vortrag werde ich aus bild- und medienwissenschaftlicher sowie aus historischer Perspektive Beispiele beleuchten, die diskutieren lassen, wo Konfliktpunkte klimawissenschaftlicher Datenbilder und ihrer Kommunikation in außerwissenschaftliche Kreise liegen. Dr. Birgit Schneider (FH Potsdam, Europäische Medienwissenschaft)	Saal (Erich-Brost-Haus, TU Dortmund)

13:00 – 14:00

W **Workshop: Eine Data-Mining-Leitlinie für guten Datenjournalismus?**

Referenten: *Anna Behrend*

Workshop-Raum (Erich-Brost-Haus, TU Dortmund)

Eine Data-Mining-Leitlinie für guten Datenjournalismus?

Wissen aus Daten generieren: Das wollen nicht nur Journalisten und Forscher, sondern auch Unternehmen. Damit nicht jeder neue Datenanalyst in die gleichen Fallen tappt, verwenden viele Firmen seit den Neunzigerjahren einen Standardprozess für sogenannte Data-Mining-Projekte. Anna Behrend hat in ihrer Bachelorarbeit gezeigt, dass dieser Prozess große Gemeinsamkeiten mit datenjournalistischen Arbeitsweisen aufweist. Kann er als Leitfaden für guten Datenjournalismus dienen? Die Ergebnisse der Bachelorarbeit dienen im Workshop als Diskussionsgrundlage für diese Frage.

Anna Behrend (Daten- und Wissenschaftsjournalistin, Berlin)

14:30 – 15:30

K **Journalismus in den Datenfriedhöfen der Medizin I Informationsansprüche und ihre Grenzen**

Referenten: *Mario Martini, Volker Stollorz*

Saal (Erich-Brost-Haus, TU Dortmund)

Datenjournalismus in Datenfriedhöfen – über die Tücken des „Public Reporting“ in der Medizin

Datenjournalismus über die Qualität der medizinischen Versorgung in Deutschland ist ein mühsames Geschäft. Es mangelt an relevanten Daten. Der Flickenteppich der vorhandenen Datenquellen erzeugt vor allem Scheintransparenz. Wie schwer es ist, sich als Journalist im Dschungel der Versorgungsdaten zu orientieren, erläutert der Vortrag anhand einer Reihe von Recherche-Werkzeugen.

Volker Stollorz (Redaktionsleiter Science Media Center Germany)

Informationsansprüche von Journalisten gegenüber Behörden - Anspruchsgrundlagen und ihre Grenzen

Die Information der Öffentlichkeit durch Journalisten ist der Nährboden demokratischer Meinungsbildung. Gesetzliche Informationsansprüche der Presse und der Öffentlichkeit kollidieren mit Schutzinteressen des Funktionierens staatlicher Einrichtungen, der Allgemeinheit und betroffener Einzelner. Wie das Recht diese Konfliktslage auflöst bzw. auflösen sollte, erläutert Prof. Martini in einem Überblick.

Prof. Mario Martini (Deutsche Universität für Verwaltungswissenschaften Speyer)

14:30 – 15:30

W **Kartographie-Workshop: So gelingen Daten-Karten**

Workshop-Raum (Erich-Brost-Haus, TU Dortmund)

Referenten: *Hannes Kröger, Julius Tröger*

Teilnehmer lernen anhand eines Beispiels aus der Praxis, wie man innerhalb kurzer Zeit mit aktuellen, kostenlosen Tools interaktive Karten erstellt. In dem Workshop werden die wichtigsten Schritte und Regeln von der Datenaufbereitung bis hin zur Gestaltung (z.B. Wahl der Farben) besprochen. Dabei werden auch typische Fehler bei der Kartenserstellung gezeigt – und wie diese vermieden werden können.

Wer mitmachen möchte, sollte sich ein kostenloses Profil bei Mapbox anlegen und die ebenfalls kostenlose Software Tilemill herunterladen – bitte diese Version für Mac und diese für Windows.

16:00 – 17:00

K **Über den Gestaltungsprozess von Visualisierungen I Daten-Storytelling (auch) unter Zeitdruck**

Referenten: *Till Nagel, Sascha Venohr*

Saal (Erich-Brost-Haus, TU Dortmund)

Mäandern. Über den Gestaltungsprozess von Visualisierungen

In diesem Vortrag wird Till Nagel über den Gestaltungsprozess von Datenvisualisierungen sprechen. Anhand von Beispielen zeigt er, wie Daten in lesbare und visuell ansprechende Form abgebildet werden können.

Diese anschaulichen Beispielprojekte reichen von einer interaktiven Darstellung von Verkehrsströmen, über eine abstrakte Visualisierung von Flugrouten, bis hin zu einem Kartogramm von Shanghai. Er dokumentiert Entwürfe und Experimente, die bei der Analyse und Synthese entstanden, und diskutiert wie Ideen erneut aufgegriffen und auch projektübergreifend eingesetzt werden können.

Till Nagel (Urban Complexity Lab, FH Potsdam)

Daten-Storytelling (auch) unter Zeitdruck

Das ZEIT ONLINE Datenteam produziert neben dem Tagesgeschäft regelmäßig umfangreiche interaktive Geschichten. Wie strukturiert das Team seine Projekte und welche Schritte sind bei der Entstehung essentiell (obwohl sie Zeit kosten)? Sascha Venohr stellt einzelne datenbasierte Geschichte vor und erzählt, wie der Workflow im Hintergrund ablief und welche Hindernisse bis zum Livegang zu überwinden waren.

Sascha Venohr (Leiter Team Datenjournalismus, ZEIT ONLINE)

- 16:00 – 17:00 W **Workshop: Daten von Behörden erfragen & erforschen** Workshop-Raum (Erich-Brost-Haus, TU Dortmund)
Referenten: Stefan Wehrmeyer
Workshop: Daten von Behörden erfragen & erforschen
Daten von Behörden finden sich in Open-Data-Portalen, hinten kruden Suchmasken oder sind erst gar nicht online. Welche Möglichkeiten habe ich an behördliche Datensätze zu kommen? Doch der Erhalt von Daten ist oftmals erst der Anfang, denn behördliche Datensätze sind oft voller Fallstricke. Worauf muss ich achten, um die Daten richtig zu verstehen und weiterzuverarbeiten?
Stefan Wehrmeyer (Fragdenstaat / Correctiv, Berlin)
-
- 17:30 – 18:30 K **The Jupyter Project: Opening cultures, from science to data-driven journalism [E]** Saal (Erich-Brost-Haus, TU Dortmund)
Referenten: Fernando Perez
The Jupyter Project: Opening cultures, from science to data-driven journalism
In this talk, Fernando Pérez will briefly describe the Jupyter project (evolved from IPython), and how its flagship, the Jupyter Notebook, allows open collaboration to emerge in a variety of contexts: A tool that was originally built for scientific computing today enables data-driven journalists, educators, bloggers and more to share narratives that weave equal parts computation, data and natural language to tell a story. We will see how these tools can support rich data-driven journalistic narratives, hopefully leading to a richer dialog between the journalists and a more informed public.
Fernando Pérez, PhD (Research scientist at U.C. Berkeley, Brain Imaging Center, founder of iPython)
-
- 17:30 – 18:30 W **Design/Dramaturgie-Workshop: Wie Daten Geschichten erzählen** Workshop-Raum (Erich-Brost-Haus, TU Dortmund)
Referenten: Michael Hörz, Jan-Erik Stange
Wie Daten Geschichten erzählen: Workshop mit konkreten Beispielen
Der Workshop zielt darauf ab, ein gemeinsam vereinbartes Thema in einer Hands-On-Session gestalterisch zu vermitteln. Die Teilnehmer berücksichtigen dabei unterschiedliche Storytelling-Methoden, Perspektiven und Ausrichtungen an Zielgruppen. Am Anfang werden wir Praxis-Beispiele untersuchen, anschließend arbeiten die Teilnehmer in kleinen Teams ihre Linie heraus und diskutieren sie in der Gruppe. Zum Abschluss ordnen wir ein, wie unterschiedliche Darstellungsformen auch auf unterschiedliche Weise die Wahrnehmung des Lesers beeinflussen können, positiv wie negativ.
-
- 19:00 – 22:00 G **Get together** _Erich-Brost-Haus

OCTOBER 24 • SATURDAY

- 10:00 – 11:00 K **How algorithms find and spread news [E]** Saal (Erich-Brost-Haus, TU Dortmund)
Referenten: Marco Maas, Neil Thurman
Coaching Computers to Detect and Verify News from Social Media
The use of social media as a source of news is entering a new phase as computer algorithms are developed and deployed to detect, rank, and verify news. The efficacy and ethics of such technology is the subject of this talk, which examines once such tool, Social Sensor. Social Sensor is being used successfully to identify trending news stories, allow journalists to search within social media, and help verify social media contributors and content. However such software also raises questions about who has influence as social media is algorithmically-filtered for the benefit of journalists and others.
Prof. Dr. Neil Thurman (Institut für Kommunikationswissenschaft und Medienforschung, LMU München)
The news will find you - the question is how
News organisations have to face a constant change in how users are consuming their stories. Be it print, desktop, tablet, smartwatch, smart-tv or other devices - we'll have to understand in which situation at which time which content ist best for our customers. The Internet of things will help us to get a better understanding on the situation. We'll need to learn that the next learning curve for us as news producer will be how we can use sensors in a privacy-acceptable way to serve news stories in the future.
Marco Maas (OpenDataCity)
-

10:00 – 11:00

W Workshop: Wer mit Wem - Vernetzungen auf der Spur! Von Visualisierung bis Analyse

Referenten: Mareike Bockholt, Katharina Anna Zweig Workshop-Raum (Erich-Brost-Haus, TU Dortmund)

Ob Geldströme zwischen Organisationen, persönliche Bekanntschaften zwischen Vorständen oder Interaktionen zwischen Zellproteinen: in der Vernetzungsstruktur von Personen und Dingen steckt Information, die sich oft gut visualisieren lässt, aber zusätzlich auch der mathematischen Analyse bedarf: wer ist am zentralsten? Welche Gruppen sind besonders stark vernetzt? Wo sind ungewöhnliche Muster zu finden? All diese Fragen beantwortet die Netzwerkanalyse. In diesem Workshop zeigt Prof. Dr. Katharina Zweig, TU Kaiserslautern, wie man aus einer Excel-Tabelle mit Vernetzungsinformationen eine schöne Visualisierung erstellt und diese mit Hilfe von Softwaretools analysiert, zeigt aber auch, dass eine zu schnelle Analyse zu Fehleinschätzungen führen kann.

Wer live mitmachen will, möge bitte die folgenden Anwendungen installieren:

1) yED von der Firma yworks <http://www.yworks.com/en/products/yfiles/yed/> (frei verfügbar)

Das Manual findet sich hier, für diejenigen, die später stärker einsteigen oder schon mal herumspielen wollen: <http://yed.yworks.com/support/tutorial/index.html>

2) visone <http://visone.info/index.html>

11:30 – 12:30

K Die Statistik hinter dem Zensus | Reproduzierbarer Datenjournalismus

Referenten: Timo Grossenbacher, Ralf Münnich

Saal (Erich-Brost-Haus, TU Dortmund)

Einwohnerzahlen via Stichprobe: die Statistik hinter der Volkszählung

2011 wurde in Deutschland der erste Register-gestützte Zensus durchgeführt. Im Rahmen des Vortrages soll die Methodik, welche Register- mit Stichprobeninformationen geeignet kombiniert, vorgestellt und diskutiert werden. Dabei sollen die Auswirkungen gesetzlicher Rahmenbedingungen auf die Festlegungen und Qualität des Zensus dargelegt werden, wie etwa die Bedeutung der statistischen Qualität, der Einfluss des Meldewesens sowie die Anforderungen an Gemeinden um 10.000 Einwohner.

Professor Dr. Ralf Münnich (Universität Trier, Fachbereich IV, Volkswirtschaftslehre)

Reproducible data-driven journalism – challenges and tools

Reproducibility is the next big thing in data journalism. Done correctly, it solves two major problems: 1) Working collaboratively on datasets, for example with Excel, is error-prone and hard to quality-control. Making data processing steps reproducible also means making them understandable and controllable - e.g. by team members. 2) While sophisticated data collection or even analysis might result in nice scoops and awesome stories, it is often hard to explain to the audience what has happened behind the curtains. In order to be credible, we should be as transparent as possible - making methods open and in the best case fully reproducible allows our audience to understand, criticize and correct or even alleviate our methodology.

In his talk, Timo Grossenbacher from SRF Data will first introduce the principles of reproducibility, why the science community has so far mostly failed achieving it and how data journalism could become better by applying it throughout. In the second part, he will demo the processes and tools he uses in his daily work, e.g. on arms exports.

Timo Grossenbacher (Schweizer Rundfunk und Fernsehen, Reporter im Data-Team)

11:30 – 12:30

W Informatik-Workshop Data-Mining: Adaptive Informationsaufbereitung aus heterogenen Quellen
Referenten: Iryna Gurevych, Kurt Jansson, Christian Meyer, Workshop-Raum (Erich-Brost-Haus, TU Dortmund)
Maximilian Schäfer

Adaptive Informationsaufbereitung aus heterogenen Quellen

Herausforderungen und Werkzeuge

1) Prof. Dr. Iryna Gurevych und Christian Meyer

(TU Darmstadt, Ubiquitous Knowledge Processing (UKP) Lab)

Intensive Recherche ist zentral für sämtliche journalistischen Tätigkeiten und unterliegt sehr engen Zeitvorgaben. Zugleich explodiert die relevante Informationsmenge und die Quellen werden immer komplexer, sind hochgradig heterogen und weisen unterschiedliche Informationsqualität auf.

Im Rahmen dieses Workshops stellen wir das von den Universitäten Darmstadt und Heidelberg sowie dem Heidelberger Institut für Theoretische Studien (HITS) gegründete Graduiertenkolleg AIPHES vor und diskutieren wesentliche Herausforderungen und Werkzeuge aus Sicht des Journalismus und der informatikorientierten Forschung zu sprach- und wissensverarbeitenden Systemen.

Das gemeinsame Anwendungsszenario des Graduiertenkollegs ist es, relevantes Wissen aus heterogenen Textquellen automatisch zu extrahieren und zu einem informativen und stilistisch homogenen Dossier aufzubereiten, das als Recherchehilfsmittel oder Schreibvorlage für Online-Redaktionen genutzt werden kann.

2) Kurt Jansson und Maximilian Schäfer (Dokumentation DER SPIEGEL, Hamburg)

Das Zu-Tage-Fördern von strukturierten Daten aus diversen Formaten und Quellen gehört beim SPIEGEL seit einigen Jahren zum Alltag. Die Szenarien sind vielfältig: Uns erreichen kiloschwere Aktenordner aus einer Behörde, ein Redakteur erhält von einem Informanten einen USB-Stick mit unklarem Inhalt oder wir benötigen Informationen, die über tausende Webseiten verstreut sind.

Mittlerweile verfügen wir über einen erprobten Werkzeugkasten, mit dem wir Daten auch aus widerspenstigen Formaten extrahieren und analysieren können. Während des Workshops geben wir einen Einblick in unsere Arbeitsweise.

13:00 – 14:00	G	Lunch	_Erich-Brost-Haus
14:00 – 16:00	K	Präsentation der Workshop-Ergebnisse Woran haben die Teilnehmer der sechs Workshops gearbeitet? Ausgewählte Ergebnisse sollen hier noch einmal dem Plenum präsentiert und mit allen Besuchern der Konferenz diskutiert werden.	Saal (Erich-Brost-Haus, TU Dortmund)
16:00 – 16:30	R	Schlussworte	Saal (Erich-Brost-Haus, TU Dortmund)
