

Big Data, Open Geography? Reflektierter Umgang mit Big Data in den raumbezogenen Wissenschaften

Montag, 26. und Dienstag, 27. März 2018

**Leibniz-Institut für Länderkunde
Leipzig**

Die Gewinnung, Analyse und Interpretation von Datensätzen sind "Brot und Butter" der Wissenschaft und Standardthema des epistemologischen Diskurses. In den letzten Jahren wird in diesem Zusammenhang immer wieder das Schlagwort „Big Data“ gebraucht, oftmals verbunden mit Verheißungen eines revolutionären wissenschaftlichen Fortschritts. Auch und gerade in der raumbezogenen Forschung werden rund um Big Data neue Perspektiven diskutiert, etwa wie räumlich referenzierte Variablen zur Identifikation von Individuen genutzt werden können und somit den Datenschutz gefährden. Oder inwiefern und in welchem Ausmaß räumliche Disparitäten durch die Anwendung von Big Data reproduziert werden. Die kritische Auseinandersetzung mit Big Data wird angesichts der steigenden Verfügbarkeit von Big Data in den kommenden Jahren eine weiter zunehmende Rolle spielen. Neben praktischen und methodischen Problemen bezüglich der Analyse und Interpretation solcher volatiler Datensätze sehen wir beim wissenschaftlichen Umgang mit Big Data zahlreiche konzeptionelle Fragen und ethische Herausforderungen.

Aus raumwissenschaftlicher Perspektive ist kritisch zu analysieren, welche Raumkonzepte verschiedensten Anwendungen von Big Data zugrunde liegen und welche Vorstellungen somit reproduziert werden. Zudem steht die Frage im Raum, für welche Forschungsthemen die Anwendung von Big Data überhaupt in Frage kommt. Für quantitative Fragestellungen scheinen die Vorteile auf der Hand zu liegen: Georeferenzierte Datensätze versprechen durch ihr Volumen, die große Anzahl unterschiedlicher Quellen und ihre Erfassungsgeschwindigkeit eine neue Perspektive auf menschliche Verhaltensweisen und naturwissenschaftliche Phänomene. Doch auch qualitativ Forschende beginnen, Daten aus Bereichen von Big Data und Social Media in ihre Studien einzubeziehen.

Wie und wo machen Big Data in unseren Forschungen Sinn? Welche Erfahrungen gibt es bisher zu den methodischen und methodologischen Fragen, vor welche uns der Umgang mit Big Data stellt? Welche Konsequenzen ergeben sich aus dem Einbezug solcher Daten für unsere wissenschaftstheoretischen Grundlagen? Wie können wir als raumbezogen Forschende einen reflektierten Umgang mit Big Data stärken? Diese Fragen wollen wir in der IfL Forschungswerkstatt #4 in den Mittelpunkt stellen.

Ablauf Montag 26. März 2018



SynPodium & Diskussionsforen

10:00 // **Auftakt** am IFL

Begrüßung durch das Organisationsteam *Kristine Beurskens, Tim Leibert, Thomas Tillmann (Leibniz-Institut für Länderkunde)*

10:30 // **SynPodium:** „Fluch oder Segen – Big Data in der raumwissenschaftlichen Forschung“

mit *Dirk Burghardt (Technische Universität Dresden), Henning Füller (Humboldt-Universität Berlin), Michael Klemm (Universität Koblenz-Landau), Jörg Scheffer (Universität Passau), Till Straube (Goethe-Universität Frankfurt)*

moderiert von *Tim Leibert (Leibniz-Institut für Länderkunde)*

12:30 // **Mittagspause** mit Imbiss

13:30 // Anmoderation und **Kurzvorstellung der Diskussionsforen**

D1: Vom Sehen zum Verstehen – Big Data Visualisierungsmöglichkeiten

Dirk Burghardt (Technische Universität Dresden)

D2: Wie passt Big Data in die raumbezogene Forschung? Kritische Perspektiven

Henning Füller (Humboldt-Universität Berlin)

D3: Big Data, Qualitative Forschung und Forschungsethik

Michael Klemm (Universität Koblenz-Landau)

D4: Raum und Raumtransformation im Zeitalter der Digitalisierung: Geographische Perspektiven auf Big Data

Jörg Scheffer (Universität Passau)

D5: Wie Big Data Verräumlichungen reproduzieren

Till Straube (Goethe-Universität Frankfurt)

13:45 // **Diskussionsforen Runde 1**

15:00 // Kaffeepause

15:30 // **Diskussionsforen Runde 2**

17:00 // **Reflexionen und Ankündigungen**

19:00 // Gemeinsames Abendessen (optional) im Volkshaus Leipzig

Ablauf Dienstag, 27. März 2018



ExperiSpaces

9.00 // ExperiSpaces

E1: Datenstadt – Workshop zu Prozessen algorithmischer Entscheidungsfindung im datafizierten Raum

Damian Paderta (Paderta.com)

E2: Tag Maps - Neue Methoden zur Visualisierung und Auswertung von kollektiver Wahrnehmung in Planung, auf Basis Social Media gewonnener Big Data

Alexander Dunkel (TU Dresden) & *Dirk Burghardt* (TU Dresden)

E3: Gamifizierte Ansätze zur Gewinnung von Big Data: Wie Liebe, Matching und Tinder Einzug in die Stadtforschung erhalten – Eine kritische und interaktive Auseinandersetzung am Beispiel von MatchMyCity.org

Laura Ziegler (Bauhaus-Universität Weimar) & *Max A. Irmer* (Bauhaus-Universität Weimar)

E4: Mehr als Wörter zählen – Quantitative Verfahren der Textanalyse

Thilo Wiertz (Universität Freiburg) & *Tobias Schopper* (Universität Freiburg)

16:00 // Zusammenkunft & Kaffeepause/Networking/Diskussion im Posterrundlauf

17:00 // **Abschluss und Verabschiedung**

„Fluch oder Segen – Big Data in der raumwissenschaftlichen Forschung“

Das *SynPodium*, eine interaktive Podiumsdiskussion zu Erfahrungen im Umgang mit Big Data in der derzeitigen raumbezogenen Forschung, wird in die Veranstaltung einführen. Im ersten Teil des SynPodiums eröffnen unsere Podiumsgäste uns ihre Perspektiven auf Big Data aus ihrer jeweiligen Forschungspraxis heraus. Schwerpunkt der weiteren Diskussion sollen unter anderem die damit verbundenen methodischen und ethischen Herausforderungen sein.

Anschließend wird Raum für Nachfragen und Kommentare der Zuhörer gegeben und die Diskussion für das gesamte Plenum geöffnet. Das *SynPodium* bildet den ersten Einstieg in das Themenfeld der *IfL Forschungswerkstatt #4* und dient als Überblick und Grundlage für die vertiefenden Diskussionen im weiteren Verlauf der Veranstaltung.

Unsere Podiumsgäste im Kurzporträt:

Dirk Burghardt ist Professor für Kartographische Kommunikation an der Technischen Universität Dresden. Seit einem Physik-Studium und einer Promotion im Bereich Kartographie/Geodäsie ist er bereits längere Zeit im Bereich der Produktion von kartographischen Darstellungen tätig. Seine gegenwärtigen Forschungsinteressen richten sich auf die neuen Herausforderungen der Abstimmungen zwischen visuellen Darstellungsmethoden und Nutzungskontexten, insbesondere im Bereich ‚citizen science‘.

Henning Füller ist wissenschaftlicher Mitarbeiter am Geographischen Institut der Humboldt-Universität zu Berlin. Durch sein Studium vereint er eine sozial- sowie politikwissenschaftliche und geographische Perspektive auf das Verhältnis von Macht und Raum. Aktuelle Forschungsfelder sind der Wandel des Städtischen und Biosicherheit. In beiden Feldern untersucht er den Einsatz von Big Data und dessen Auswirkungen aus einer kritischen raumwissenschaftlichen Perspektive.

Michael Klemm ist Professor für Medienwissenschaft mit sprachwissenschaftlicher Ausrichtung am Institut für Kulturwissenschaft der Universität Koblenz-Landau, Campus Koblenz. Schwerpunkte seiner Forschung sind unter anderem die politische Online-Kommunikation (z.B. Audiovisuelle Politik, Twitter), die kulturvergleichende Medienforschung und multimodale Diskursanalysen (u.a. zu Erinnerungsdiskursen, Wissenschaftspopularisierung). Seine Arbeit im Schnittfeld von Medien- und Kulturwissenschaften mit Bezug auf die politische Nutzung von Social Media erfordert eine Reflexion des Umgangs mit Big Data, insbesondere für die qualitative Forschung und in Hinsicht auf forschungsethische Fragestellungen.

Jörg Scheffer ist Akademischer Oberrat an der Universität Passau in der Abteilung Geoinformatik und Geographie. Neben Schwerpunkten in den Bereichen Kulturtheorie, Stadtgeographie und der Interkulturellen Kommunikation forscht und lehrt Jörg Scheffer im Bereich der Geoinformatik und Neue Medien. Seine laufende Habilitation "Sozialisation im Kontext von Big Data" setzt sich mit dem Einfluss der Digitalisierung auf die ungleichen Aufstiegschancen im Stadtraum auseinander.

Till Straube ist wissenschaftlicher Mitarbeiter am Institut für Humangeographie an der Goethe-Universität Frankfurt am Main. Nach abgeschlossener Ausbildung und einem Studium zum Diplom-Betriebswirt (BA) gründete er zunächst eine Softwarefirma und war als Programmierer und Projektmanager tätig. In seinem aktuellen Promotionsprojekt untersucht er polizeiliche Prognoseprogramme und deren Evaluationspraktiken.

Diskussionsforen

Am Nachmittag des ersten Tages wird es die Möglichkeit geben, bestimmte Themen und Fragestellungen rund um den Gebrauch von Big Data in der raumbezogenen Forschung mit den Diskutanten des SynPodiums weiter zu vertiefen. Dazu werden wir uns in kleinere Runden aufteilen, in der spontane Nachfragen wie auch Beispiele und Problemstellungen aus der eigenen Forschungsarbeit intensiv besprochen werden können. Die Diskussionsforen finden in 2 Runden statt, geteilt durch eine Kaffeepause. Ein Wechsel in der Pause ist möglich, aber nicht zwingend. Für die Planung ist eine grobe Abschätzung der Teilnehmendenzahlen hilfreich, daher bitten wir in der Anmeldung um Interessensbekundungen für die Wahl der Diskussionsforen.

ExperiSpaces

Die Workshops am 2. Tag finden parallel statt. Folgende stehen zur Auswahl:

E1: Datenstadt – Workshop zu Prozessen algorithmischer Entscheidungsfindung im datafizierten Raum

Damian Paderta (Offene Kommunen.NRW)

„Big Data“ und die daraus resultierenden Prozesse algorithmischer Entscheidungsfindung gehören längst zum täglichen Leben. Die gegenwärtige ethische Debatte um die Folgen von Big Data und Automatisierung fokussiert auf individualrechtliche Aspekte. Jedoch weisen algorithmische Prozesse, die die Automatisierung vor allem ausmachen, primär eine kollektive Dimension auf. Personenbezogene Daten sind nicht per se private Daten. Das Konzept der informationellen Privatheit reduziert die Vorstellung von Privatheit auf die reine Kontrolle von personenbezogenen Informationen und übersieht, dass eine Kontrolle der personenbezogenen Daten möglich ist, ohne damit Privatheit zu genießen. Statt dem *ob* steht vielmehr das *wie, warum* und *wer* im Vordergrund datenbezogener Praxis. Im Workshop werden daher Ambivalenzen gängiger Praxis im Bereich Big Data diskutiert und individuelle ethisch-moralische Schwellenwerte reflektiert. Ein Schwerpunkt bildet die kritische Auseinandersetzung mit Big Data Narrativen.

Technologien die Entscheidungen besser fällen sollen als Menschen sind keine Errungenschaft der Neuzeit. Doch die schiere Anzahl der beteiligten Systeme, ihre Komplexität und ihre globalen

Auswirkungen haben ein neues unvergleichliches Ausmaß angenommen. Weder sind die Prozesse algorithmischer Entscheidungsfindung neutral, noch sind sie von denjenigen, die von ihnen betroffen sind, in ihrer Kausalität erfahrbar. Das Versprechen mithilfe von Big Data zu neuen Erkenntnissen zu gelangen, erweist sich somit in einer kleinen aber entscheidenden Differenzierung als zutreffend (wenn auch nicht unumstritten): es erlaubt die Steigerung privatisierter Erkenntnisse einiger Weniger - in weitestgehender Unkenntnis der Konsequenzen Vieler.

Dazu ist das Verständnis eines datifizierten Raums notwendig. Weniger als explizierendes Konzept, sondern als Perspektive und Mindset, um das Potenzial der Wechselwirkungen und Implikationen zwischen der Trias „Technologie-Raum-Gesellschaft“ bewusst zu machen. Der datifizierte Raum kann neue Perspektiven auf den physischen Raum, sowie auf die Genese neuer sozialer Praktiken eröffnen. Datifizierte Räume können zur Repression genutzt werden oder zum nachhaltigeren Wirtschaften mit Ressourcen. Sie können Mythen zerstören und neue schaffen, Wissenshierarchien verflachen und an anderer Stelle vertikalisieren. Diese datifizierten Räume bilden die Grundlage von allen Konzepten, die sich unter dem Begriff „Datenstadt“ vereinen lassen. Damit verwebt, sind viele Vorstellungen von Steuerung, Verwaltung und Führung von Gesellschaft, die sich ideologisch in der Technik und damit im Raum, abbilden.

Ablauf Workshop:

1. Skizzieren der persönlichen Alltagsentscheidungen in Form von Storyboards & Analyse derer im Hinblick auf zukünftige Prozesse der automatisierten Datenerfassung und Entscheidungsalgorithmen.
2. Ausgesuchte Datensätze werden nach den Statuten *Permissiv* (offen & teilbar), *Selektiv* (reduziert & kontrolliert) und *Restriktiv* (geheim) verortet und diskutiert. Dabei entsteht eine kollaborative Karte der ausgehandelten Zugänglichmachung von Daten einer fiktiven Stadt.
3. Planspiel zu kreativen Anwendungsentwicklungen anhand von randomisierten Datensätzen.

Literaturempfehlungen:

Mayer-Schönberger, V. & Cukier, K. (2013): *Big Data. Die Revolution, die unser Leben verändern wird*. Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz (2015) 58: 788. S. 7ff.

Morozov, E. & Bria, F. (2017): *Die smarte Stadt neu denken. Wie urbane Technologien demokratisiert werden können*. Berlin, Rosa-Luxemburg-Stiftung. S.66ff

Calzada I. (2018): *From Smart Cities to Experimental Cities?* In: Giorgino V., Walsh Z. (Eds.) *Co-Designing Economies in Transition*. Cham, Palgrave Macmillan. S.192ff

Finn, E. (2017): *What algorithms want: Imagination in the age of computing*. Cambridge, MA and London, UK: The MIT Press.

E2: Tag Maps - Neue Methoden zur Visualisierung und Auswertung von kollektiver Wahrnehmung in Planung, auf Basis Social Media gewonnener Big Data

Alexander Dunkel (Technische Universität Dresden) und *Dirk Burghardt* (Technische Universität Dresden)

Als Wahrnehmung wird der Bewusstseinsprozess des subjektiven Verstehens der Umwelt bezeichnet. Grundlage für diesen Prozess ist die Gewinnung von Informationen über die Sinne, also aus visuellen, olfaktorischen, akustischen und anderen Reizen. Die Wahrnehmung ist aber auch wesentlich durch interne Prozesse beeinflusst. Das menschliche Gehirn ist fortlaufend damit beschäftigt, sowohl bewusst als auch unbewusst Sinneswahrnehmungen mit Erinnerungen abzugleichen, zu vereinfachen, zu assoziieren, vorherzusagen oder zu vergleichen. Aus diesem Grund ist es schwierig, die Wahrnehmung von vielen Menschen zu erfassen und in Planungsprozessen zu berücksichtigen. Jedoch wird genau dies von der Europäischen Landschaftskonvention gefordert, die Landschaft als einen bestimmten Bereich definiert, so wie er von Besuchern und Einwohnern wahrgenommen wird.

Während viele Fortschritte und Erkenntnisse, zum Beispiel aus den Kognitionswissenschaften, heute helfen, die Wahrnehmung einzelner Menschen zu verstehen, konnte die Stadt- und Landschaftsplanung kaum profitieren. Es fehlt an Kenntnissen über das Zusammenwirken der Wahrnehmung vieler Menschen. Schon Stadtplaner Kevin Lynch beschäftigte dieses gemeinsame, kollektive ‚Bild‘ der menschlichen Umwelt.

Der vorgeschlagene Workshop zu „Tag Maps“ ist auf diese Schnittstelle zwischen Data Science und angewandter Forschung ausgerichtet und soll ausdrücklich auch anwendungsbezogene Teilnehmer wie Stadt- und Landschaftsplaner ansprechen. Eine bisher in der Planung ungenutzte Informationsquelle für die Erfassung der Wahrnehmung vieler Menschen bietet sich in Form von crowdsourced Daten (auch ‚Big Data‘), also großen Mengen an Daten die von vielen Menschen im Internet zusammengetragen werden. Im Vergleich zu konventionellen Daten, zum Beispiel solchen die durch Experten erhoben werden und durch öffentliche Träger zur Verfügung stehen, eröffnet sich durch crowdsourced Daten eine bisher nicht verfügbare Quelle für Informationen, um die komplexen Zusammenhänge zwischen Raum, Identität und subjektiver Wahrnehmung zu verstehen. Anhand dieser Daten wird die Entwicklung neuer Visualisierungstechniken exemplarisch vorgestellt.

Im Workshop werden Teilnehmende angeleitet, „Tag Maps“ für frei wählbare Bereiche selbstständig zu erstellen. Die Methodik wurde im Rahmen einer Dissertation erarbeitet und in einer Reihe von Workshops weiterentwickelt. Der Workshop ist so organisiert, dass Teilnehmende mit unterschiedlich tiefem Hintergrundwissen teilnehmen können. Für schnelle Anwender besteht die Möglichkeit, sich selbstständig anhand der zur Verfügung gestellten Materialien durch den Workshop zu arbeiten. Gleichzeitig wird Schritt für Schritt der Ablauf anhand eines Beispieldatensets vorgestellt (Berlin Spreeinsel). Teilnehmende, die einige Schritte verpassen, können zu einem späteren Zeitpunkt wieder einsteigen, da alle Zwischenschritte der Datenverarbeitung zur Verfügung gestellt werden. Ein hoher Stellenwert liegt auf dem Vermitteln eines breiten Überblicks der Hürden, die bei der Arbeit mit tausenden, von vielen Menschen erstellten und im Internet geteilten Fotos auftreten. Ziel ist es, Anwendern die Grundkenntnisse zu vermitteln um diese Daten selbstständig in Forschung und Praxis einzusetzen. Gegen Ende des

Workshops besteht die Möglichkeit, weitere Visualisierungsmethoden und -richtungen kurz vorzustellen, wie die Generierung von Sichtachsen aus Fotodaten.

Ablauf Workshop:

1. Vorstellung der Methodik
2. Datenauswahl und exploratives filtern von Informationen basierend auf 4 Facetten: Raum (Space), Zeit (Temporal), „Social“ (User Origin) und „Thematic“ (Tags)
3. Säubern der Daten sowie manuelle und automatische Vorselektion für thematisches Clustern
4. Aggregieren der Daten über Clustern in zwei Layern: Foto Standort und Foto Tags
5. Visualisieren der Tag Maps in ESRI ArcGIS mit abschließender Interpretation der Ergebnisse

Literaturempfehlungen:

Dunkel, A. (2015). *Visualizing the perceived environment using crowdsourced photo geodata*. *Landscape and Urban Planning* 142. S. 173-186.

Lynch, K. (1960). *The image of the city*. Cambridge, Mass.: Technology Press.

E3: Gamifizierte Ansätze zur Gewinnung von Big Data: Wie Liebe, Matching und Tinder Einzug in die Stadtforschung erhalten – Eine kritische und interaktive Auseinandersetzung am Beispiel von MatchMyCity.org

Laura Ziegler (Bauhaus-Universität Weimar) und *Max A. Irmer* (Bauhaus-Universität Weimar)

Tinder, Dating oder Matching sind Begriffe, die – zumindest auf den ersten Blick – nichts mit der Stadtforschung zu tun haben. Dennoch tauchen aktuell vermehrt eben solche Formate auch in raumwissenschaftlichen Forschungen auf. Vom spielerischen Test um herauszufinden, ob man eher zu Berlin oder zu München passt über die Beantwortung eines Fragebogens, der einem direkt die am besten passendste Stadt ausgibt bis hin zur interaktiven Einbindung der BewohnerInnen und BesucherInnen in Santa Monica, die verschiedene Faktoren des städtischen Lebens durch swipen bewerten können. Dabei sind die Ziele, wenn es sich um einen wissenschaftlichen Zugang handelt, häufig dieselben: Möglichst viele junge Menschen (Stichwort Generation Y und Millennials) erreichen. Der Zugang erfolgt über den Transfer von medial präsenten und in der Alltagskultur etablierten Funktionsweisen und Formaten wie Tinder in andere Kontexte wie z.B. die Stadtforschung. Dabei variieren sowohl der (methodologische) Erfolg als auch die Anzahl der Teilnehmenden je Projekt stark.

Im Rahmen unseres Workshops stellen wir uns folglich den Fragen, die solche Funktionsübernahmen aus anderen Themengebieten und deren Forschungsintegrationen in raumbezogene Wissenschaften mit sich bringen: Wer kann damit erreicht werden? Können die Ergebnisse solcher Formate repräsentativ sein? Dürfen wissenschaftliche Standards eingebüßt

werden, um die Anzahl der Teilnehmenden zu erhöhen? Wie viel Integration von spielerischen Elementen können wissenschaftliche Methoden – insbesondere Umfragen – vertragen? Müssen und sollen die Befragten darüber aufgeklärt werden, dass ihre Angabe für wissenschaftliche Zwecke genutzt werden? In welcher Form und welchen raumbezogenen Wissenschaften sind diese Big-Data-Ansätze zielführend?

Um sich diesen komplexen und disziplinübergreifenden Fragen zu nähern, untersuchen wir die Webseite MatchMyCity.org als Anwendungsbeispiel. Die „erste Datingplattform für Städte“ erreichte in zwei Monaten über 20.000 Teilnehmende und generierte damit in Bezug auf den urbanen Kontext einen für Deutschland neuartigen Datensatz. Auf der Webseite können in Form einer spielerischen Umfrage 33 urbane Merkmale (z.B. Anbindung an einen internationalen Flughafen oder Vielfältigkeit des kulturellen Angebots) auf einer Wichtigkeitsskala von 0 bis 3 bewertet werden. Ein Algorithmus errechnet anschließend die Stadt, die am besten zum/zur Teilnehmenden passt. Die anonymisierten Ergebnisse aus der Umfrage wurden schon im Rahmen einer explorativen Forschungsarbeit verwendet und ausgewertet. Konsequenterweise kann durch die Untersuchung und Analyse von MatchMyCity.org der gesamte Prozess eines gamifizierten Big-Data-Ansatzes untersucht, debattiert und reflektiert werden. Die Annäherung an den Forschungsansatz erfolgt über ein interaktives Format und die Präsentation der Forschungsergebnisse. Aus diesem Einstieg werden die Herausforderungen und Probleme des Big-Data-Ansatzes im Detail entwickelt und anschließend (in Gruppen) analysiert und diskutiert.

Ablauf Workshop:

1. Einführung: Städte-Speed-Dating & Testen der Umfrage auf MatchMyCity.org
2. Ergebnisse: Vorstellung der Forschungsergebnisse hinter MatchMyCity.org & Aufzeigen von Herausforderungen und Problemen mit gamifizierten Big-Data-Ansätzen (insbesondere hinsichtlich methodischer und methodologischer Aspekte)
3. Analyse und Diskussion folgender Themenfelder:
 - a. Gamifizierte Ansätze zur Datengewinnung: Interdisziplinäre Perspektive & Reichweite
 - b. Methodische und methodologische Herausforderungen: spielerische Elemente im wissenschaftlichen Kontext
 - c. Ethische Fragen hinsichtlich Big-Data-Anwendungen: Informed Consent & Repräsentativität von gamifizierten Ansätzen in der Forschung
 - d. Zukunftsperspektiven: Mehrwert von gamifizierten Ansätzen für die Stadtforschung- & -planung sowie Big Data vs. Smart Data

Literaturempfehlungen:

Vereinigung für Stadt-, Regional- und Landesplanung (Hrsg.) (2017): *Play the City – Stadt & Spiel*. PlanerIn 3/2017, Berlin.

Fuchs, M., Fitze, S., Ruffino, P. & Schrade, N. (Hrsg.) (2014): *Rethinking Gamification*. Lüneburg: Meson press.

Eckardt, F. (Hrsg.) (2014): *Stadtforschung - Gegenstand und Methoden*. Wiesbaden: Springer VS.

E4: Mehr als Wörter zählen – Quantitative Verfahren der Textanalyse

Thilo Wiertz (Universität Freiburg) und *Tobias Schopper* (Universität Freiburg)

Im Zuge der Digitalisierung werden zunehmend umfangreiche Textquellen digital zugänglich. Webscraping, Textmining, Natural Language Processing oder Methoden der Korpuslinguistik ermöglichen es, immer umfangreichere Datenbanken (digitaler) Textquellen zu erstellen und quantitative Analysen durchzuführen. Dabei eröffnen sich neue analytische Möglichkeiten, auch für ansonsten qualitativ ausgerichtete humangeographische Forschungsprojekte, wie Erfahrungen aus der humangeographischen Diskursforschung in den vergangenen 10 Jahren gezeigt haben. Jüngere methodische und technische Entwicklungen sowie Veränderungen in der sozial-medialen Textproduktion werfen jedoch neue Fragen auf, die auch die theoretischen Zugänge betreffen. So lassen sich Netzwerk-Formationen und affektive Ballungen in sozialen Medien wie Twitter oder Facebook mit diskurstheoretischen Ansätzen und korpuslinguistischen Methoden nur eingeschränkt erfassen. Umgekehrt sind theoretische Kontextualisierungen jenseits der (poststrukturalistischen) Diskurstheorie denkbar und „kollaborative“ Textkorpora versprechen einen zunehmend einfacheren Zugriff auf sozialwissenschaftliche relevante Quellen.

Die Möglichkeiten quantitativer Textanalysen in Sozialwissenschaften und Humangeographie sind keinesfalls erschöpft und Diskussionen um Big Data und Digitale Geographien verweisen auf ein steigendes Interesse und eine Erweiterung der Anwendungsfelder. Um einen kritischreflektierten Umgang zu ermöglichen, ist dabei ein substantielles Verständnis der technischen Abläufe, methodischen Vorannahmen und analytischen Verfahren erforderlich. Im Rahmen dieses Workshops möchten wir gemeinsam mit den TeilnehmerInnen einige Möglichkeiten der quantitativen Textanalyse kennenlernen und ausprobieren. Über eine Weboberfläche erschließen wir uns korpuslinguistische Verfahren wie Häufigkeits-, Kollokations- und Keywordanalysen, die wir in kleinen Projekten auf bestehende Medien und Politikorpora anwenden. Zudem befassen wir uns mit einigen Möglichkeiten der Analyse von sozialen Medien und Netzwerken mit R. Die im Workshop gesammelten Erfahrungen diskutieren wir vor dem Hintergrund unterschiedlicher Forschungsinteressen und theoretischer Zugänge.

Ablauf Workshop:

1. Vorstellung & Begrüßung
2. Einführung: Ansätze der quantitativen Textverarbeitung – Erfahrungen und aktuelle Entwicklungen in der (geographischen) Diskursforschung, Linguistik und Computational Sciences
3. Interaktiv: Ideen für die Verwendung quantitativer Textanalysen in eigenen/anderen Forschungsfeldern entwerfen
4. Praxisworkshop: Korpuslinguistische Analysen mit CQPWeb (Material: Deutschsprachige Medienkorpora, Bundestagsdebatten, etc.)
5. Präsentation: Auswertung von Social Media-Daten mit R
6. Projektarbeit: Entwickeln kleiner Projekte und Durchführung mit CQPWeb
7. (Eventuell Präsentation zu Verfahren wie Kollokationsnetzwerken, Ähnlichkeitsanalysen, Wort-Vektoren, Sentiment Analysis, Named Entity Recognition)
8. Diskussion: Ergebnisse aus Projekten, kritische Reflexion und theoretische Kontextualisierung

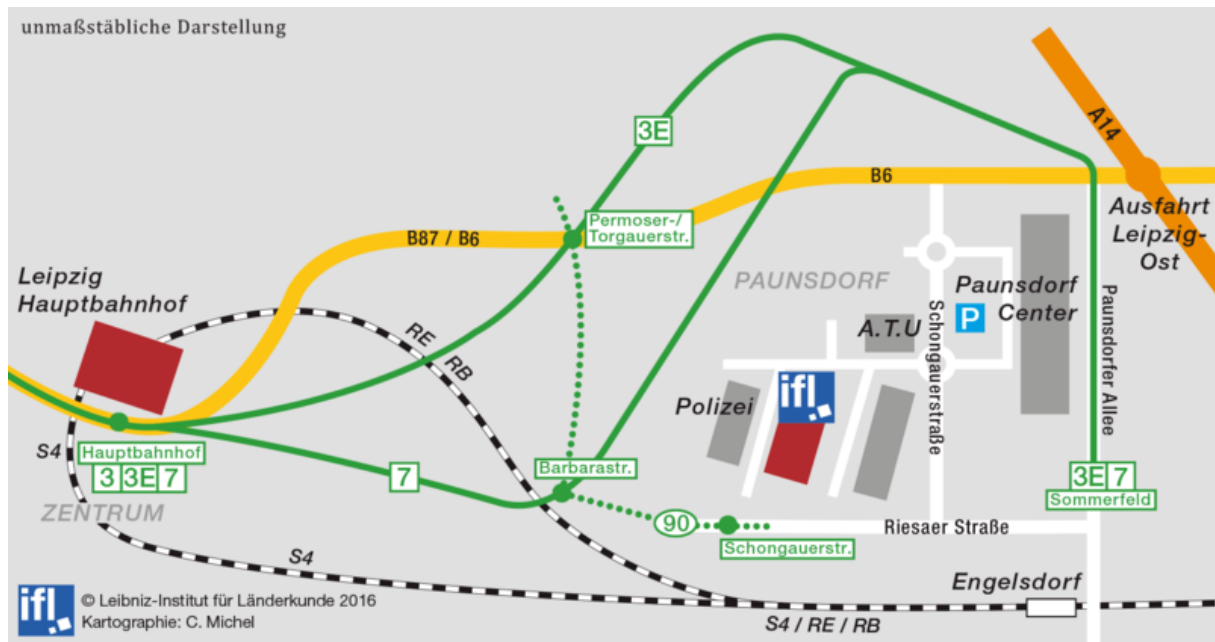
Literaturempfehlungen:

Linnemann, K. (2014): *Im Feld Der Worte: Eine Reflexion Quantitativ-Lexikometrischer Methoden Der Diskursanalyse*. Berichte. Geographie Und Landeskunde 88 (1): S. 5–20.

Mattisek, Annika, Carmella Pfaffenbach, and Paul Reuber (2013): *Methoden der empirischen Humangeographie*. 2. Aufl., Neubearb. Das Geographische Seminar. Braunschweig: Westermann. Kapitel 5.3.1 Makroverfahren der Text- und Korpusanalyse, S. 272-282.

Wiertz, Thilo (im Erscheinen). *Quantitative text analysis in Geography: facilitating access and collaboration*. DIE ERDE. (Entwurfssfassung).

Anfahrt Leibniz-Institut für Länderkunde



Tram & Bus

1. Variante (Dauer: 26 min. Umst.: 0)

Tram 3E

Richtung: Sommerfeld (Gleis 2)

Leipzig Hauptbahnhof → Sommerfeld

2. Variante (Dauer: 27 min. Umst.: 0)

Tram 7

Richtung: Sommerfeld (Gleis 4)

Leipzig Hauptbahnhof → Sommerfeld

3. Variante (Dauer: 27 min. Umst.: 1)

Tram 7

Richtung: Sommerfeld (Gleis 4)

Leipzig Hauptbahnhof → Barbarastraße

Bus 90

Richtung: Sommerfeld/Paunsdorf Center

Barbarastraße → Schongauerstraße

4. Variante (Dauer: 36 min. Umst.: 0)

Bus 72

Richtung: Paunsdorf Straßenbahnhof (Gleis 2)

Leipzig Hauptbahnhof → Schongauerstraße

SBahn (Dauer: 19 min. Umst.: 0)

S4

Richtung: Wurzen/Oschatz

Leipzig Hauptbahnhof (tief) → Leipzig Engelsdorf

RE/RB (Dauer: 6 min. Umst.: 0)

RE/RB

Richtung: Wurzen, Döbeln, Meißen, Dresden

Leipzig Hauptbahnhof → Leipzig Engelsdorf

PKW

A14, Abfahrt Leipzig Ost

Notfall Kontakte

Kristine Beurskens: 01577 / 5331275

Tim Leibert: 0160 / 8368206

Thomas Tillmann: 0157 / 39252498