

Die soziale Genese von „Big Data in Social Sciences“ in der Entstehungsphase zentraler Beobachtungs- und Analyse-Plattformen ab 2000

Hans Dieter Hellige

Zusammenfassung

Der Beitrag rekonstruiert am Beispiel der Genese der Computational Social Science den Wandel der sozialen „eScience“-Architekturen und -Wissenschaftskulturen unter dem Einfluss der Rezentralisierung des Computing und der Herausbildung zentraler Beobachtungs- und Analyse-Plattformen nach 2000. Er zeigt, wie sehr die Agenda der soziometrischen Verhaltensforschung mit ihrem Ziel einer durchgängigen Erfassung des Alltagsverhaltens von Gruppen und sozialen Netzwerken von dem durch die militärische Förderung nach 9/11 und der Platform Economy ausgelösten Data-Analytics-Boom geprägt wurde. Entstehungsbedingt und gegenstandsspezifisch entfalten Big-Data-Konzepte in den einzelnen Wissenschaften eine unterschiedliche Wirkmächtigkeit, sie sollten deshalb wie auch wegen der großen Datendiversität nicht unter ein einheitliches quantitativsbetontes normatives „Fourth Paradigm“ subsumiert werden, das die epistemischen Divergenzen zwischen Sozial- und Naturwissenschaften von vornherein einebnet.¹

1 Einleitung

Aus technik- und wissenschaftshistorischer Sicht lassen sich in der Genese von Big-Data-Wissenschaftskulturen und von Modellen der „data governance“ deutlich zwei Etappen unterscheiden, die wesentlich auf dem Wandel der sozialen Architekturen im Computing und in den informationellen Infrastrukturen beruhen. Die ersten Big-Data-Pionierbereiche waren mit ihren charakteristischen kollaborativen Wissenschaftsplattformen und föderativen Grid-Architekturen noch stark beeinflusst von der Endphase der in den 1960er/1970er Jahren entstandenen Ära des „Distributed Computing“. In ihr hatte sich der entscheidende Wandel des Architekturkonzepts der computerbasierten Wissensorganisation vom zentralistischen Versorgungsmodell der Mainframe-Ära und vom „public utility“-Konzept der frühen Time-Sharing-Ära zum PC-basierten dezentralen Kommunikationsmodell einer sich selbst organisierenden kooperativen Wissensproduktion vollzogen. Die vom Leitkonzept zentraler Beobachtungs- und Analyse-Plattformen und Cloud-Systemen geprägte zweite Etappe der sozialen Architektur und epistemischen Kultur der „data-rich eSciences“ war das Ergebnis der in den 1990er Jahren einsetzenden Rezentralisierung im Computing und der Rückverlagerung eines Großteils der

¹ Die Beiträge „Die soziale Genese kollaborativer e-Science-Plattformen ab 1990 am Beispiel von Big Data in Astronomy“ und „Die soziale Genese von Big Data in Social Sciences in der Entstehungsphase zentraler Beobachtungs- und Analyse-Plattformen ab 2002“ wurden 2019 als Gesamttext unter dem Titel „Etappen der sozialen Genese von Big Data und ihr prägender Einfluss auf Wissenschaftskonzepte und Forschungsperspektiven“, (56 S.) abgegeben. Die Erstfassung wurde in der jetzigen Fassung von 2021 überarbeitet, erweitert und in zwei Beiträge aufgeteilt. Diese erscheinen 2023 in: Hashagen, Ulf; Seising, Rudolf (Hrsg.): Algorithmische Wissenskulturen: Der Einfluss des Computers auf die Wissenschaftsentwicklung. Wiesbaden: Springer 2023

Datenverarbeitung und -speicherung von den dezentralen Endsystemen in zentrale Serverfarmen. Treiber der Entwicklung waren die großen Hardware-, Software- und Content-Anbieter, die unter der Devise des Endes des PC-Zeitalters einen grundlegenden Wandel in der Governance und in der Arbeitsteilung zwischen verschiedenen Computing-Ressourcen anstrebten, um den durch die PC-Revolution verlorenen Einfluss auf die Computernutzung wieder zurückzugewinnen. Ein Vorbild hierfür bildete die Zukunftsvision der Designer der ersten Version der TCP/IP-Protokolle Robert Kahn (*1938) und Vinton G. Cerf (*1943) für den Aufbau einer „Digital Library System“ genannten „National Information Infrastructure“, die wieder an die „Big Library“-Metapher und die soziale Systemarchitektur klassischer Infrastrukturversorgungsnetze anknüpfte (Kahn, Cerf 1988; Hellige 2012, S. 24–27). Inspiriert hiervon entwickelten die IT-Firmen Oracle, IBM, Sun Microsystems, Hewlett-Packard und Microsoft zentral organisierte Systemarchitekturen und Geschäftsmodelle für das „Post-PC-Era Paradigm“, die sie unter den Leitmetaphern „Service Grid“, „On Demand Computing“, „Utility Computing“ und 1996 erstmals auch „Cloud Computing“ medienwirksam inszenierten (Hellige 2012, S. 33–51). Die IT-Giganten schufen damit zwar wesentliche Ideenkonzepte und Technologien für die Renaissance der Datacenter und die Entstehung von Service-Plattformen, doch wegen der Fixierung auf ihre angestammten Geschäftsfelder übernahmen nicht sie die Führung bei der Web-Zentralisierung und beim Aufbau von Big-Data-Infrastrukturen, sondern überließen dies zwei anderen Akteursgruppen, staatlich-militärischen Instanzen und den großen Internet-Service-Providern. Diese waren es auch, die das zentrale Datenbank- und Plattform-Konzept und das zentral organisierte Modell der „data governance“ (Micheli et al. 2020) mit den Technologien zur automatischen Datenanalyse, Webfilterung und Wissensextraktion zusammenbrachten, die sich in der „customer-centric economy“ bereits seit den 1980er Jahren für die Personen- bzw. Kundenbeobachtung in der Entwicklung befanden. Welche wesentlichen Anstöße die Big-Data-Kulturen und insbesondere die Computational Social Science durch diese beiden Akteursgruppen erhielt, soll im folgenden Kapitel skizziert werden.