



Torsten Eymann
Universität Bayreuth



Michael Goedicke
Universität Duisburg-Essen



Ulrike Lucke
Universität Potsdam

E-Science als Herausforderung für die Informatik

Dieses Heft des Informatik Spektrum beschäftigt sich mit dem Thema E-Science. Von diesem Begriff gibt es vermutlich ebenso viele Schreibweisen wie Definitionen. E-Science beschreibt in unseren Augen in einer weit gefassten Definition die Nutzung von Informatikmethoden und Technologien für Zwecke der wissenschaftlichen Erkenntnisgewinnung. Dies schließt High-Performance Computing und Datenanalyseverfahren ebenso ein wie Forschungsdatenmanagement. In der Definition von Bohle, die wir im Positionspapier der GI nutzen, wird E-Science beschrieben als „die Anwendung von Informations- und Kommunikationstechnologie (ICT) in der modernen Wissenschaft und umfasst ein Spektrum von der Vorbereitung und Durchführung von Untersuchungen, über die Datenerfassung und Verbreitung von Ergebnissen bis hin zur Langzeitspeicherung und dem Zugang zu relevanten Materialien und Artefakten“¹. Ähnliche und andere Definitionen werden in den Einzelbeiträgen nochmals ausgeführt.

Man könnte also einwenden, dass es „nur“ um die Anwendung dieser Verfahren, z. B. in den Naturwissenschaften oder Geisteswissenschaften (Digital Humanities), ginge. Dieser Wahrnehmung folgend, waren in den letzten Jahren die wesentlichen Treiber von E-Science eher die Anwender, z. B. die Geowissenschaften oder die Physik, oder die Betreiberinnen und Betreiber von E-Research-Infrastrukturen, häufig die Bibliotheken von Hochschulen und Akademien.

In unseren Augen trifft das nicht die Essenz des Begriffs, denn E-Science beschreibt eine Vision des zukünftigen wissenschaftlichen Arbeitens und die mittel- bis langfristige Entwicklung der wissenschaftlichen Praxis dorthin. Wir möchten dieses Heft daher nutzen, den Informatikerinnen und Informatikern näherzubringen, warum es sich beim „E-“ in „E-Science“ nicht einfach um die Übertragung bestehender Anwendungen in ein digitales Umfeld handelt, sondern sich hinter dem Begriff neuartige und spannende Fragestellungen auch für die Informatik in Wissenschaft und Praxis verbergen. Die in diesem Heft vorgestellten Beiträge zeigen sowohl für die Wissenschaftslandschaft in Deutschland als auch im Vergleich für andere europäische Länder, welche Dynamik sich hinter dem Begriff verbirgt und warum es sinnvoll ist, sich gerade als Informatikerin oder Informatiker damit zu beschäftigen.

Den Anfang macht der Beitrag „Thesen zur disruptiven Digitalisierung von Hochschulforschung – Faktoren der Skalierung und ein Zukunftsszenario“, der in einer früheren Fassung auf der GI-Jahrestagung 2017 präsentiert wurde. Von der Heyde et al. binden damit das Thema E-Science in einen größeren Zusammenhang der Digitalisierung der Hochschulen allgemein ein.

Der Beitrag von Ramin Yahyapour zeigt den derzeitigen Stand der E-Science-Infrastruktur in Deutschland auf. Yahyapour ist Mitglied des Rates für Informationsinfrastrukturen (RfII) der Gemeinsamen Wissenschaftskonferenz, in der Wissenschafts- und Finanzminister aus Bund und Ländern über Leitlinien der Wissenschaftsförderung diskutieren. Der Rat hat sich bereits 2016 für die Gründung einer Nationalen Forschungsdateninfrastruktur (NFDI) ausgesprochen, die der Wissenschaft eine Grundversorgung an Speichermöglichkeiten und Dienstleistungen für die Nachnutzung der Daten anbieten soll. Auch für die europäische Wissenschaftslandschaft gibt es ähnliche Aktivitäten, wie Yahyapour darstellt.

Der Artikel von Bernard und Mäs über E-Science in den Geo- und Umweltwissenschaften zeigt exemplarisch, welche Anwendungsfragen an die Informatik und hier speziell die Geoinformatik herangetragen werden. Beispiels-

¹ Bohle S (2013) „What is E-science and How Should it Be Managed?“ Nature.com, Spektrum der Wissenschaft (Scientific American), http://www.scilogs.com/scientific_and_medical_libraries/what-is-e-science-and-how-should-it-be-managed/.

weise wird gerade hier mit großen Datenmengen gearbeitet, insofern ergibt sich ein reiches Feld für alle Fragen rund um die Speicherung und Verarbeitung von Big Data. Die Unmöglichkeit, Ergebnisse dieser Verarbeitung manuell zu evaluieren, führt zu Fragen der Reproduzierbarkeit wissenschaftlicher Erkenntnisse.

Kraft stellt in seinem Beitrag „Hybride Edition und analoges Erbe“ den Stand der Technologie in den Digital Humanities vor. Am Beispiel der Digitalisierung einer Handschrift von Alexander von Humboldt wird nicht nur der technologische Ansatz beschrieben. Es wird auch deutlich, wie viel menschliches Wissen derzeit noch für die Generierung qualitativ hochwertiger Forschungsdaten benötigt wird und dass wir in der Digitalisierung an dieser Stelle noch einigen Weg vor uns haben.

Der Blick über die Landesgrenzen zeigt in drei Beiträgen, welche Aktivitäten in den Niederlanden, in Schweden und Großbritannien durchgeführt werden. Besonders in diesen Artikeln, bei denen Autorinnen und Autoren mit einem existierenden Erfahrungsschatz in E-Science zu Wort kommen, zeigt sich, dass die Beschränkung auf eine technologische Fragestellung zu kurz greift. Ohne den Einzelbeiträgen vorgreifen zu wollen, ist dies eine gute Entwicklung für die Informatik: es geht nicht nur um die Nutzung von Technologie. Gerade die Gestaltung der Benutzerschnittstelle, die Organisation von Software-Entwicklungsprozessen oder die Gewährleistung eines sicheren IT-Betriebs sind Fragestellungen, die von Informatikern durch ihre Ausbildung und Erfahrungen besser als von jeder anderen Fachrichtung beantwortet werden können. Die dort beschriebene Einsetzung einer „human e-cloud“ oder von „research engineers“ zeigt, dass die Zukunft der Wissenschaft im Zusammenspiel von hochqualifizierten Wissenschaftlerinnen und Wissenschaftlern mit einer komplexen technischen Infrastruktur liegt und dafür zentrale Support-Institutionen geschaffen werden sollten.

Im Beitrag von Hazeleger et al. wird dargestellt, wie die Niederlande von der Bereitstellung von eScience-Ressourcen durch das zentrale niederländische eScience-Center profitieren. Dabei ist der wesentliche Aspekt eben nicht die Technologie, sondern ein Stab von sehr gut ausgebildeten und multidisziplinären eScience Research Engineers sowie eine offene Wissensdatenbank zu Softwaretools für die Forschung und das Wissen, wie damit umzugehen ist.

Die schwedischen Aktivitäten werden im Artikel von Eriksson et al. vorgestellt. Auch Schweden nutzt ein dediziertes e-Science Research Center (SeRC), und auch dort ist nicht die Bereitstellung der Technologie der Engpass, sondern das Wissen über den effizienten und effektiven Umgang damit. Die schwedische Antwort ist ebenfalls die Bereitstellung von beratenden Wissenschaftlern als „human e-cloud“, also als Teil einer gedachten zentralen und unterstützenden Infrastruktur. Ausgehend von einer Kollaboration zwischen mehreren schwedischen Universitäten hat sich SeRC in der Zwischenzeit zu einem eigenständigen Forschungsinstitut und einem zentralen Player in der schwedischen Wissenschaftslandschaft entwickelt.

Schließlich beschreibt der Beitrag von Neil Chue Hong das Vorgehen in der britischen Wissenschaftslandschaft. Mit dem Software Sustainability Institute (SSI) wird eine eindeutige Linie zwischen der Bereitstellung forschungunterstützender Software und der Durchführung von Wissenschaft gezogen. Besonders deutlich wird das am Beispiel der Reproduzierbarkeit wissenschaftlicher Ergebnisse, welche im Bereich des E-Science unbedingt auf die beweisbare Korrektheit von Code angewiesen ist – hier kann die Informatik nicht nur helfen, sondern ist sogar unverzichtbar.

Die derzeitigen Möglichkeiten und Anstrengungen der Deutschen Forschungsgemeinschaft (DFG) stellt anschließend der Beitrag von Katerbow dar. Dies mag insbesondere für Hochschulen, welche in das Thema E-Science verstärkt einsteigen möchten, nützlich sein, aber es zeigt auch sehr gut den gegenwärtigen Stand vor Einführung einer NFDI wie im Vortrag von Yahyapour.

Die Ausgabe wird durch das Positionspapier der Gesellschaft für Informatik zu „E-Science“ beschlossen. Unsere Motivation war es, die Rolle der Informatik in diesem Kontext zu stärken und in den wissenschaftspolitischen Diskurs einzubringen. Das Positionspapier stellt hierfür einen ersten Schritt dar, es sollte in der nächsten Zeit jedoch, zum Beispiel durch aktives Einbringen der deutschen Informatik-Community in den Aufbau der Nationalen Forschungsdateninfrastruktur, mit Leben gefüllt werden.

Torsten Eymann
Michael Goedicke
Ulrike Lucke