

# Big Data : den digitalen Wandel aktiv gestalten

Autor(en): **Estermann, Beat**

Objekttyp: **Article**

Zeitschrift: **Arbido**

Band (Jahr): - **(2016)**

Heft 3: **Vernichten um zu bewahren? = Détruire pour conserver? =  
Distruggere per conservare?**

PDF erstellt am: **30.06.2024**

Persistenter Link: <https://doi.org/10.5169/seals-769440>

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

d'exemples où le choix de la conservation de toutes les données a déjà été effectué et ce malgré la masse considérable d'informations.

La tendance en informatique est à la conservation des données pour des raisons de traçabilité, de transparence et pour pouvoir reconstruire l'histoire du document. Néanmoins, ne pas éliminer l'information ne veut pas dire ne pas filtrer l'information. Ainsi, sans faire disparaître l'information définitivement, l'intérêt de classer, catégoriser, voire d'enrichir, a encore tout son sens.

L'espace physique est fini. L'espace numérique, qui est in fine physique, l'est tout autant. Cependant, le numérique est, sans aucune comparaison ou mesure, le champion de la compression d'espace physique. Les exemples et idées développés ici montrent que les limites actuelles et surtout futures sont

plutôt le fait d'utilisation extrême tel l'archivage du Web ou des expériences du CERN avec leurs tailles de stockage de plus de 200 Po (soit 200 000 To).

Paradoxalement, si la conservation et la diffusion sont facilitées par les technologies, il est urgent de trouver des moyens de mieux collecter l'information publique, voire l'information privée (courriels, réseaux sociaux fermés, photographies et films produits en masse comme autant de souvenirs qui ne dureront pourtant guère, faute de réflexion), car des millions de documents disparaissent tous les jours. Sur le web par exemple, la durée de vie d'une page est estimée entre 50 et 100 jours en moyenne selon Brewster Kahle, fondateur principal de l'«Internet Archive».

Mais la bonne nouvelle, c'est que la page Wikipédia du rat-taupe nu du

Kenya est bien à l'abri parmi les Peta données d'archive.org et ne «souffre» d'aucun risque d'élimination!

*L'auteur remercie Esther Bettiol pour sa relecture.*

Contact: [arnaud.gaudinat@hesge.ch](mailto:arnaud.gaudinat@hesge.ch)

Twitter: @AGaudinat

#### L'auteur

Arnaud Gaudinat est ingénieur en informatique de formation. Après 15 ans d'expérience en recherche appliquée, il intègre en 2010 la filière information documentaire de la HEG. Il est actuellement adjoint scientifique et enseigne les sciences de l'information, la programmation et la gestion de contenu sur le Web. Il effectue principalement sa recherche dans le domaine de la fouille de données sur le Web et de la webométrie.

## Big Data: Den digitalen Wandel aktiv gestalten

Beat Estermann, Wissenschaftlicher Mitarbeiter, Berner Fachhochschule

**Vor dem Hintergrund der zunehmenden Digitalisierung von Wirtschaft und Gesellschaft und als Antwort auf Entwicklungen, welche gemeinhin unter dem Oberbegriff «Big Data» zusammengefasst werden, hat der Bundesrat im April 2016 die Strategie «Digitale Schweiz» verabschiedet. Der vorliegende Artikel gibt einen Überblick über die Implikationen für Gedächtnisinstitutionen und gibt Denkanstösse, wie diese den digitalen Wandel aktiv gestalten können. Ein aktuelles Forschungsprojekt zeigt auf, dass die Chancen für Gedächtnisinstitutionen vor allem bei der verbesserten Sichtbarkeit der Institutionen und der Bestände liegen sowie bei der zunehmenden Vernetzung und der Neugestaltung der Beziehung zu den Nutzerinnen und Nutzern. Allerdings legen die Forschungsergebnisse auch nahe, dass die Schweizer Institutionen im internationalen Vergleich Aufholbedarf haben.**

Vor dem Hintergrund der zunehmenden Digitalisierung von Wirtschaft und Gesellschaft hat der Bundesrat im April 2016 die Strategie «Digitale Schweiz» verabschiedet. Sie dient dem Zweck, die digitale Entfaltung der Gesellschaft zu fördern, indem die digitale Transformation und der damit verbundene Strukturwandel proaktiv und vernetzt gestaltet werden. Der Bundesrat reagiert mit der neuen Strategie unter anderem auch auf Entwicklungen, welche gemeinhin unter dem Oberbegriff «Big Data» zusammengefasst werden.

Der vorliegende Artikel gibt einen Überblick darüber, was es mit Big Data auf sich hat, welche Chancen und Risiken für Gedächtnisinstitutionen damit verbunden sind und wie sich die Stossrichtungen der bundesrätlichen Strategie auf den Bereich der Bibliotheken und Archive übertragen lassen. Anhand der Ergebnisse einer Umfrage unter Gedächtnisinstitutionen in zehn Ländern wird aufgezeigt, wo die Schweizer Institutionen mit Blick auf die digitalen Transformationsprozesse

im internationalen Vergleich stehen. Den Abschluss bilden einige Denkanstösse und Handlungsempfehlungen, um in der Debatte über Bewahren versus Ausscheiden den Blick aufs Wesentliche nicht zu verlieren: die aktive Mitgestaltung des digitalen Wandels.

#### Chancen und Risiken von Big Data

Im Auftrag des Bundesamts für Kommunikation hat die Berner Fachhochschule 2015 eine Studie durchgeführt, um die Chancen und Risiken von Big Data zu erörtern und den kollektiven Handlungsbedarf aufzuzeigen (Jarchow/Estermann 2015). Grundlage der Studie waren qualitative Interviews sowie eine quantitative Erhebung bei über 800 Personen in der Schweiz, welche mit Daten zu tun haben – sei es privat, beruflich oder im Rahmen ihres politischen oder zivilgesellschaftlichen Engagements.

Wie aus der Studie hervorgeht, ist Big Data – definiert als «Praktiken, welche sich dadurch kennzeichnen, dass grosse Datenmengen aus sehr unter-

schiedlichen Quellen zusammengeführt und – oftmals sehr zeitnah – analysiert und ausgewertet werden» – sowohl mit Chancen als auch mit Risiken verbunden.

Was die positiven Aspekte von Big Data angeht, ergab die Studie, dass rund 80% der Befragten davon ausgehen, dass Big Data es ermöglicht, bisher unbekannte Zusammenhänge aufzudecken und neue Erkenntnisse zu gewinnen. Konkret wurde Big Data von rund zwei Dritteln der Befragten als Mittel gesehen, um genauere Prognosen und Risikoabschätzungen zu ermöglichen, um vermehrt daten- und faktenbasiert zu handeln und um operative und logistische Abläufe zu optimieren. Bei den Risiken standen die grundsätzliche Bedrohung der Privatsphäre und die missbräuchliche Verwendung von Daten, inklusive Datendiebstahl, im Vordergrund. Rund 80% der Befragten sahen darin ein ernst zu nehmendes Risiko von Big Data.

Interessanterweise liessen sich die Befragten nicht in zwei Lager von Big-Data-Pessimisten und Big-Data-Optimisten einteilen, sondern sie unterschieden sich primär darin, wie hoch sie das Potenzial von Big Data einschätzten – im Positiven wie auch im Negativen.

### Kollektiver Handlungsbedarf

Ist das Potenzial von Big Data erkannt, so stellt sich die grundsätzliche Her-

ausforderung, die Entwicklungen im Bereich Big Data so zu kanalisieren, dass sie zum Nutzen der Menschen reichen. Hier ist aus Sicht der Befragten insbesondere die öffentliche Hand gefragt, wobei die Studienergebnisse eine Drei-Punkte-Strategie nahelegen: den Aufbau und die Pflege einer nationalen Dateninfrastruktur, Massnahmen zum Schutz vor Missbräuchen und Datenschutzverletzungen sowie die Umsetzung des Prinzips der persönlichen Datenhoheit, mit dem langfristigen Ziel, die Menschen zu befähigen, die Kontrolle über ihre Daten besser auszuüben.

Diese drei Stossrichtungen finden sich in der Strategie des Bundesrates für eine «Digitale Schweiz» wieder, in der unter anderem die drei folgenden Ziele vorgegeben werden (Schweizerische Eidgenossenschaft 2016):

- **Die Schweiz verfügt über eine kohärente und zukunftsorientierte Datenpolitik.**
- **Die Schweiz verfügt über eine nationale Dateninfrastruktur:** «Im Sinne einer «Good Data Governance» wird die Bereitstellung von Daten als Infrastrukturressourcen zur freien Weiterverwendung verstärkt. Durch den Aufbau einer nationalen Dateninfrastruktur, analog zur bereits bestehenden Geodateninfrastruktur, soll dem Open-Data-Prinzip über die gesamte Verwaltung hinweg, in der Forschung und teilweise auch in der Privatwirt-

schaft zum Durchbruch verholfen werden» (ibid., S. 6).

- **Die Einwohnerinnen und Einwohner der Schweiz können die Kontrolle über ihre eigenen Daten ausüben:** «Damit das Grundrecht jeder Person auf informationelle Selbstbestimmung umgesetzt werden kann und um Missbräuchen und Ungleichgewichten entgegenzuwirken, müssen die Datenschutzregeln neu gestaltet werden. Zudem gilt es, in enger Kooperation mit allen Mitarbeiterinnen und Mitarbeitern von Daten Mechanismen zu schaffen und Dienste bereitzustellen, die es den Einzelnen erlauben, sie betreffende Personen- und Sachdaten – unter grösstmöglicher Kontrolle – für die Weiterverwendung durch Dritte freizugeben oder deren Weiterverwendung zu unterbinden» (ibid., S.6).

Eine Sondierung der Berner Fachhochschule bei Schlüsselakteuren mit Blick auf den Aufbau und die Pflege einer nationalen Dateninfrastruktur ergab folgendes Bild (Estermann et al. 2016):

Eine nationale Dateninfrastruktur ist eine schweizweite (verteilte) technische Infrastruktur (Portale, Plattformen, Services etc.), die den Zugang zu und den Austausch von Daten auf Basis von definierten Regeln ermöglicht. Ihr Zweck besteht darin, die datenbasierte Wertschöpfung zu unterstützen und den potenziellen Mehrwert bestehender Daten freizusetzen.

Eine nationale Dateninfrastruktur ist aus vier Perspektiven zu betrachten (siehe Abb. 1):

- Die **Open-Data-Perspektive** stellt die Sekundärnutzung von Daten durch Dritte in den Fokus. Um diese zu vereinfachen, ist in einem ersten Schritt die rechtliche und technische Offenheit der Daten und Inhalte (freie Lizenzen, freie Datenformate, offener Zugang) zentral; während in einem weiteren Maturitätsschritt die Herstellung der semantischen Interoperabilität zwischen Datensätzen unterschiedlichen Ursprungs (Linked Data) im Vordergrund steht.
- Die **Perspektive der Basis-Register** stellt die gemeinsame Verwendung von Referenzdaten ins Zentrum. Im Vordergrund stehen Authentizität, Vollständigkeit und Interoperabilität

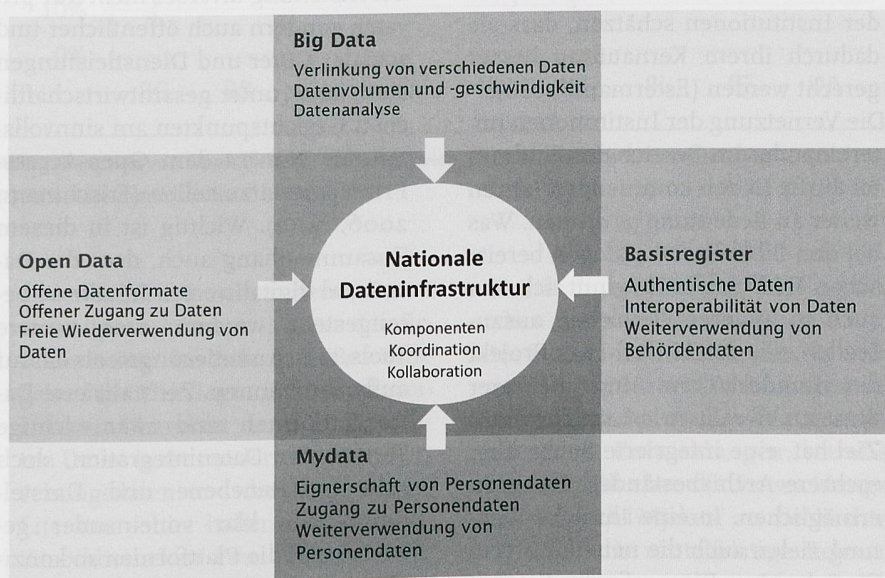


Abb. 1: Vier Perspektiven einer nationalen Dateninfrastruktur.

(Quelle: Estermann et al. 2016, CC BY-SA 4.0)

der Daten, aber auch die Bereitschaft verschiedener Akteure, einen gemeinsamen Pool von Basisdaten (z.B. ein nationales Unternehmensverzeichnis, ein zentrales Adressverzeichnis, ein Gemeindeverzeichnis oder Geo-Basisdaten) zu pflegen und zu nutzen.

- Die **Big-Data-Perspektive** fokussiert auf neue Methoden der Datenanalyse und das Gewinnen von neuen Erkenntnissen dank dem Aufbrechen von Datensilos und der Analyse von immer grösseren Mengen von Daten, welche oftmals in immer höherer Kadenz vorliegen.
- Und schliesslich stellt die **«Mydata»-Perspektive** das Empowerment des Einzelnen ins Zentrum, was die Verfügungsgewalt über die einen betreffenden Personendaten anbelangt, und damit einhergehend die Befähigung zu ihrer gezielten Freigabe für die Weiterverwendung zu ausgewählten Zwecken (z.B. für die medizinische Forschung).

Eine nationale Dateninfrastruktur ist kein monolithischer Block und soll auch nicht als Schweizer Insellösung funktionieren, sondern muss international vernetzt sein. Der Staat ist mitverantwortlich für die Bereitstellung der Daten und der technischen Infrastruktur, wobei es zum Aufbau einer nationalen Dateninfrastruktur eines schritt-

80% der Befragten gehen davon aus, dass Big Data es ermöglicht, neue Erkenntnisse zu gewinnen.

weisen Prozesses bedarf, der sowohl einer Bottom-up- als auch einer Top-down-Logik folgt: Die Bottom-up-Logik steht im Vordergrund, wenn einzelne Akteure ihre Daten veröffentlichen, Bürger und Bürgerinnen sowie die Privatwirtschaft mit einbeziehen und Anwendungsfälle aufzeigen. Nach der Top-Down-Logik sollten der strategische Rahmen abgesteckt und politische Aufträge formuliert werden. Für einen erfolgreichen Aufbau einer nationalen Dateninfrastruktur müssen diese beiden Entwicklungslogiken miteinander korrespondieren, wobei der Zivilgesellschaft eine Vermittlerrolle zukommt.

Neben diesen Grundpfeilern, welche bisher unumstritten sind, hat die Sondierung auch einiges an weiterem Klärungs- und Koordinationsbedarf zutage gefördert, was die Governance einer solchen Dateninfrastruktur anbelangt.

### Relevanz für Bibliotheken und Archive

Aus der Perspektive der Bibliotheken und Archive stellt sich die Frage der Relevanz dieser Entwicklungen für ihre

Die «Mydata»-Perspektive stellt das Empowerment des Einzelnen ins Zentrum.

eigenen Wandlungsprozesse angesichts der zunehmenden Digitalisierung der Gesellschaft. Diese soll im Folgenden in Form von kurzen Thesen skizziert werden, wobei diese als Diskussionsbeitrag in einer laufenden Debatte zu verstehen sind und nicht als abschliessende Bewertung:

- Wie aus einer internationalen Befragung von Gedächtnisinstitutionen hervorgeht, können Gedächtnisinstitutionen dank Open Data und Open Content die Sichtbarkeit der Institutionen und ihrer Bestände verbessern, die Vernetzung zwischen den Institutionen erhöhen, die Inhalte für Benutzerinnen und Benutzer einfacher zugänglich machen und die Interaktion mit ihren Nutzerinnen und Nutzern verbessern. Rund 70% der Institutionen schätzen, dass sie dadurch ihrem Kernauftrag besser gerecht werden (Estermann 2016).
- Die Vernetzung der Institutionen untereinander im Bereich der Findmittel dürfte in den kommenden Jahren weiter an Bedeutung gewinnen. Was bei den Bibliothekskatalogen bereits schon Tradition hat, beginnt sich nun auch in anderen Bereichen auszubreiten, wie das Linked-Data-Projekt des Bundesarchivs und mehrerer Staatsarchive illustriert, welches zum Ziel hat, eine integrierte Suche über mehrere Archivbestände hinweg zu ermöglichen. In eine ähnliche Richtung zielen auch die neue Plattform für Schweizer Fotografie des Fotobüros Bern und die geplante Plattform der darstellenden Künste der Schwei-

zerischen Theatersammlung und des Schweizer Tanzarchivs. Wir haben es mit einer Entwicklung zu tun, bei der im Rahmen von thematischen Plattformen die Metadaten in diversen Bereichen unter anderem dank Linked Data immer weiter integriert werden, was nicht nur die Suche, sondern auch die Verknüpfung mit weiteren Daten vereinfacht und im Endeffekt auch neuen Forschungserkenntnissen den Weg bereiten dürfte.

- Die Gedächtnisinstitutionen sollten sich an der Diskussion zu den Basisregistern beteiligen: Einerseits geht es darum, bestehende Register aus anderen Bereichen zu identifizieren, welche bei der Anreicherung und Vernetzung der Metadaten eine zentrale Rolle spielen könnten; andererseits gibt es auch im Bereich der Gedächtnisinstitutionen selber Verzeichnisse, die am besten gemeinsam gepflegt und allen Institutionen als Linked-Data-Dienste zur Mitbenutzung bereitgestellt werden. Der Betrieb eines Registers aller Schweizer Gedächtnisinstitutionen wäre ein naheliegendes Vorhaben.
- Zentral ist die Erkenntnis, dass Daten den Charakter von Infrastruktur-Ressourcen haben. Diese zeichnen sich dadurch aus, dass sie durch ihre Nutzung nicht aufgebraucht werden und gleichzeitig als Inputs für eine breite Palette von nachgelagerten Wertschöpfungsprozessen dienen. Insofern sie Voraussetzung für die Bereitstellung diverser nicht nur privater, sondern auch öffentlicher und sozialer Güter und Dienstleistungen sind, ist es unter gesamtwirtschaftlichen Gesichtspunkten am sinnvollsten, sie gemäss dem Open-Access-Prinzip bereitzustellen (Frischmann 2006, 2012). Wichtig ist in diesem Zusammenhang auch, dass die Daten und digitalisierten Inhalte so bereitgestellt werden, dass weitere Tools, z. B. Annotierungstools darauf aufbauen können. Zentralisierte Datenplattformen sind zwar wichtige Treiber der Datenintegration, doch sollten Datenebene und Darstellungsebene klar voneinander getrennt und die Plattformen so konzipiert werden, dass Dritte die Daten oder Teile davon auch in anderen Kontexten nutzen und im Rahmen

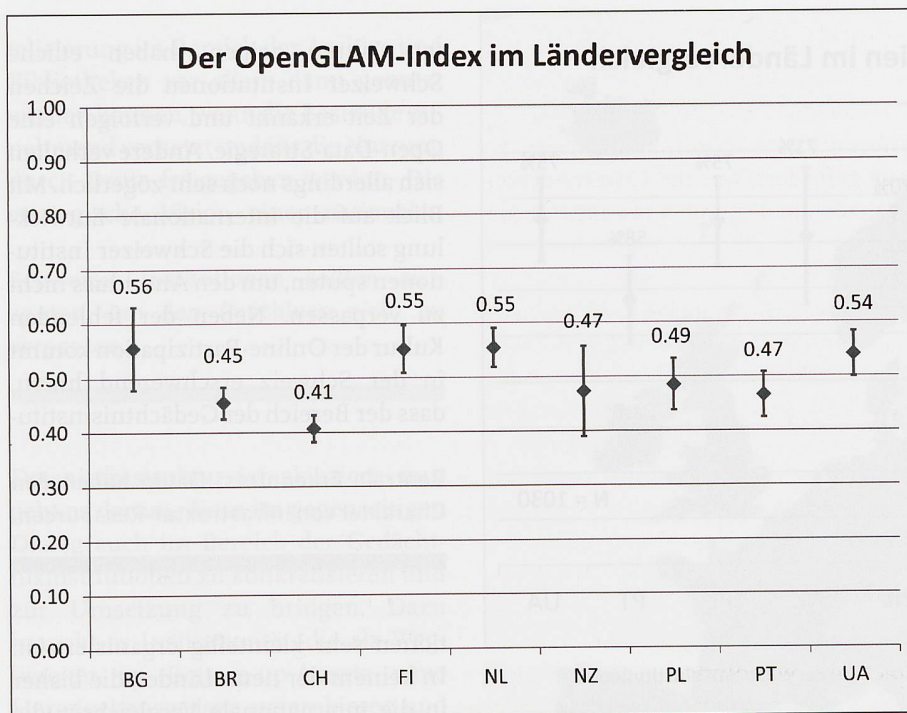


Abb. 2: Der OpenGLAM-Index im Ländervergleich. (Quelle: Estermann 2016, CC BY 4.0)

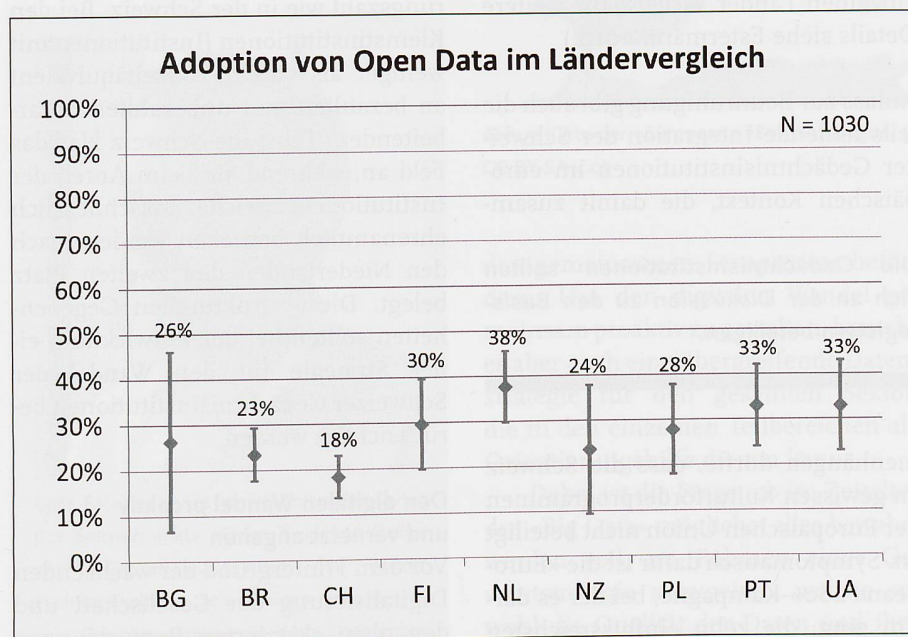


Abb. 3: Adoption von Open Data im Ländervergleich, in % der Gedächtnisinstitutionen. (Quelle: Estermann 2016, CC BY 4.0)

anderer Dienste mit weiteren Daten verknüpfen können. Neben der Erschliessung historischer Bestände gelten die Überlegungen zum Infrastruktur-Charakter der bereitgestellten Daten insbesondere auch für Forschungsdaten und die akademische Literatur – hier haben die Universitätsbibliotheken eine zentrale Rolle zu spielen – sowie für Datenbanken der Verwaltungsgeschäfte, deren Archivierung und Langzeitbetrieb im

Zuständigkeitsbereich der staatlichen Archive liegt.

- In Sachen informationelle Selbstbestimmung dürften Bibliotheken und Archive nicht unmittelbar von den aktuellen Entwicklungen betroffen sein. Es gilt aber, die Entwicklungen im Bereich Mydata zu beobachten und auf die mittelfristige Relevanz für Archive zu überprüfen, welche personenbezogene Daten in ihren Beständen haben, die für weiterfüh-

rende Auswertungen von Interesse sein könnten. Des Weiteren sollten die Vor- und Nachteile sowie die ethischen Implikationen im Zusammenhang mit der Personalisierung von Angeboten und dem Tracking der Nutzung von Beständen mit Blick auf den Datenschutz sorgfältig abgewogen werden. Wie die Big-Data-Studie der Berner Fachhochschule zeigt, stehen die Nutzerinnen und Nutzer der zunehmenden Personalisierung von Angeboten eher zwiespältig gegenüber: Einerseits kann daraus eine deutliche Verbesserung von Dienstleistungen resultieren, andererseits ruft sie bisweilen auch Unbehagen hervor und wird als invasiv empfunden.

- Im Zusammenhang mit Big Data dürfte der Fokus bei den Bibliotheken und Archiven vor allem auf dem Aspekt der Verknüpfung von heterogenen Datenbeständen liegen. Was die Verarbeitung grosser Datenmengen anbelangt, so werden bei den klassischen Daten eher geringe Volumina erreicht. Eine Ausnahme dürften die Entwicklungen im Bereich Bildanalyse und Bilderkennung bilden, wo beim eigentlichen Content durchaus mit grossen Datenmengen zu rechnen ist. Von einer hohen Aktualisierungskadenz dürfte bei der Datenanalyse hingegen kaum etwas zu spüren sein. Im Zusammenhang mit der Integration von heterogenen Datenbeständen dürfte es lohnenswert sein, ergänzend zum Linked-Data-Ansatz auch das Potenzial von Wikidata auszuloten, welches sich immer stärker in Richtung einer globalen Datendrehscheibe entwickelt und dessen Daten mittlerweile auch über einen SPARQL-Endpoint abrufbar sind, womit sie sich nahtlos ins Semantic Web (Linked Data) einfügen.

### Schweizer Institutionen international im Hintertreffen

Besondere Aufmerksamkeit ist auch dem internationalen Kontext zu widmen. Im Ländervergleich spielt die Schweiz, was die technische Infrastruktur anbelangt, in der besten Liga. Dies könnte mit Blick auf die aktuelle Digitalisierung der Gesellschaft durchaus zuversichtlich stimmen. Was die Kultur der Online-Partizipation (gemessen am

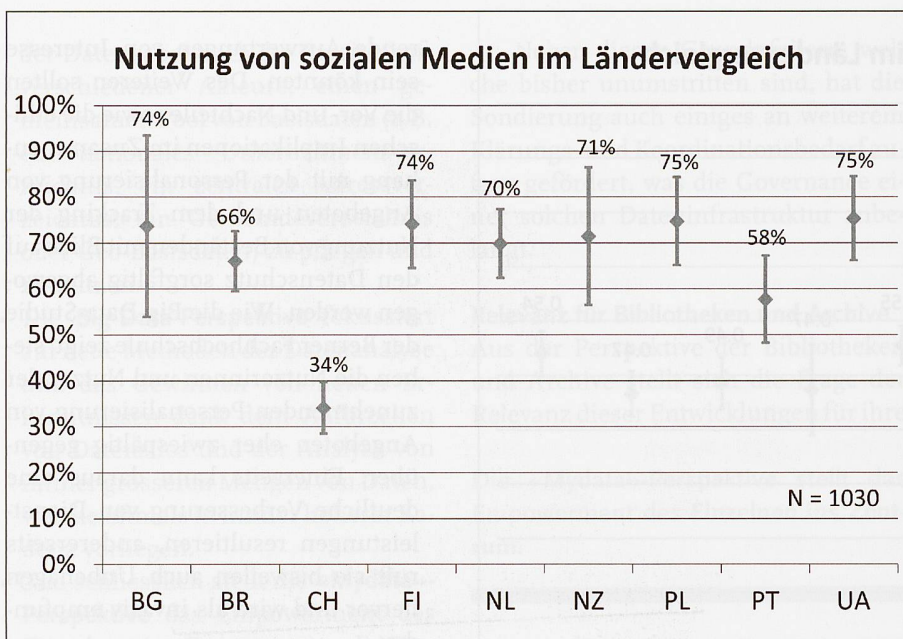


Abb. 4: Nutzung von sozialen Medien im Ländervergleich, in % der Gedächtnisinstitutionen. (Quelle: Estermann 2016, CC BY 4.0)

UNPACS-E-Participation-Index) be- trifft, rangiert die Schweiz allerdings im hinteren Mittelfeld. Im Ergebnis sind die Schweizer Gedächtnisinstitutionen, was die Öffnung der Daten und der Inhalte anbelangt, wie auch in Bezug auf die Nutzung der Möglichkeiten der Online-Partizipation im europäischen Vergleich auf den hinteren Rängen zu finden. Abbildung 2 zeigt den OpenGLAM-Index im internationalen Vergleich. Der Index stellt eine Messgrösse für die drei Dimensionen «Einbezug von Nutzern», «Open Data» und «Open Content» dar, womit er einen guten Anhaltspunkt gibt, wo die Gedächtnisinstitutionen eines Landes bezüglich des digitalen Wandels stehen.

**Eine nationale Dateninfrastruktur soll nicht als Schweizer Insellösung funktionieren, sondern international vernetzt sein.**

Die Abbildungen 3 und 4 zeigen die Anteile der Gedächtnisinstitutionen in den jeweiligen Ländern, bei denen das Bereitstellen von Daten als «Open Data» bzw. die Nutzung von sozialen Medien im Kontakt mit den Nutzerinnen und Nutzern gängige Praxis ist. (Die Balken zeigen 95%-Konfidenzintervalle an, deren Länge aufgrund der unterschiedlichen Stichprobengrössen der

einzelnen Länder variiert; für weitere Details siehe Estermann 2016.)

Anlass zur Beunruhigung gibt auch die teils fehlende Integration der Schweizer Gedächtnisinstitutionen im europäischen Kontext, die damit zusam-

**Die Gedächtnisinstitutionen sollten sich an der Diskussion zu den Basisregistern beteiligen.**

menhängen dürfte, dass die Schweiz an gewissen Kulturförderprogrammen der Europäischen Union nicht beteiligt ist. Symptomatisch dafür ist die «Europeana 280»-Kampagne, bei der es darum ging, die zehn einflussreichsten Kunstwerke jedes der europäischen Länder zu identifizieren und online verfügbar zu machen. Wie in Abbildung 5 ersichtlich ist, erscheint die Schweiz dort als einsame Insel inmitten von Europa; ein Phänomen, das im kulturellen Bereich auch schon bei früheren Aktionen zu beobachten war. Vor diesem Hintergrund stellt sich die Frage, was es für Schweizer Institutionen längerfristig bedeutet, bei europäischen Initiativen aussen vor zu bleiben.

Wie die Erfahrungen im Zusammenhang mit der schweizerischen OpenGLAM-Arbeitsgruppe in den letz-

ten Jahren zeigen, haben etliche Schweizer Institutionen die Zeichen der Zeit erkannt und verfolgen eine Open-Data-Strategie. Andere verhalten sich allerdings noch sehr zögerlich. Mit Blick auf die internationale Entwicklung sollten sich die Schweizer Institutionen sputen, um den Anschluss nicht zu verpassen. Neben der fehlenden Kultur der Online-Partizipation kommt in der Schweiz erschwerend hinzu, dass der Bereich der Gedächtnisinstitu-

**Zentrale Erkenntnis: Daten haben den Charakter von Infrastruktur-Ressourcen.**

tionen sehr kleinteilig organisiert ist: In keinem der neun Länder, die bisher in die internationale Vergleichsstudie einbezogen wurden, gibt es so viele Institutionen im Verhältnis zur Bevölkerungszahl wie in der Schweiz. Bei den Kleinstinstitutionen (Institutionen mit weniger als einem Vollzeitäquivalent an bezahlten und unbezahlten Mitarbeitenden) führt die Schweiz klar das Feld an, während sie beim Anteil der Institutionen, welche ausschliesslich ehrenamtlich betrieben werden, nach den Niederlanden den zweiten Platz belegt. Diese strukturellen Gegebenheiten sollten bei der Entwicklung einer Strategie für den Wandel der Schweizer Gedächtnisinstitutionen berücksichtigt werden.

**Den digitalen Wandel proaktiv und vernetzt angehen**

Vor dem Hintergrund der wachsenden Digitalisierung der Gesellschaft und den oben skizzierten Entwicklungen liegt es nahe, dass die Schweizer Gedächtnisinstitutionen den digitalen Wandel proaktiv und vernetzt angehen sollten: Proaktiv, weil sich der Wandel nicht automatisch aus der komfortab-

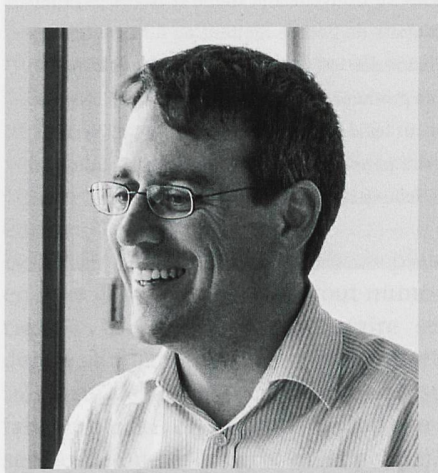
**Was die Kultur der Online-Partizipation betrifft, rangiert die Schweiz im hinteren Mittelfeld.**

len Situation hinsichtlich der technischen Infrastruktur und der finanziellen Ausstattung ergibt. Vernetzt, weil die Chancen der fortschreitenden Digi-

talisierung im Bereich der Archive und Bibliotheken vor allem dann genutzt werden können, wenn die Bestände zunehmend vernetzt und für die Nutzung durch Dritte freigegeben werden. Die strategische Vision einer nationalen

Schweizer Institutionen sollten sich sputen, um den Anschluss nicht zu verpassen.

Dateninfrastruktur ist skizziert; nun geht es darum, diese im gegenseitigen Dialog auch im Bereich der Gedächtnisinstitutionen zu konkretisieren und zur Umsetzung zu bringen. Dazu braucht es Institutionen, die als Wegbereiter im Bottom-up-Ansatz neue Herangehensweisen ausprobieren und



Beat Estermann ist stellvertretender Leiter des Schwerpunkts «Open & Linked Data» am E-Government-Institut der Berner Fachhochschule. Seit einigen Jahren gilt sein Forschungsinteresse dem Themenbereich «Open Data» und im speziellen «OpenGLAM», d.h. Open Data und Crowdsourcing unter Gedächtnisinstitutionen. Er koordiniert die schweizerische OpenGLAM-Arbeitsgruppe des Vereins «opendata.ch», welche sich für Open Data im Bereich der Gedächtnisinstitutionen einsetzt, sowie die eCH-Fachgruppe «Open Government Data», welche sich Standardisierungsfragen rund um Open Government Data widmet. Er erbringt Consulting-Dienstleistungen für Bundesämter und Gedächtnisinstitutionen und hat 2015 im Auftrag des Bakom die Studie «Big Data: Chancen, Risiken und Handlungsbedarf des Bundes» mitverfasst.



Abb. 5: Karte der «Europeana 280»-Kampagne. (Quelle: Llywelyn2000, Wikimedia Commons, CC BY-SA 4.0)

den gemeinsamen Lernprozess befördern. Um den digitalen Wandel gemeinsam proaktiv zu gestalten, braucht es aber auch eine übergreifende Datenstrategie für den gesamten Sektor, die in den einzelnen Teilbereichen als Orientierungshilfe dienen kann.

Dabei ist die Frage, ob im Zeitalter des «Big Data» möglichst alles bewahrt werden soll, im Rahmen einer Gesamtstrategie anzugehen, welche sowohl die Qualität der Daten und ihre Vernetzung als auch datenschützerische Belange berücksichtigt. In die entsprechenden Debatten sollten sich auch die Fachleute I+D einbringen.

Kontakt: [beat.estermann@bfh.ch](mailto:beat.estermann@bfh.ch)  
Twitter: @beatestermann

#### Bibliographische Angaben

- Estermann Beat, OpenGLAM Benchmark Survey – Measuring the Advancement of Open Data/Open Content in the Heritage Sector. Paper presented at the International Symposium on the Measurement of Digital Cultural Products, 9-11 May 2016, Montreal, Canada.
- Estermann Beat, Fraefel Marianne, Neuron Alessia, Conceptualizing a National Data Infrastructure for Switzerland. Paper Presented at the EGPA 2016 Conference, 24-26 August 2016, Utrecht, The Netherlands.
- Frischmann Brett M., An Economic Theory of Infrastructure and Commons Management. *Minnesota Law Review*, 89, 2006, pp. 917-1030.
- Frischmann Brett M., Infrastructure: The Social Value of Shared Resources. *Oxford Scholarship Online*, 2012.
- Jarchow Thomas und Estermann Beat, Big Data: Chancen, Risiken und Handlungsbedarf des Bundes. Studie im Auftrag des Bundesamts für Kommunikation, Berner Fachhochschule, Oktober 2015.
- Schweizerische Eidgenossenschaft, Strategie «Digitale Schweiz», April 2016.

## *Big data: façonner le virage numérique de manière proactive*

Face à la numérisation croissante de l'économie et de la société et pour répondre aux évolutions liées au «big data», le Conseil fédéral a adopté en avril 2016 la stratégie «Suisse numérique». Celle-ci a pour but de promouvoir et d'encadrer la transformation numérique de la société, de sorte que la mutation numérique et les changements de structures y relatifs soient façonnés de manière proactive et en réseau. Avec sa nouvelle stratégie, le Conseil fédéral réagit notamment aux développements de ce que l'on appelle communément les «big data». Ceux-ci recèlent autant de chances que de risques: parmi les chances, on notera surtout la possibilité de gagner de nouvelles connaissances en découvrant des relations jusque-là inconnues; du côté des risques, le plus inquiétant concerne les possibles violations de la protection des données. Une des parties-clés de la stratégie «Suisse numérique» consiste dans la mise en place d'une infrastructure de données nationale fondée sur le principe de l'open data et sur l'application du principe de l'autodétermination informationnelle dans le domaine des données personnelles. L'infrastructure de données nationale est tout d'abord une infrastructure technique (portails, plateformes, services, etc.) répartie à l'échelle de la Suisse qui permet l'accès et l'échange de données sur la base de règles définies. Son but est de soutenir la création de richesses à partir de données et surtout de libérer la valeur ajoutée potentielle des données déjà existantes. Une infrastructure de données nationale peut être examinée sous quatre perspectives différentes, chacune d'entre elles mettant l'accent sur un aspect particulier, comme l'exploitation secondaire des données (open data), l'utilisation en commun de données de référence (registres de base), de nouvelles méthodes d'analyse des données (big data) ou encore l'autodétermination informationnelle dans le domaine des données personnelles (mydata).

Les institutions patrimoniales devraient s'engager activement dans le développement d'une infrastructure de données nationale, car celle-ci leur permet d'être mieux adaptées à leur mission. Ainsi, le renforcement de la mise en réseau entre les institutions dans le cas des instruments de recherche est particulièrement significatif, car ce n'est pas seulement la recherche que cela simplifie, mais également la relation avec d'autres données, et par conséquent l'acquisition de nouvelles connaissances. D'autres opportunités découlent de l'exploitation collaborative des registres de base, ainsi que de la séparation claire entre les niveaux de la donnée et de sa représentation, ce qui permet à un tiers d'utiliser les données également dans d'autres contextes.

Comme le montre une étude comparative internationale, les institutions patrimoniales suisses se trouvent à la traîne d'autres pays par rapport à la mise en œuvre de l'open data et de la participation en ligne des publics. Les résultats de l'étude mettent en lumière que les institutions devraient aborder le virage numérique de façon proactive et en réseau, car l'excellence des infrastructures techniques et le bon équipement financier des institutions suisses ne suffisent pas. Pour que le virage numérique se produise avec succès, il faut aussi bien des institutions innovatrices qui se lancent dans de nouvelles approches, qu'une stratégie globale sur les données pour l'ensemble du secteur. Ainsi, on le voit, la problématique du «tout-conserver» chère aux partisans des «big data» est à penser dans une stratégie globale tenant compte à la fois de la qualité des données et de leur mise en réseau que de l'évaluation des risques sur la protection des données. Et les professionnels I+D sont tout désignés pour y apporter leurs compétences.

## Le coût environnemental de la gestion documentaire

Sophie Roy, responsable du blog *Convergence* de l'Association des archivistes du Québec

**De plus en plus, l'environnement est une préoccupation de nos sociétés et cela se reflète dans notre volonté de réduire la consommation de papier. Pourtant, est-ce que vraiment la production, l'utilisation et la conservation de données analogiques polluent moins que leurs consœurs numériques? Rien n'est moins sûr d'après plusieurs études.**

**L'environnement a la cote actuellement**  
Depuis l'avènement de l'ère industrielle, la santé de notre planète soulève

des questions. Le contexte de croissance économique fait s'accélérer la destruction de nos ressources naturelles et pousse l'humanité à agir. En 2015, la Conférence internationale de Paris sur le climat (COP21) semble avoir eu des résultats encourageants pour l'avenir de la planète. Les décideurs, gouvernements et entreprises, ont désormais la responsabilité d'agir. De plus, l'environnement a la cote dans notre quotidien avec les tendances zéro déchets, légumes moches, mobilité douce et voitures électriques, compostage et autres réutilisations des ressources.

L'industrialisation passe aujourd'hui à une étape de dématérialisa-

tion et de robotisation dans un but d'efficacité, de mobilité, mais aussi de réduction de la consommation de papier. Mais, dans le contexte de croissance mondiale, qu'en est-il du coût environnemental de la conservation et de la duplication sous toutes ses formes de la masse documentaire numérique? Et comment le spécialiste de la documentation peut-il intervenir, réagir face à ces enjeux de taille?

### **Quel coût environnemental des TI?**

Le calcul du coût environnemental des technologies de l'information n'est pas encore courant dans les organisations. Les réseaux et les serveurs consomment