

Die Erschließung des Kontexts. Neue Perspektiven auf ein bewährtes Prinzip

von Gerhard Müller und Silke Jagodzinski

Die Programmkommission des Internationalen Archivrats (ICA) berief 2012 eine Expertengruppe, die Expert Group on Archival Description (EGAD)¹, mit dem Auftrag ein, die elementaren Aspekte der archivischen Erschließung neu zu betrachten. Ausgangspunkt sind die vier internationalen Richtlinien:

- General International Standard Archival Description, ISAD (G), (1994/1999) – Richtlinie für die Erschließung von Archivbeständen –,
- International Standard Archival Authority for Corporate Bodies, Persons, and Families, ISAAR (CPF) (1996/2003) – Richtlinie für die Beschreibung von Aktenbildnern –,
- International Standard for Describing Functions, ISDF (2007) – Richtlinie für die Beschreibung von Funktionen von Körperschaften – sowie

- International Standard for Describing Institutions with Archival Holdings, ISDIAH (2008) – Richtlinie für die Beschreibung von Archiven.

Im Vorwort zu ISDIAH, der zuletzt veröffentlichten Richtlinie, konstatierten die Autoren: „ICA/CBPS² is aware that in the future a single reference model should be developed to reconcile and harmonize the four international descriptive standards developed since the 1990s.“³ Nach einer vierjährigen Phase (2012–2016) legte die EGAD im Septem-

1 International Council on Archives (Hrsg.), EGAD. <https://www.ica.org/en/about-egad> [Stand 14.01.2019, gilt ebenfalls für alle nachfolgenden Hinweise auf Internetseiten].

2 Die EGAD übernahm 2012 in Teilen die Aufgaben des vormaligen Committee on Best Practices and Standards (CBPS).

3 ISDIAH 2008, S. 6.

ber 2016 den Entwurf eines vollständig neu erarbeiteten Modells unter dem Titel „Records in Context – A Conceptual Model“, RiC-CM, für ein öffentliches Review vor. Bis zum Ende der Kommentierungsphase im Januar 2017 gingen 62 Kommentare aus 19 Ländern im Umfang von über 200 Seiten ein. Die Beiträge haben das große Interesse der Fachöffentlichkeit an der Weiterentwicklung der Archivstandards gezeigt, aber auch verdeutlicht, dass vor den Mitgliedern der EGAD noch Arbeit liegt, um Widersprüche und Lücken im Modell aufzuklären, den Text zu konsolidieren und den Entwurf einer Ontologie (RiC-O) vorzulegen. Dieser Beitrag beruht auf einem Vortrag auf dem 19. Deutsch-Niederländischen Archivsymposium in Leeuwarden im Oktober 2018. Im ersten Abschnitt werden die Intention hinter „Records in Context – A Conceptual Model“ (RiC-CM) vorgestellt sowie dessen wesentliche Begriffe eingeführt. Grundlage hierfür ist der vorliegende erste Entwurf sowie abzusehende konzeptionelle Anpassungen, ohne aber dem für 2019 avisierten zweiten Entwurf vorzugreifen. Im Fokus steht der Kontext-Begriff⁴, der einerseits dem Provenienzprinzip, dem leitenden methodischen Ansatz archivischer Erschließung, und andererseits dem Fachgebiet der Wissensrepräsentation (*knowledge representation*), hier vor allem den Methoden für die Deklaration logischer Aussagen (Graphen), entlehnt ist. Beide Aspekte sind für RiC-CM konstitutiv und als komplementär zu verstehen. Bereits an dieser Stelle soll ergänzend erwähnt werden, dass RiC-CM kein Datenmodell ist: Etablierte Erschließungsmethoden oder auch die vorherrschende Bereitstellung von Erschließungsergebnissen in Form von Findbüchern werden durch RiC-CM weder infrage gestellt noch ungültig. Gleichwohl wird auch gezeigt, dass RiC-CM nicht nur das Spektrum der Erschließung erfasst, sondern fraglos die Basis zur Datenmodellierung für Archivinformationssysteme sein soll. Es wird daher auch für ein neues Datenverständnis geworben, das die etablierte Erschließungsmethodik sublimiert. Nur eine neue, die Erschließung flankierende Methodik zur Erfassung und Analyse von Archivdaten kann, dann aber gewinnbringend, dazu beitragen, innovative digitale Dienste und Dienstleistungen zu entwickeln.

Die fortschreitende Digitalisierung von Aufgaben und Prozessen von Archivträgern und Archiv sowie die veränderten Verhaltensweisen und Erwartungen von Nutzerinnen und Nutzern sind ein treibender Faktor hinter RiC-CM und wohl auch der Grund für das Interesse an dieser Entwicklung. Exemplarisch werden abschließend drei Initiativen vorgestellt, die Potenziale für neue Datendienste untersuchen: *Social Network and Archival Context* (SNAC), *Pilote d'interopérabilité pour les Autorités Archivistiques françaises* (PIAAF) und *Interfaces to Data for Historical Social Network Analysis and Research* (SoNAR [IDH]). Sie sind unabhängig voneinander entstanden, gehen von verschiedenen Anfangspunkten aus und verfolgen unterschiedliche Ziele. Gemein ist ihnen, dass sie ein kooperatives Selbstverständnis mit strukturierten Daten und modernen Technologien für neue Dienste erproben.

Von Richtlinien zum Referenzmodell

Die digitale Transformation erfordert auch in Bereichen wie der Schriftgutverwaltung (Records Management) oder der archivischen Erschließung neue Konzepte. So erzwingt die Digitalisierung methodische Anpassungen bei der Organisation von Datenströmen unter Berücksichtigung von rechtlichen, prozessualen und technologischen Aspekten. Die Retrodigitalisierung oder die digitale Bereitstellung von Findbüchern sind dagegen nur noch Teilaspekte eines wesentlich umfangreicheren digitalen Transformationsprozesses. Nicht zuletzt vor diesem Hintergrund werden in RiC-CM hohe Erwartungen gesetzt; denn es kann ein zentrales Instrument für die Entwicklung von weitergehenden Konzepten und Lösungen für einzelne archivische Aufgabenbereiche in diesem Prozess sein. Dabei bleibt der Archivträger notwendigerweise der Bezugspunkt, doch wird die aktive Gestaltung etwa der einrichtungs- und auch der spartenübergreifenden Kuration und Vernetzung von Daten und Datendiensten bisherige Aufgaben wie Bewertung, Ordnung und Verzeichnung flankieren müssen.

Die Anforderungen, die an die Expertengruppe gestellt sind, sind vielschichtig und repräsentieren die fachliche und organisatorische Komplexität, die systematisch, sprach- und kulturübergreifend sowie konsensorientiert angegangen wird:

- Konsolidierung und Harmonisierung der internationalen Richtlinien der archivischen Erschließung,
- Berücksichtigung der zwei Bereiche des Archivwesens Records Management und Endarchiv,
- Verständigung über einen internationalen Konsens über Elemente archivischer Erschließung,
- Anschlussfähigkeit an angrenzende Domänen des kulturellen Erbes wie Museen und Bibliotheken,
- Erschaffung eines Bezugsrahmens zur Datenmodellierung in Archivinformationssystemen,
- Berücksichtigung von modernen Kommunikationstechnologien und semantischen Netzen,
- Aufnahme von Kritik (wissenschaftlicher) Archivnutzung an die traditionelle Erschließungspraxis.

Das auf diesen Anforderungen beruhende Konzeptmodell weist die Merkmale von Referenzmodellen auf, die in Bereichen wie den Informations- und Kommunikationstechnologien oder den neueren Steuerungs- und Governancemodellen in Wirtschaft und Verwaltung seit den 1980er-Jahren Konjunktur haben. Es sind Ontologien resp. Begriffssysteme, die branchen- oder aufgabenspezifische Klassen von Sachverhalten abbilden und konsequent Wissens- und Tätigkeitsbereiche erschließen.⁵ Mit Referenzmodellen kann es gelingen, institutionelles Erfahrungswissen über Objekte und Prozesse zu abstrahieren, um auf die zunehmende

4 Vgl. Thomas Karbe, *Conceptions and Context as a Fundament for the Representation of Knowledge Artifacts*, Berlin 2011, 98ff. [http://www.iiisci.org/Journal/CV\\$/sci/pdfs/CN11014.pdf](http://www.iiisci.org/Journal/CV$/sci/pdfs/CN11014.pdf).

5 Vgl. Klaus Peter Kratzer, *Neuronale Netze. Grundlagen und Anwendungen*, München 1990, S. 21.

Komplexität – wie sie beispielhaft durch die fortschreitende Digitalisierung wahrgenommen wird – zu reagieren. Sie sind so Bezugspunkte für fachliche Diskurse und offen für Adaption und Fortschreibung, sodass sie Instrumente für Veränderungsprozesse sein können.⁶

Das ausgearbeitete Referenzmodell Records in Context besteht aus zwei Teilen: das Konzeptmodell (RiC-CM) und eine Ontologie (RiC-O). Die Regelungsgegenstände und Beschreibungskategorien der bisherigen internationalen Richtlinien für Archivbestände, Aktenbildner, Funktionen und Archive werden in ein kohärentes Begriffssystem von Entitätentypen (entity types), Merkmalen (properties) und Beziehungen (relations) überführt. Erstmals werden, in Anlehnung an das Library Reference Model (LRM) für Bibliotheken oder CIDOC-CRM für Museen, die Komponenten der archivarischen Domain umfassend, aber notgedrungen auch nicht erschöpfend beschrieben. Die Formulierung von Anforderungen an Software und Schnittstellen sowie die Standardisierung von Daten und Prozessen gewinnen mit dem Referenzmodell einen konkreten, internationalen Bezugspunkt. Auf der Basis des Modells kann es unter anderem auch gelingen, Analogien und Differenzen von Methoden und Daten der einzelnen Kulturerbesparten strukturierter zu vergleichen.

Referenzmodell von Entitäten und Relationen

RiC-CM erfasst und beschreibt Entitätentypen, Merkmale und Beziehungen in einem Begriffssystem. In diesem werden primäre Entitätentypen in Anlehnung an die Regelungsgegenstände der bisherigen Richtlinien nach so genannten Top-Level-Entitätentypen gruppiert: record resources (Archivobjekte), agents (Akteure) und activity (Tätigkeitsbereiche).⁷ Sie sind Abstraktionen der Gemeinsamkeiten von spezifischeren Entitätentypen. Die folgende Liste zeigt eine Auswahl:

- Archivobjekte (record resources)
 - Bestände (record sets)
 - Verzeichnungseinheiten (records)
 - Einzelstücke (record components)
- Akteure (agents)
 - Personen (persons)
 - Gruppen (groups)
 - Körperschaften (corporate bodies)
 - Familien (families)
 - Software-Agenten (delegate agents)
 - Positionen in Gruppen (positions)
- Tätigkeitsbereiche (activities)
 - Aufgaben (functions)
 - Vorgänge (processes)
 - Ereignisse/Handlungen (actions)

Die Top-Level-Entitäten Archivobjekte und Akteure können als Klassen, ihre Sub-Level-Entitäten als Unterklassen betrachtet werden, das heißt, dass die Merkmale der Klassen auch für die Unterklassen gelten und Unterklassen durch zusätzliche Merkmale differenziert werden. Eine Ausnahme

bildet die Top-Level-Entität „Tätigkeitsbereiche“: Die Sub-Level-Entitäten Aufgaben und Vorgänge sind keine Unterklassen, sondern als je eine von zwei Seiten von Tätigkeitsbereich zu verstehen: Die Aufgabe begründet die Tätigkeit eines Akteurs und so die Existenz von Archivobjekten (Warum?). Ein Vorgang beschreibt dagegen anhand einzelner Ereignisse die Prozesse einer Tätigkeit, die durch Archivobjekte dokumentiert sind (Wie?).

Zusätzlich sind weitere, die Beschreibung primärer Entitäten unterstützende Entitätentypen definiert (supporting entities). Hierzu zählen vor allem die Entitätentypen:

- Rechtliche Grundlagen/Mandate (rules/mandate)
- Zeitpunkte/Laufzeiten (date)
- Orte/geographische Punkte (places)

Jeder Entitätentyp wird in RiC-CM definiert und beschrieben. Als Merkmale der Entitätentypen sind die Beschreibungskategorien zu verstehen, zu denen die Werte bzw. Ausprägungen einer Entität erhoben resp. erfasst werden können, z. B. für den Entitätentyp Verzeichnungseinheit (record) sind (nicht vollständig) folgende Merkmale in RiC-CM aufgeführt:

- global persistenter Identifier (global persistent identifier),
- Bezeichnung des Objekts, auch: Titel (name),
- Umfangsangabe (content extent),
- Sprache (language information),
- Inhaltsangabe (scope and content),
- Datenträger (medium).

Die entscheidende Neuerung in RiC-CM ist aber die Einführung von Relationen. Sie benennen die Art der Beziehung zwischen den Entitäten. Die Grafik in Abbildung 1 deutet diese Beziehungen an. Bisher war es die Regel, dass z. B. für diverse Zeiten unterschiedliche Beschreibungskategorien angelegt wurden: das Entstehungsdatum einer Urkunde, die Laufzeit einer Akte, das Gründungsdatum einer Behörde, die Geburts- und Sterbedaten einer Person, Zugangs- oder Übernahmedatum von Archivgut oder ein Enddatum für die Aufbewahrungs- oder Sperrfrist.

Durch die Abstraktion von Entitäten und Relationen können nun semantisch logische und maschinell auswertbare Aussagen nach dem Schema Subjekt – Prädikat – Objekt getroffen werden:

- date – was creation date of – record, z. B. 1948 ist das Entstehungsdatum der Akte,

⁶ Vgl. Rainer Koch, Strategischer Wandel des Managements öffentlicher Dienste. Design-orientierte Managementlehre und Modernisierung öffentlicher Dienste, Wiesbaden 2008, 32 ff.

⁷ Die Termini sind Gegenstand der laufenden Arbeit am zweiten Entwurf von RiC-CM und unterliegen Modifikationen. Sie sind in einem Compendium, das den Sachstand der Arbeit der Expertengruppe dokumentiert, veröffentlicht (vgl. EGAD [Hrsg.], Records in Context – Compendium. <https://web.esrc.unimelb.edu.au/ICAD/>).

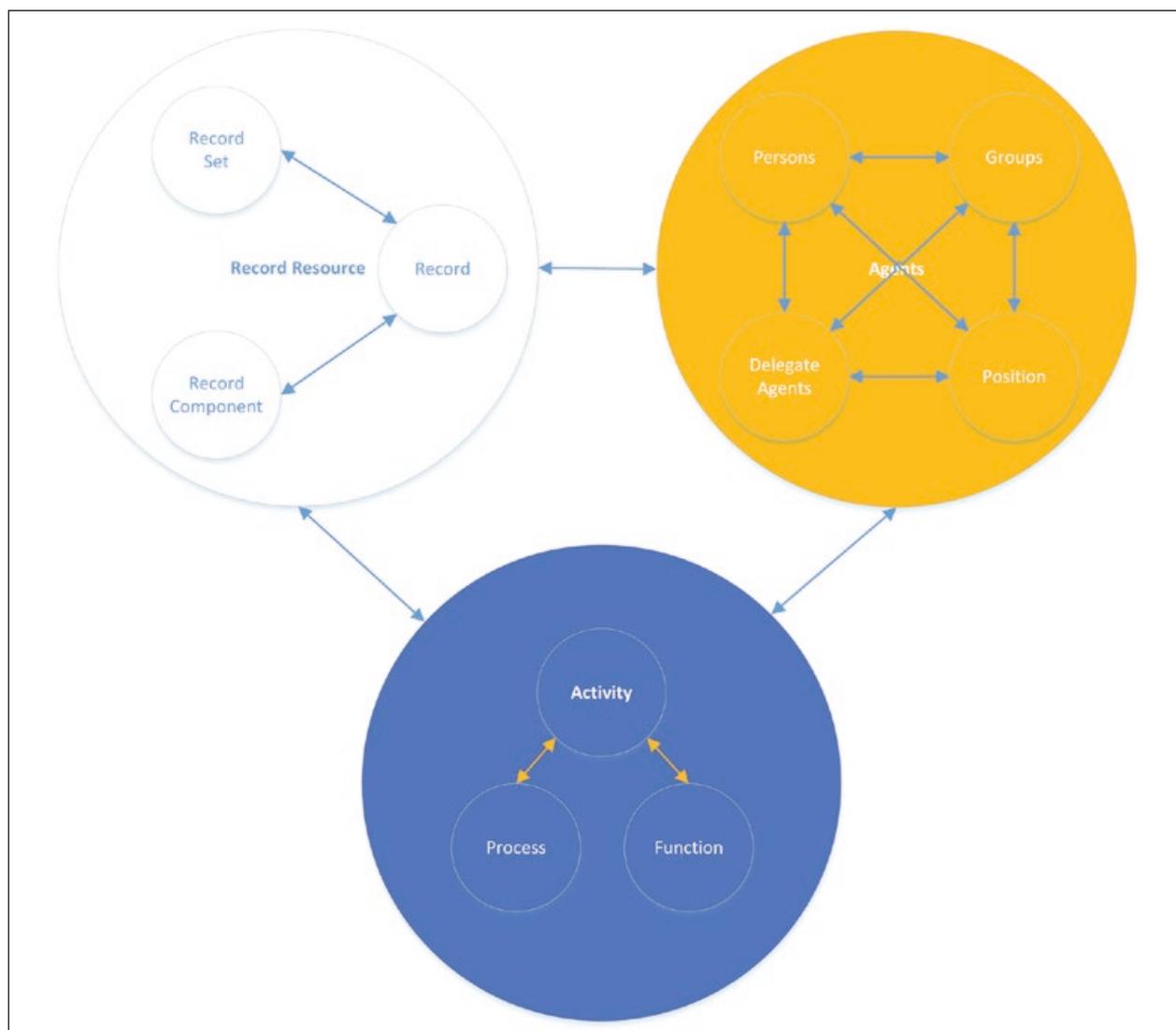


Abb. 1: Auswahl der Entitätentypen und Relationen zwischen Top-Level- und Sub-Level-Entitäten

- date – was start date of – agent (group), z. B. 1950 ist das Gründungsdatum des Suhrkamp Verlags,
- date – was birth date of – agent (person), z. B. 1906 ist das Geburtsdatum von Hannah Arendt.

Diese Form expliziter Aussagen über die Beziehung zwischen Entitäten unterstützt die Entwicklung des „Semantic Web“, und es erleichtert die Bereitstellung, den Austausch und die Nachnutzung von Daten durch standardisierte Formate und Schnittstellen sowohl zwischen Archiven und innerhalb der Domain der Kulturerbeinstitutionen als auch in sonstigen, noch unbekanntem Szenarien.

Die Liste der Relationen bzw. der Aussagen über die Art der Beziehungen zwischen Entitäten sind mit „RiC-O“ im Format der Web Ontology Language (OWL) formalisiert. OWL beruht auf der RDF-Syntax⁸, die von Software verarbeitet werden kann. Sie ist zudem für formalisierte Aussagen in verschiedenen natürlichen Sprachen offen. Für Archivobjekte (record resources) können über formale Aussagen beispielsweise die hierarchischen Strukturen eines Findbuchs zum Ausdruck gebracht werden, z. B. record –

is part of – record set oder auch: record – has part – record component. Diese maschinell interpretierbaren Aussagen sind ein Äquivalent für konventionelle Findbücher. Die Beziehungen zwischen Akteuren können, wie Abbildung 1 zeigt, horizontal und vertikal sein. Die Entitätentypen Aufgaben (functions) und Vorgänge (processes) konstituieren gemeinsam den Entitätentyp Tätigkeitsbereiche (activities).

Das Referenzmodell berücksichtigt sowohl die historische Perspektive, das sind die Archivobjekte und Überlieferungsbildner der Vergangenheit, als auch die Perspektive der Gegenwart, das sind die heute aktiven Akteure im Aufgabenbereich eines Archivs. Das Archiv, das selbst ein beschreibbarer Akteur nach RiC-CM ist, bildet die Brücke zwischen diesen beiden Perspektiven.

Datenmodellierung und Datendienste

Mit Records in Context geht die EGAD nicht nur terminologisch und methodisch einen neuen Weg. Das Proveni-

⁸ RDF steht für Resource Description Framework, einem Datenmodell und Basis des Semantic Web (s. hierzu auch den Beitrag von Ed de Heer in diesem Heft).

enzprinzip wird um ein neues Kontextverständnis erweitert. Dies deutet der Wechsel von Mehrebenenerschließung, die den ISAD(G) als Ableitung des Provenienzprinzips zugrunde liegt, hin zu einer multidimensionalen Erschließungsperspektive an. Diese schließen sich aber nicht aus: „Le système de description archivistique doit être en mesure de représenter, documenter et préserver les relations entre les diverses entités.“⁹ Der Entstehungs- und Benutzungszusammenhang der durch das Archiv überlieferten Archivobjekte, ist für die Analyse und das Verstehen historischer Zeugnisse entscheidend. Kontextinformationen sichern die Integrität und Authentizität der Archivobjekte. Der Erhalt dieser Informationen auch für digitale Daten ist eine Kernaufgabe von Archiven.

Das Provenienzprinzip ist bisher Methode und zugleich auch Maßnahme, um die Ursprünglichkeit der Zeugnisse zu erhalten: einmal durch die Aufstellung der physischen Bestände im Archiv (Kontext der Überlieferung), einmal durch die logische Ordnung der Verzeichnungseinheiten im Findbuch (Kontext von Entstehung und Nutzung der Archivobjekte). Zur Einführung des Provenienzprinzips konstatierte der damalige Direktor des Geheimen Staatsarchivs (GStA), Paul Bailleu:

„Es kamen in das Archiv z. B. die Akten der preußischen Gesandtschaft in Petersburg, unter ihnen etwa Akten über die Verhandlungen mit Rußland über die polnischen Theilungen. Der Archivbeamte, der die Registratur der Petersburger Gesandtschaft übernimmt, erinnert sich, daß er ja schon Akten über diese Ereignisse in XI, 175 besitzt, und er stopft nun die betreffenden Stücke ... in diese Nummer hinein und vertheilt die anderen Stücke an entsprechenden Stellen des Archivs, wo sie ihrem Inhalte nach Anschluß zu finden scheinen. Er sieht sich dann im Archiv um und findet unter den Nachlässen in den Papieren des Prinzen Heinrich, Bruders Friedrichs des Großen, gleichfalls Korrespondenzen zur Geschichte der polnischen Theilungen: flugs steckt er auch diese in die Packete XI, 175.“¹⁰

Die Neusystematisierung des GStA „nach den Registraturen, nach den Behörden, wie sie im Laufe der Geschichte entstanden, gewachsen und untergegangen sind, entspricht in gleichem Maße unserm historischen Denken“ (ebd.). Neben der Sicherung des Kontexts war es aber die Not der Archive im 19. Jahrhundert, mit dem stark anwachsenden Schriftgut zunehmend modernisierter Verwaltungen umzugehen, die zur breiten Einführung des Provenienzprinzips beitrug¹¹.

Eine ähnliche Modernisierung stellt die Digitalisierung der Verwaltungen und Unternehmen, des öffentlichen und privaten Lebens dar. Sie setzt Archive angesichts legislativer und exekutiver Vorgaben für den Einsatz z. B. von Dokument-Management-Systemen unter Handlungsdruck¹². Records in Context überführt das Provenienzprinzip, das in Archivteknik und Findbuch Ausdruck findet, in ein

graph-basiertes semantisches Netz von Entitäten und Relationen¹³. So wird nicht mehr nur der engere Kontext von Entstehung und Nutzung der Quellen (Bestand und Bestandsbildner), sondern der soziale und rechtliche Kontext in die Erschließung einbezogen. Dies ist nicht neu¹⁴; die einleitenden Erläuterungen eines Findbuchs beziehen idealerweise Aussagen über Tätigkeitsbereiche und soziale Beziehungen der Bestandsbildner ein. Jedoch leitet RiC-CM zu einem neuen Ansatz der Datenmodellierung über: Informationen über Entitäten werden nicht deskriptiv als Texte, sondern als Aussagen über Entitäten explizit in Metadaten mit eindeutigen Identifikatoren erfasst¹⁵, sodass die Aussagen beispielsweise anwendungs- und bestandsübergreifend vernetzt werden können.

Die Orientierung der Erschließung an logischen Aussagen, Graphen¹⁶, kann darüber hinaus eine neue Erschließungsdynamik begründen. Sie muss nicht mehr zwingend das abgeschlossene, publizierte und nicht mehr veränderte Findbuch zum Ziel haben. Vielmehr erlaubt der graphorientierte Ansatz eine stärker fall- und nutzerorientierte Erschließung der Bestände, Akteure und Tätigkeitsbereiche. Sie unterstützen auch das Erproben von neueren Ansätzen wie Crowdsourcing, die etwa für Citizen Science charakteristisch sind.¹⁷ Das Modell von RiC-CM unterstützt das Bilden von aussagekräftigen Metadaten etwa für Datenbanken und Dokument-Management-Systeme, die direkt in Archivinformationssysteme übernommen werden können. Graphen können ebenfalls die Erstellung von dynamischen Organigrammen in Verbindung mit Aufgaben und rechtlichen Grundlagen fördern. Auch diese Daten können für Endarchive direkt übernommen werden. Das Findbuch wird so auch nicht mehr das ausschließliche Ziel der archivistischen Erschließung, sondern je nach Anforderung eine mögliche Form zur Repräsentation von logischen Aussagen über Entitäten und ihre Beziehungen.

9 Nils Brübach/Robert Nahuët/Claire Sibille-de Grimouard, Une évolution dans les pratiques descriptives. Vers un modèle conceptuel archivistique?, in: *arbidio* 14 (2012), 2. 4–8.

10 Paul Bailleu, Das Provenienzprinzip und dessen Anwendung im Geheimen Staatsarchiv, in: *Die Blätter für deutsche Landesgeschichte* 10/11 (1902), S. 193–195.

11 Peter Müller, Vollregist, Findbuch oder Informationssystem. Anmerkungen zu Geschichte und Perspektiven der archivistischen Erschließung, in: *Der Archivar* 58 (2005), S. 6–15.

12 Christoph Schmidt, Aussonderung und Archivierung von elektronischen Akten im Landesarchiv NRW, in: VdA (Hrsg.), *Kompetent. Archive in der Wissensgesellschaft*. 86. Deutscher Archivtag in Koblenz (Tagungsdokumentation zum Deutschen Archivtag 21), Fulda 2018, S. 79–84.

13 Vgl. RiC-CM 2016, 3 ff.

14 Kathrine M. Wisser, Investigating the Small World of literary archival collections. The impact of EAC-CPF on archival descriptive practices. (Relationships, description, and the archival community. 1), in: *The Journal of Contemporary Archival Studies* Vol. 2 (2015), <http://elischolar.library.yale.edu/jcas/vol2/iss1/1>.

15 Vgl. Daniel Pitti/Gavan McCarthy/Bogdan-Florin Popovici, Records in Context. An Archival Description Draft Standard, Mexiko 2017, Präsentation, <http://www.alaarchivos.org/wp-content/uploads/2018/01/1.-Daniel-V.-Pitti.pdf>.

16 Vgl. RiC-CM 2016, 9.

17 Thekla Kluttig, Die Citizen Science Strategie 2020 für Deutschland und die Archive, in: VdA (Hrsg.), *Kompetent. Archive in der Wissensgesellschaft*. 86. Deutscher Archivtag in Koblenz (Tagungsdokumentation zum Deutschen Archivtag 21), Fulda 2018, S. 33–41.

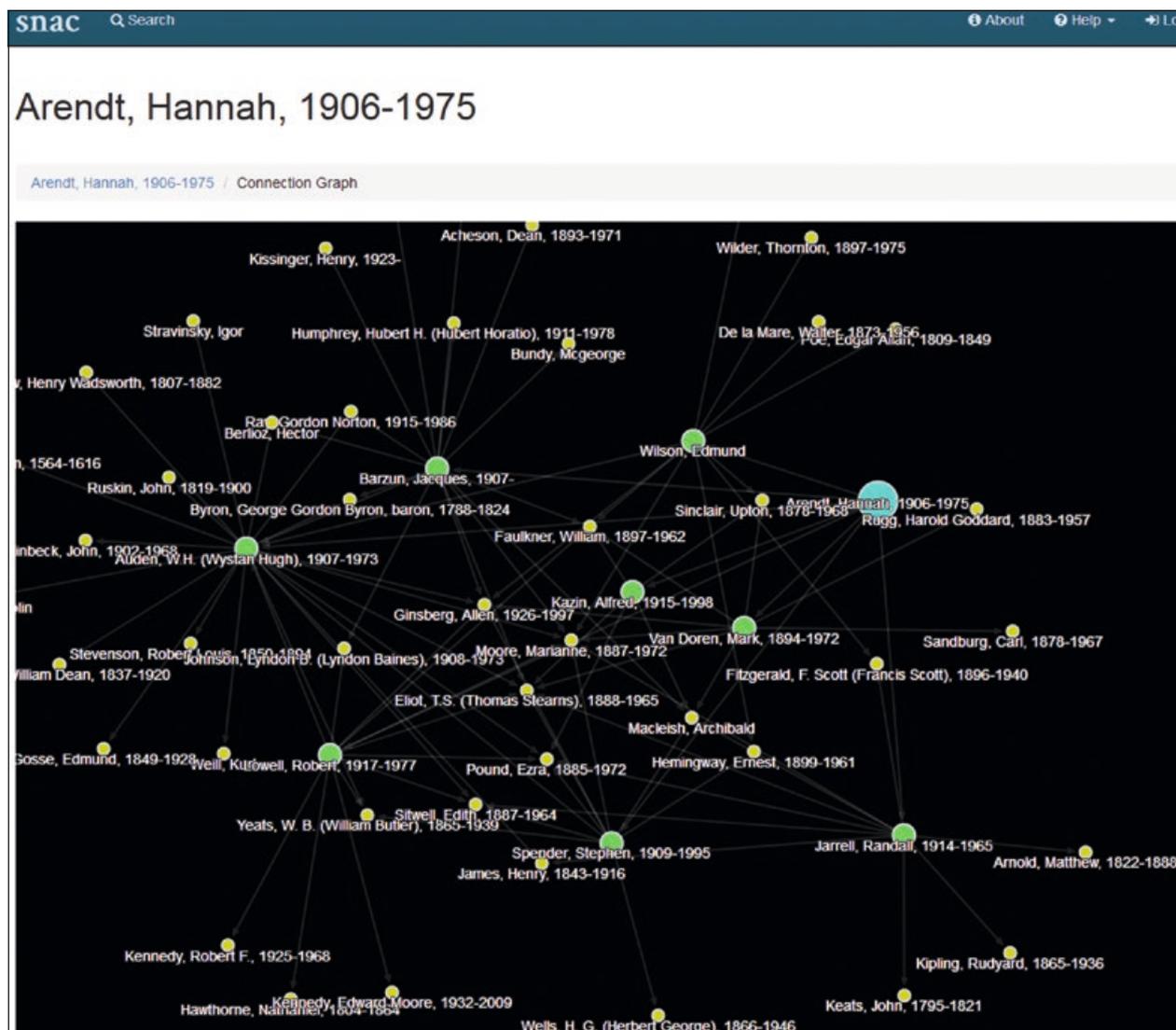


Abb. 2: Netzwerk von Hannah Arendt im SNAC-Datenbestand

Die Einführung einer differenzierten Datenerfassung entlang des Entitäten-Relationen-Modells von RiC-CM kann die Entwicklung neuer Dienstleistungen und Datendienste stimulieren. Die Nutzung von Normdaten ist ein erster, wie von Karsten Uhde in diesem Heft dargelegt, schon jetzt geforderter Ansatz. Die Vorteile von Normdaten in der Erschließung sind: Eindeutigkeit der Entitäten, Konsistenz der Orthographie, Persistenz der Datensätze, Konformität mit einschlägigen Regeln sowie Format- und Schnittstellenstandards, Nachnutzbarkeit und letztlich Vernetzung von Daten und Diensten.

Mit dem sukzessiven Aufbau von Social Network and Archival Context, SNAC¹⁸, unterstützt von der National Endowment for the Humanities und der A. Mellon Foundation, ist seit 2010 eine Normdatei für Personen, Körperschaften und Familien im Bereich archivischer Erschließung entstanden. Durch SNAC wurden im ersten Schritt aus einer großen Anzahl von Findbüchern Personen-, Körperschafts- und Familiennamen mittels Algorithmen als Entitäten identifiziert, disambiguiert und durch Abgleich mit der aggregierenden internationalen Normdatei VIAF verlinkt. In

der aktuellen Projektphase wird SNAC als ein produktiver Dienst mit Möglichkeiten zur Erfassung und redaktionellen Betreuung von Daten über Personen, Körperschaften und Familien implementiert¹⁹ und für das z. B. die National Archives and Records Administration ein Qualifizierungskonzept entwickeln. Mit SNAC stehen aber nicht nur Normdatensätze für die Erschließung zur Verfügung, sondern auch die Art von Beziehungen sowohl zwischen Akteuren (s. Abbildung 2) als auch zwischen Akteuren und Archivbeständen.

Einen alternativen, aber flankierenden Ansatz verfolgen das Archives nationales in enger Kooperation mit der Bibliothèque Nationale de France und dem Service interministériel des Archives de France im Projekt Pilote d'interopérabilité pour les Autorités Archivistiques françaises (PIAAF)²⁰. Das Ziel ist es, die Anwendung von Records in Context an konkreten Daten zu erproben. Es umfasst Fragen an die Struk-

18 <http://snaccooperative.org/>.

19 Daniel Pitti et al., Social Networks and Archival Context. From Project to Cooperative Archival Program, in: *Journal of Archival Organization* 12 (2015), H. 1–2, S. 77–97.

20 <http://piaaf.demo.logilab.fr/>.

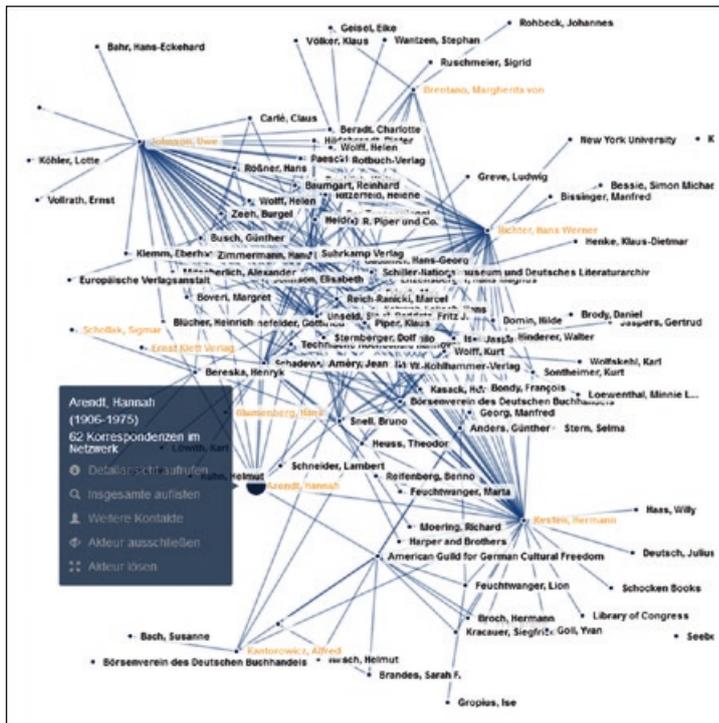


Abb. 3: Netzwerkgraph zu Hannah Arendt (Kalliope-Verbund)

turierung von Erschließungsdaten und Konzepte ihrer Präsentation. Dies wird beispielhaft am Archivbestand des Ministère de l'Éducation nationale erprobt und eindrucksvoll demonstriert. Die Erschließungsdaten liegen im Format RDF vor. Die Entitäten sind eindeutig in einzelnen Datensätzen erfasst und mittels persistenten Identnummern adressierbar. Die Akteure im Kontext des Bestands sind mit einem ISNI (International Standard Name Identifier, ISO 27729) ausgezeichnet und so mit ergänzenden Datenangeboten vernetzbar. Die Daten des Bestandes werden visuell und textlich für explorative und klassische Recherchen angeboten²¹. Die Demonstration des Prototyps verzichtet auf die Darstellung der Erschließungsdaten in Form eines Findbuchs. Stattdessen zielt der Ansatz auf die in RiC-CM angelegte multidimensionale Präsentation. Der Einstieg in das Informationsangebot gelingt über diverse Entitäten (Körperschaften, Tätigkeitsbereiche, Positionen, Personen oder Bestände). Das Potenzial des Ansatzes etwa für die Suche auch in größeren Datenmengen wird durch die Demonstration erlebbar.

Einen dritten, zu PIAAF und SNAC ebenfalls komplexen Ansatz wird das Projekt SoNAR (IDH), Interfaces to Data for Historical Social Network Analysis and Research, verfolgen. Dieses Proof-of-Concept-Projekt, das die Deutsche Forschungsgemeinschaft ab Mitte 2019 fördern wird, untersucht den Aufbau einer Forschungstechnologie zur Analyse historischer sozialer Netzwerke. Ausgangspunkt sind bereits strukturierte Metadaten der Zeitschriftendatenbank, des Kalliope-Verbunds und der Gemeinsamen Normdatei sowie Volltexte von Zeitungen (OCR) und Briefen (Transkriptionen). Im Projektverlauf werden Softwarekomponenten für eine Prozesskette erprobt, um Aussagen über Beziehungen zwischen Entitäten in den Ausgangsdaten in

Graphen zu transformieren sowie, in einem zweiten Schritt, diese zu visualisieren und mit statistischen und netzwerkanalytischen Maßzahlen zu beschreiben. Die Projektpartner evaluieren im Projektverlauf eine Reihe von Faktoren, die auf die Leistungsfähigkeit einer Forschungstechnologie für die Analyse von historischen sozialen Netzwerken wirken. Hierzu zählen etwa der Umgang mit fehlenden oder fehlerhaften Daten oder die Grenzen der Interaktion mit multimodalen Netzwerken.

Zusammenfassung

RiC-CM ist ein Referenzmodell. Entlang von Entitätentypen, Merkmalen und Beziehungen werden der Blickwinkel und die Komponenten der Erschließung von Archivquellen modellhaft erfasst. Dadurch gewinnen die Elemente der bisherigen internationalen Standards Kohärenz. Das Modell reflektiert einen internationalen Konsens und auch einen Konsens zwischen Records Managern und Archivaren. Ein Ergebnis kann sein, dass Datenangebote künftig nicht nur geteilt, sondern im Zusammenhang mit der Erschließung von Ressourcen des kulturellen Erbes auch intensiver nachgenutzt werden können. Durch die Datennachnutzung, das ist vor allem die eindeutige Referenzierung auf Beschreibungen gleicher Entitäten, entsteht ein überregional vernetztes und, z. B. mit dem Angebot Deutsche Digitale Bibliothek/Archivportal D, vernetzbares Informationsangebot. Überregionale Informationsangebote profitieren erheblich von Anpassungen der Datenerfassung entlang des

²¹ Florence Clavaud/Anila Angjeli/Stéphanie Roussel, *Réprésenter en RDF, interconnecter et visualiser en graphe des jeux de métadonnées archivistiques de provenances multiples. Un projet de prototype*, in: *Gazette des archives* 245 (2017), S. 157–172.

Entitäten-Relationen-Modells durch Eindeutigkeit von Entitäten, etwa durch die Referenz auf Normdaten der GND.²² Die Chancen, die eine Orientierung der Gestaltung und Entwicklung von Diensten und Dienstleistungen an RiCCM haben kann, wird durch die beispielhaft genannten Initiativen oder auch den Beiträgen von Karsten Uhde und Ed de Heer in diesem Heft deutlich. Es wird aber auch deutlich, dass die digitale Dividende, die Archive aus der digitalen Transformation gewinnen können, kaum vorzustellen ist, wenn die Voraussetzungen durch einzelne Archive geleistet werden sollten. Die notwendigen Investitionen für die Konzeption, Planung und Entwicklung technischer Umgebungen setzen sehr wahrscheinlich eine engere, verbundorientierte regionale und überregionale Zusammenarbeit und Spezialisierung voraus. Die Hürden sind vorhanden. Den-

noch zeigen beispielsweise Hessen und Niedersachsen erfolgreiche Kooperationen bei der Entwicklung von Archivinformationssystemen. Als Referenzmodell bietet Records in Context einen geeigneten Bezugspunkt für sehr erfolgreiche Archivdienste. ■

Gerhard Müller
Staatsbibliothek zu Berlin – Preußischer Kulturbesitz
gerhard.mueller@sbb.spk-berlin.de

Silke Jagodzinski
Bundesarchiv, Berlin
s.jagodzinski@bundesarchiv.de

²² Wolfgang Krauth, Archive und Online-Portale. Thesen für den weiteren Erfolg, in: Der Archivar 68 (2015), H. 1, S. 6–9.