Sarah Hartmann und Heinz Pampel

GND und ORCID: Brückenschlag zwischen zwei Systemen zur Autorentidentifikation

GND and ORCID: Building a bridge between two systems for author identification

DOI 10.1515/bd-2017-0062

Zusammenfassung: Die Uneindeutigkeit von Personennamen erschwert im Rahmen der bibliothekarischen Erschließung die eindeutige Zuordnung von Autorinnen und Autoren zu ihren Werken. Bibliotheken im deutschsprachigen Raum adressieren das Problem der Mehrdeutigkeit von Namen durch den Einsatz der Gemeinsamen Normdatei (GND). Die internationale Initiative ORCID (Open Researcher and Contributor ID) verfolgt das gleiche Ziel. Akteur ist hier jedoch die einzelne Wissenschaftlerin oder der einzelne Wissenschaftler. Das Projekt „ORCID DE – Förderung der Open Researcher and Contributor ID in Deutschland“¹ hat sich, dank der Förderung der Deutschen Forschungsgemeinschaft (DFG), unter anderem zum Ziel gesetzt, einen Brückenschlag zwischen den beiden Systemen – GND und ORCID – zu schaffen, um damit die Datenqualität beider Systeme wechselseitig zu erhöhen. Der vorliegende Beitrag gibt einen Überblick über ORCID und das Projekt ORCID DE. Dabei wird insbesondere auf die angestrebte Verlinkung von GND und ORCID eingegangen.

Schlüsselwörter: ORCID, Autorentidentifikation, Gemeinsame Normdatei

Abstract: In the process of library indexing, the ambiguity of personal names makes the clear relation of authors to their works difficult. Libraries in the German-speaking area address the problem of ambiguity of names by applying a commonly used authority file (Gemeinsame Normdatei, GND). The international initiative ORCID (Open Researcher and Contributor ID) has the same object, but with the individual scientist acting. Thanks to the promotion of the German

1 <http://www.orcid-de.org/> [Zugriff: 18.04.2017].

Research Foundation (DFG), the project „ORCID DE – Promotion of the Open Researcher and Contributor ID in Germany“ aims, among other things, to build a bridge between the two systems – GND and ORCID – in order to increase mutually the data quality of both systems. This article provides a survey of ORCID and the project ORCID DE, focussing on the aspired linking of GND and ORCID.

Keywords: ORCID, author identification, authority file

1 Einleitung

Die Uneindeutigkeit von Personennamen erschwert im Rahmen der bibliothekarischen Erschließung die eindeutige Zuordnung von Autorinnen und Autoren zu ihren Werken. Die Herausforderung der eindeutigen Identifikation und der damit verbundenen Disambiguierung von Personennamen tritt allerdings nicht nur in Bibliotheken auf, sondern tangiert viele Wissenschaftlerinnen und Wissenschaftler, deren akademische Leistung an ihren Veröffentlichungen gemessen wird. Abhilfe schaffen Identifikationssysteme, die eine eindeutige und dauerhafte Zuordnung zwischen Personen und ihren Werken ermöglichen.

Bibliotheken im deutschsprachigen Raum adressieren das Problem der Mehrdeutigkeit von Namen durch den Einsatz der Gemeinsamen Normdatei² (GND): Einer Person wird ein eindeutiger Identifikator (ID) zugewiesen, indem ein Normdatensatz in strukturierter Form nach festgelegten Regeln angelegt wird. Dafür werden öffentlich zugängliche Informationen (z. B. aus vorliegenden Publikationen oder Webseiten) ausgewertet. Die jeweiligen Publikationen, an denen die Person mitgewirkt hat oder die über diese Person veröffentlicht wurden, werden mit den GND-Einträgen in Beziehung gesetzt.

Die internationale Initiative ORCID (Open Researcher and Contributor ID) verfolgt das gleiche Ziel, jedoch mit einem anderen Ansatz: Hier sind die Wissenschaftlerinnen und Wissenschaftler selbst für die Pflege ihrer eigenen Profile, in denen Daten und Informationen zu einer Person und den mit dieser Person verbundenen Werken hinterlegt werden können, verantwortlich.

Das Projekt „ORCID DE – Förderung der Open Researcher and Contributor ID in Deutschland“³ hat sich unter anderem zum Ziel gesetzt, einen Brückenschlag

2 <http://www.dnb.de/gnd> [Zugriff: 18.04.2017].

3 <http://www.orcid-de.org/> [Zugriff: 18.04.2017].

zwischen den beiden Systemen – GND und ORCID – zu schaffen, um damit die Datenqualität beider Systeme wechselseitig zu erhöhen.

Im Folgenden wird ein Überblick über ORCID und das Projekt ORCID DE gegeben. Dabei wird insbesondere auf die Verlinkung von GND und ORCID eingegangen.

2 ORCID

Der ORCID-Leitsatz lautet: „enter once – reuse often“⁴. Die einmal im ORCID-Profil hinterlegten Daten und Informationen können über die ORCID ID mit sämtlichen wissenschaftlichen Beiträgen einer Person – wie z. B. Textpublikationen, Forschungsdaten oder auch Softwareprodukten – verknüpft werden. Mittels Programmierschnittstellen kann die Aktualisierung und der Austausch der hinterlegten Daten zentral über das ORCID-Profil organisiert werden. Dabei unterstützt ORCID den gesamten Ablauf des Publikationsmanagements: Wird die ORCID ID zum Beispiel bei der Einreichung eines Zeitschriftenartikels angegeben, kann – vorausgesetzt, der Artikel wird von dem Journal angenommen – eine automatische Übertragung der bibliografischen Daten in das ORCID-Profil des Publizierenden selbst sowie z. B. in Publikationsdatenbanken von wissenschaftlichen Institutionen oder Förderorganisationen umgesetzt werden.⁵

Die Idee zu einem internationalen, disziplinübergreifenden und offenen Identifier zur eindeutigen Identifikation von Wissenschaftlerinnen und Wissenschaftlern und zur Verwaltung eines digitalen Lebenslaufes ist 2009 entstanden.⁶ Seit 2012 wird ORCID als Service angeboten.⁷ ORCID ist im Gegensatz zu bereits länger bestehenden wissenschaftlichen Identifikatoren als offenes System konzipiert.⁸ Jede Person kann sich in dem ORCID-Verzeichnis kostenlos anmelden und Daten und Informationen pflegen. Dafür ist lediglich die Angabe des Vornamens,

4 <https://orcid.org/content/collect-connect> [Zugriff: 18.04.2017].

5 <https://orcid.org/blog/2016/06/24/auto-update-has-arrived-orcid-records-move-next-level> [Zugriff: 18.04.2017].

6 Credit where credit is due. (2009). *Nature*, 462(7275), S. 825–825. <http://doi.org/10.1038/462825a> [Zugriff: 18.04.2017].

7 Butler, D. (2012): Scientists: your number is up. *Nature*, 485(7400), S. 564–564. <http://doi.org/10.1038/485564a> [Zugriff: 18.04.2017].

8 Haak, L. L.; Fenner, M.; Paglione, L.; Pentz, E.; Ratner, H. (2012): ORCID: a system to uniquely identify researchers. *Learned Publishing*, 25(4), S. 259–264. <http://doi.org/10.1087/20120404> [Zugriff: 18.04.2017]. Sowie: Fenner, M. (2011): Author Identifier Overview. *LIBREAS. Library Ideas*, (18), 24–29. <https://nbn-resolving.org/urn:nbn:de:kobv:11-00183866> [Zugriff: 18.04.2017].

einer E-Mailadresse sowie eines Passwortes notwendig. Optional können weitere identifizierende Merkmale, wie alternative Namensformen, Webseite, Biografie, Förderinformationen (z. B. Auszeichnungen und Projekte) sowie assoziierte Länder und Institutionen (z. B. Angaben zur Ausbildung und Arbeitgeber) angegeben werden.

Abb. 1: Screenshot des ORCID-Profiles <http://orcid.org/0000-0003-3334-2771>.

Die Pflege der Publikationsliste in ORCID erfolgt wahlweise durch die automatisierte Verknüpfung mit anderen Informationssystemen, durch den Import von Metadaten als BibTeX-Datei oder durch die manuelle Eingabe der bibliografischen Information. Erstgenannte Option steht dank der sogenannten „Search&Link“-Funktionalität beispielsweise für die DOI-Registrierungsdienste Crossref⁹ und

⁹ <https://www.crossref.org/> [Zugriff: 18.04.2017].

DataCite¹⁰ sowie für die Datenbank Scopus¹¹ und das Portal Web of Science¹² zur Verfügung.

Der Datenaustausch erfolgt über die ORCID-Programmierschnittstellen, hierzu muss die Person eines ORCID-Profiles einem Informationssystem – wie z. B. der Publikationsdatenbank ihrer Institution – Rechte zur Nutzung ihres Profils zuweisen. Hintergrund ist die bereits erwähnte Hoheit einer Person über ihre eigenen im ORCID-Profil gespeicherten Daten. Die ORCID-Nutzerin oder der ORCID-Nutzer entscheidet immer selbst darüber, ob und für wen die gespeicherten Informationen zugänglich sind.¹³ Anhand eines Ampel-Systems kann zwischen drei Einstellungen unterschieden werden: a) Öffentlichkeit („public“), b) definierte Dritte („trusted parties“) oder c) privat („private“).

Die öffentlich freigegebenen Daten stehen jedem kostenfrei zur Nachnutzung zur Verfügung. Die Daten werden einmal im Jahr von ORCID in Form eines Datendumps unter der Creative-Commons-Deed CC0¹⁴ im JSON- und XML-Format bereitgestellt und sind über die öffentlich frei zugängliche REST-Schnittstelle¹⁵ „Public Open API“ abfragbar.

ORCID ermöglicht auch die Verlinkung zu weiteren Identifikatoren, wie z. B. Scopus Author Identifier, ResearcherID oder ISNI (International Standard Name Identifier) über die sogenannte „Search&Link“-Funktionalität.

Mit Stand März 2017 waren weltweit ungefähr 3,3 Mio. ORCID IDs registriert, davon sind circa 670.000 Einträge mit mindestens einem Werk (Textpublikation etc.) verknüpft.¹⁶ Innerhalb des Jahres 2016 sind eine Mio. neu registrierte Personen hinzugekommen¹⁷ – die Zahl der ORCID-Profile steigt stetig an und zeigt die wachsende Bedeutung von ORCID als Standard im Bereich des wissenschaftlichen Publikationsmanagements.

¹⁰ <https://www.datacite.org/> [Zugriff: 18.04.2017].

¹¹ <https://www.scopus.com/> [Zugriff: 18.04.2017].

¹² <https://webofknowledge.com/> [Zugriff: 18.04.2017].

¹³ <https://orcid.org/about/what-is-orcid/our-principles> [Zugriff: 18.04.2017].

¹⁴ <https://orcid.org/content/download-file> [Zugriff: 18.04.2017].

¹⁵ https://de.wikipedia.org/wiki/Representational_State_Transfer [Zugriff: 18.04.2017].

¹⁶ ORCID statistics. <https://orcid.org/statistics> [Zugriff: 18.04.2017].

¹⁷ Meadows, Alice; Brown, Josh; Haak, Laurel; Paglione, Laura; Peters, Robert; Wright, Douglas (2017): ORCID Annual Report 2016.pdf. figshare. <https://doi.org/10.6084/m9.figshare.4810213.v1>, S. 10 [Zugriff: 18.04.2017].

3 ORCID Inc.

Betrieben wird ORCID von einer internationalen gemeinnützigen Organisation mit Sitz in den USA, die sich durch Mitgliedsbeiträge und zusätzliche Drittmittel finanziert. Zu den 654 institutionellen Mitgliedern¹⁸ von ORCID Inc. zählen wissenschaftliche Einrichtungen, Verlage, Bibliotheken und andere Akteure im Bereich der Wissenschaft. Die Mitgliedsgebühr¹⁹ ist gestaffelt und abhängig von der jeweiligen Art der Nutzung von ORCID. Beispielsweise kann eine Premium-Mitgliedschaft erforderlich sein, wenn auf die ORCID-Schnittstelle mit mehreren Systemen einer Institution zugegriffen werden soll.

Als institutionelles Mitglied bekommt eine Institution zusätzlich zur öffentlich bereitgestellten Schnittstelle einen Zugriff auf eine Schnittstelle, die lesen- und schreibenden Zugriff auf die ORCID-Profile der eigenen Wissenschaftler erlaubt („trusted party“) – sofern ein individueller Nutzer diese Option freigegeben hat. Des Weiteren können sich dank dieser Schnittstelle Angehörige der eigenen Institution aus der eigenen Anwendung (z. B. Redaktions- oder Publikationssystem) heraus bei ORCID registrieren, Metadaten ergänzen bzw. aktualisieren. Darüber hinaus besteht für Mitgliederinstitutionen die Möglichkeit, den technischen Support von ORCID in Anspruch zu nehmen und in den ORCID-Gremien aktiv zu werden.²⁰

Zur Community-Arbeit von ORCID zählt, dass spezifische Themen in Arbeitsgruppen²¹ behandelt werden. Die Arbeitsgruppen behandeln aktuell beispielsweise Themen wie die Ausarbeitung von Best-Practice-Leitfäden zur Integration von ORCID in den Publikationsworkflow von Büchern oder die Nutzung von Identifikatoren für Organisationen.²²

¹⁸ <https://orcid.org/members> [Zugriff: 18.04.2017].

¹⁹ <https://orcid.org/about/membership> [Zugriff: 18.04.2017].

²⁰ Wie Anm. 19.

²¹ <https://orcid.org/about/community> [Zugriff: 18.04.2017].

²² Der Bedarf für eindeutige Identifikatoren für Organisationen entstand aus dem Wunsch heraus, im ORCID-Profil eine Affiliation nicht nur als Namen (Freitext) angeben zu können, sondern eindeutig zu einer Organisation zu referenzieren. Siehe hierzu auch: <http://orcid.org/blog/2017/03/30/oi-project-gets-underway-planning-open-organization-identifier-registry> [Zugriff: 18.04.2017].

4 Verbreitung von ORCID

Die Zahl der wissenschaftlichen Verlage, die ORCID in ihre Editorial-Management-Systeme und damit auch in die Metadaten von wissenschaftlichen Veröffentlichungen integrieren, steigt. Mit Stand März 2017 hatten 27 Verlage eine Absichtserklärung zur Integration von ORCID in ihre Systeme unterzeichnet. Unter anderem haben folgende Institutionen ORCID bereits verpflichtend in ihre Publikationssysteme integriert: American Association for the Advancement of Science, American Chemical Society, American Geophysical Union, eLife, EMBO, PLOS, The Royal Society und Royal Society of Chemistry²³ Die IDs werden meist, z. B. unter Nutzung der Journal Article Tag Suite (JATS)²⁴ auch in den jeweiligen Publikationen und deren Metadaten veröffentlicht. Auch im in der Verlagsindustrie für die Weitergabe von bibliografischen Daten weit verbreiteten Metadatenformat ONIX²⁵ ist neben anderen Standardidentifikatoren wie ISNI und GND auch vorgesehen, eine ORCID ID mitzuliefern.

Auch bei einigen Forschungsförderern, wie z. B. dem Wellcome Trust in Großbritannien²⁶ oder dem FWF in Österreich,²⁷ ist die Angabe einer ORCID ID eine Voraussetzung für die Beantragung von Fördergeldern. Dies dient sowohl der eindeutigen Zuweisung einer Person im Rahmen des Genehmigungsprozesses als auch im weiteren Verlauf der Projekte beispielsweise zur Erfassung von Forschungsleistungen und -ergebnissen.

In Europa ist die Verwendung von ORCID mittlerweile weit verbreitet.²⁸ In Deutschland steht die Implementierung von ORCID in den Publikationsdatenbanken und Open-Access-Repositoryen deutscher Hochschulen und außeruniversitären Forschungseinrichtungen zwar noch am Anfang, die technischen

23 D. h. bei Publikationen in einer Zeitschrift der genannten Verlage muss i. d. R. zumindest der „Corresponding author“ eine ORCID iD angeben. Siehe dazu: Meadows, Alice; Haak, Laurel (2017): ORCID Open Letter – One Year On Report. <http://doi.org/10.6084/m9.figshare.4828312.v1> [Zugriff: 18.04.2017].

24 <https://jats.nlm.nih.gov/index.html> [Zugriff: 20.04.2017].

25 <http://www.editeur.org/8/ONIX/> [Zugriff: 18.04.2017].

26 <https://wellcome.ac.uk/funding/open-researcher-and-contributor-id-orcid> [Zugriff: 18.04.2017].

27 <https://www.fwf.ac.at/de/news-presse/news/nachricht/nid/20151116-2155/> [Zugriff: 18.04.2017].

28 Wie Anm. 17, S. 7 [Zugriff: 18.04.2017].

Vorraussetzungen für eine ORCID-Integration sind in den gängigen Publikationsmanagementsystemen meist jedoch bereits vorhanden.^{29, 30}

5 ORCID DE

Initiiert durch die Deutsche Initiative für Netzwerkinformation (DINI) wurde das Projekt „ORCID DE – Förderung der Open Researcher and Contributor ID in Deutschland“ gestartet, um ORCID in Deutschland zu fördern. Das Projekt wird von der Deutschen Forschungsgemeinschaft (DFG) seit Mai 2016 für drei Jahre gefördert.³¹ Projektpartner sind das Helmholtz Open Science Koordinationsbüro am Deutschen GeoForschungsZentrum GFZ, die Deutsche Nationalbibliothek (DNB) und die Universitätsbibliothek Bielefeld.

Ziel des Projekts ORCID DE ist es, Hochschulen und außeruniversitäre Forschungseinrichtungen, die die Implementierung von ORCID erwägen, nachhaltig zu unterstützen, z. B. durch den Aufbau einer zentralen Anlaufstelle, um sich informieren und austauschen zu können. Dazu zählt auch die Unterstützung durch ein im Projekt zu erarbeitendes Rechtsgutachten, das datenschutzrechtliche Fragen der institutionellen Implementierung von ORCID behandelt. Des Weiteren soll ORCID als Standard in einer offenen Informationsinfrastruktur etabliert werden, indem ORCID in Diensten wie Bielefeld Academic Search Engine (BASE)³² integriert und mit der GND verzahnt wird. Die Agenda des Projektes sieht demnach eine Bandbreite an organisatorischen sowie technischen und rechtlichen Thematiken vor.³³

29 Siehe hierzu ausführlich: Fuchs, C.; Pampel, H.; Vierkant, Paul (2017): ORCID in Deutschland – Ergebnisse einer Bestandsaufnahme im Jahr 2016. Im Erscheinen. o-bib. Das offene Bibliotheksjournal, 2/2017.

30 Summann, F.: Die Verwendung von Autorenidentifikatoren in Wissenschaftlichen Repositorien: Ansätze, konkrete Umsetzungen und Herausforderungen“, 2016. <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0290-opus4-25950> [Zugriff: 18.04.2017].

31 Bertelmann, R.; Niggemann, E.; Pieper, D.; Elger, K.; Fenner, M.; Hartmann, S.; Höhnnow, T.; Jahn, N.; Müller, U.; Pampel, H.; Schirrwagen, J.; Summann, F. (2015): ORCID DE – Förderung der Open Researcher and Contributor ID in Deutschland. <http://doi.org/10.2312/lis.16.01> [Zugriff: 18.04.2017]. <http://doi.org/10.2312/lis.16.01> [Zugriff: 18.04.2017].

32 <https://www.base-search.net/> [Zugriff: 18.04.2017].

33 Siehe auch: Fenner, M.; Hartmann, S.; Müller, U.; Pampel, H.; Reimer, T.; Scholze, F.; Summann, F. (2016): Autorenidentifikation für wissenschaftliche Publikationen. O-Bib. Das Offene Bibliotheksjournal, 3(4), S. 286–293. <http://doi.org/10.5282/o-bib/2016H4S286-293> [Zugriff: 20.04.2017].

Angestoßen durch das Projekt ORCID DE, wurde Ende 2016 für Deutschland ein nationales ORCID-Konsortium gebildet, wie es auch in anderen europäischen Ländern existiert. Die Technische Informationsbibliothek (TIB) in Hannover hat die administrative Führung dieses nationalen Konsortiums übernommen. Die Mitglieder des Konsortiums profitieren von einem im Vergleich zur Einzelmitgliedschaft reduzierten Beitrag. Mit Stand März 2017 umfasste das Konsortium bereits 20 Mitglieder.³⁴ Alle an einer ORCID-Mitgliedschaft interessierten wissenschaftlichen Einrichtungen in Deutschland haben die Möglichkeit sich daran zu beteiligen. Informationen hierzu finden sich auf der Website des Projektes.³⁵

6 ORCID und GND

6.1 Verlinkung von ORCID und GND

Die GND ist eine kooperativ geführte Normdatei, die u. a. zum Zweck der Identifikation von Autorinnen und Autoren und anderen Personen, die mit einer Publikation oder anderen Ressourcen in Verbindung stehen, eingesetzt wird. Neben Personen verzeichnet die GND auch Körperschaften, Kongresse, Geografika, Sachschlagwörter und Werke aus allen Ländern, Epochen sowie Sach- und Fachgebieten. Diese Entitäten werden eindeutig identifiziert und durch ihre Beschreibung von anderen Entitäten gleichen Namens unterschieden. Die GND stellt ein zentrales Arbeitsinstrument und ein eindeutiges Bezugssystem zur Erschließung von Beständen insbesondere in Bibliotheken im deutschsprachigen Raum, aber auch in Archiven, Museen und anderen Wissenschafts- und Kultureinrichtungen dar. Die Gemeinsame Normdatei umfasst derzeit insgesamt 14 Mio. Datensätze,³⁶ davon ca. 4,4 Mio. Datensätze für Personen.

Im Gegensatz zu ORCID werden die Datensätze für Personen in der GND nicht von den Personen selbst erfasst und gepflegt. Die Neuerfassung und die Pflege von Datensätzen werden von Fachpersonal in Kultur- und Wissenschaftseinrichtungen betrieben. Die Idee hinter der GND und ihrem kooperativen Konzept ist allerdings ähnlich zu ORCID: es soll zusammengeführt werden, was zusammen-

³⁴ <http://www.orcid-de.org/universitaet-bremen-ist-neues-mitglied-im-orcid-deutschland-konsortium> [Zugriff: 18.04.2017].

³⁵ <http://www.orcid-de.org/konsortium/> [Zugriff: 18.04.2017].

³⁶ <https://wiki.dnb.de/download/attachments/43518047/gnd-statistik-2017.pdf> [Zugriff: 18.04.2017].

gehört und zu einer Arbeitersparnis im Erschließungsprozess führen, so dass einmal eingegebene Daten in verschiedenen Zusammenhängen nachnutzbar und für verschiedene Zwecke einsetzbar und auffindbar sind.

Im Gegensatz zum ORCID-Profil soll in GND-Datensätzen keine ausführliche Biografie oder Bibliografie abgebildet werden. Die erfassten Angaben dienen primär dazu, eine Person eindeutig zu identifizieren. Allerdings kann und soll von den GND-Datensätzen aus zu allen Stellen verlinkt werden, die weitere Informationen zu einer Entität bereitstellen. Dies kann im Falle einer Person beispielsweise ein ORCID-Eintrag sein.

Der Unterschied zur ORCID-Anwendung liegt aber nicht nur in den Nutzerinnen und Nutzern begründet, sondern zeigt sich aktuell auch im Grad der Verlinkung zu anderen Entitäten und Ressourcen. In der GND wird empfohlen – wo möglich – eine eindeutige Referenzierung zu einer anderen GND-Entität herzustellen. Um das Beispiel der Affiliation aufzugreifen: Im Datensatz der Person wird ein Link zum jeweiligen GND-Datensatz für die Organisation, an der eine Person tätig ist, eingefügt. Während ORCID auch viele Registrierungen ohne Verlinkungen zu Werken (s. o.) aufweist, ist bei fast jedem Datensatz der GND ein Bezug zu einem Kultur- oder Wissensobjekt nachweisbar. Das bedeutet, dass der GND-Identifizierer mit mindestens einer Publikation, einem Museumsobjekt oder einer Archivalie in Beziehung steht, andernfalls wäre er nicht angelegt worden, um eben ein solches Objekt näher oder eindeutig zu beschreiben.

Eine Herausforderung für die GND besteht darin, dieses Netzwerk, das verteilt über verschiedene Kataloge, Sammlungen und Projekte existiert, sichtbar zu machen. Dies ist u. a. ein Ziel der GND-Kooperative³⁷. Eine prototypische Entwicklung einer solchen Zusammenführung von Informationen aus verschiedenen Katalog- und Informationssystemen, die mit einer ORCID und einer GND verknüpft sind, soll im Projekt ORCID DE entwickelt werden.

Bereits Anfang 2015 wurde in der Gemeinsamen Normdatei eine Möglichkeit geschaffen, ORCID IDs in Datensätzen für Personen anzugeben. Von dieser Funktion wird auch zunehmend Gebrauch gemacht: waren es im März 2016 erst 106 GND-Datensätze, die eine ORCID aufwiesen, so waren es ein Jahr später bereits ca. 1.200 Datensätze. Die in der GND enthaltenen ORCID IDs werden auch in den verschiedenen Exportformaten der GND ausgegeben und veröffentlicht, z. B. in der Linked-Data-Repräsentation³⁸ der GND. Tauchen ORCID- und GND-Datensatz

³⁷ Kett, Jürgen: GND-Kooperative – Zusammenarbeit für die Zukunft. <http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:0290-opus4-25556> [Zugriff: 18.04.2017].

³⁸ <http://www.dnb.de/lds> [Zugriff: 18.04.2017].

in einem gemeinsamen VIAF³⁹-Cluster auf, wird diese Information im Entity-Facts-Service⁴⁰ bzw. zukünftig auch in der Linked-Data-Repräsentation der GND ausgewiesen. Die aktuelle Planung sieht vor, mit den Informationen aus verschiedenen Quellen entsprechend umgehen zu können und die Provenienz der Verlinkungen z. B. zwischen ORCID und GND transparent zu machen.

Da zunehmend Verlage (z. B. auch Springer Nature) wie oben beschrieben in den bibliografischen Metadaten ORCID IDs für Publizierende mitliefern (dies ist z. B. möglich in den Formaten MARC 21/MARC-XML und ONIX), können diese genutzt werden, um (semi-)automatisch bibliografische Daten mit GND-Datensätzen zu verlinken – vorausgesetzt, eine Verlinkung von ORCID und GND existiert bereits.

6.2 Ausbau der Verlinkung von ORCID und GND

Wie kann also die Verlinkung von ORCID und GND ausgebaut und unterstützt werden?

Im Rahmen des Projektes ORCID DE soll ein Abgleich (Matching) der Datenbestände von GND und ORCID erfolgen, um Verknüpfungen (semi-)automatisch zu generieren. Dieses Matching kann lediglich auf Kriterien beruhen, die in beiden Systemen zu einer Person vorhanden sind, dies sind der Name, alternative Namen, assoziierte Länder und Institutionen sowie weitere Identifier, wie z. B. VIAF, ISNI, Scopus Author Identifier oder ResearcherID. Da diese Kriterien voraussichtlich für ein sicheres Matching nicht ausreichen werden, zumal sie nicht in allen Einträgen auftreten, werden die in Beziehung stehenden Publikationen in den Abgleich mit einbezogen.

Eine weitere Aufgabe im Projekt ORCID DE ist es, eine einfache Möglichkeit zum „Nachschlagen“ eines Identifikators im jeweils anderen System für Anwender zu schaffen. Die oben bereits kurz skizzierte „Search&Link“-Funktionalität des ORCID-Verzeichnisses, wie sie für andere Identifier bereits existiert, soll auch für die Verknüpfung zur GND angeboten werden. Ein solcher Workflow kann beispielsweise wie folgt aussehen: Der Anwender initiiert eine Suche in der GND, daraufhin werden auf Basis des Namens und ggf. weiterer Attribute GND-Datensätze vorgeschlagen, die potentiell dieselbe Person repräsentieren. Durch die Auswahl des entsprechenden GND-Eintrages durch die Wissenschaftlerin oder

39 Virtual International Authority File, <http://viaf.org/> [Zugriff: 18.04.2017].

40 <http://www.dnb.de/entityfacts> [Zugriff: 18.04.2017].

den Wissenschaftler kann der GND-Identifizier in das ORCID-Profil übernommen werden.

Abb. 2: Beispielworkflow Verlinkung von ORCID und GND.

Eben eine solche Verlinkungsmöglichkeit soll beiderseitig angeboten werden. Idealerweise verfügen beide Systeme über einheitliche Schnittstellen für eine solche Abfrage und nutzen dafür gemeinsame Softwaremodule. Eine Synchronisationsfunktion zwischen den beiden Systemen ist unerlässlich und wird voraussichtlich eine der größeren Herausforderungen in diesem Projekt darstellen.

Neben den Use-Cases, die auf die Verlinkung von GND und ORCID abzielen, sind die GND-Anwender auch daran interessiert, via ORCID frühzeitig Kenntnis über eine Wissenschaftlerin oder einen Wissenschaftler zu erlangen, der ggf. am Anfang seiner Karriere steht und/oder in der GND nicht verzeichnet ist. In diesem Fall, z. B. beim Anlegen eines neuen ORCID-Profiles oder bei der Verknüpfung eines Eintrages mit dem ersten „Werk“, soll ein Hinweis an die GND erfolgen. Wie ein solcher Hinweis aussehen kann, muss in der nächsten Projektphase noch erarbeitet und mit den GND-Kooperationspartnern abgestimmt werden. Aus konzeptioneller Sicht bietet sich eine Art vorläufiger GND-Datensatz („Kandidat“) an. Die Datenbasis für einen solchen „Kandidaten“ könnte aus freigegebenen Informationen eines ORCID-Profiles inklusive verlinkter Beiträge stammen. Diese Datensätze sollten allen GND-Anwendern zur Verfügung stehen, so dass sie im Rahmen der Erschließung verwendet werden können. Es bleibt zu diskutieren, ob die Erzeugung eines „Kandidaten“ in der GND an ein einschränkendes Kriterium geknüpft wird, wie das Vorhandensein einer Verknüpfung zu mindestens einem Werk.

6.3 Aggregation von Metadaten aus verschiedenen Quellen

Die Verlinkung der Identifikatoren ORCID und GND soll auf der einen Seite die Zusammenführung der verschiedenen Repräsentationen einer Person ermöglichen und auf der anderen Seite die Zusammenführung aller mit einer Person verknüpften Publikationen, Objekte und Daten fördern, die in verschiedenen Informationssystemen verzeichnet sind. Durch diese Aggregation werden wissenschaftliche Beiträge sichtbarer. Darüber hinaus wird die Sicht auf die Forschungsleistung einer Person verbessert. Die aggregierten Daten könnten z. B. perspektivisch auch in Dienste wie Wikidata eingebunden werden und damit auch die Qualität der Wikipedia weiter verbessern.

7 Ausblick

Seit Beginn des ORCID DE Projekts ist die Anzahl der in Deutschland registrierten ORCID IDs von ca. 44.000 ORCID IDs im April 2016 auf ca. 71.000 im März 2017 weiter angestiegen.⁴¹ Die im Projekt bereits entwickelten Informationsangebote, die sich in Arbeit⁴² befindliche Verknüpfung zwischen ORCID und BASE sowie weitere geplante Aktivitäten werden die Sichtbarkeit von ORCID in Deutschland weiter steigern.

Das offene Konzept, die breite Organisationsstruktur und die steigende Mitgliederzahl von ORCID zeigen, dass der Dienst auf dem besten Weg zu einem nachhaltigen Dienst der Informationsinfrastruktur ist. Dies ist insbesondere für Kooperationen mit anderen Non-Profit-Institutionen oder -systemen, wie z. B. der GND, ein entscheidender Faktor in der Zusammenarbeit.

Die Resonanz auf den ersten vom Projekt veranstalteten Workshop⁴³ und die ansteigenden Interessenbekundungen am ORCID Deutschland Konsortium zeigen, wie groß das Interesse von Hochschulen und außeruniversitären Forschungseinrichtungen an der Implementierung von ORCID ist und wie notwendig es ist die Aktivitäten zu koordinieren. Ein zweiter ORCID-DE-Workshop wird im September 2017 in der DNB in Frankfurt am Main stattfinden und sich mit der

⁴¹ <http://www.orcid-de.org/immer-mehr-forschende-besitzen-eine-orcid-id-in-deutschland/> [Zugriff: 18.04.2017].

⁴² Stand: März 2017.

⁴³ <http://www.orcid-de.org/1-orcid-outreach-workshop/> [Zugriff: 18.04.2017].

technischen Implementierung und ihren Voraussetzungen und Rahmenbedingungen beschäftigen.⁴⁴

Die Interoperabilität verschiedener Identifikatoren für Personen zu gewährleisten und diese – wo möglich – miteinander zu verknüpfen, um ein Netz von Daten zu spannen, ist ein Konzept, das im Projekt ORCID DE mit der Verlinkung von ORCID und GND verfolgt wird. Von der Kombination der verschiedenen Ansätze profitieren beide Systeme: Metadaten von Publikationen und öffentlich zugänglichen Quellen (GND) sowie Informationen der Autorinnen und Autoren selbst (ORCID) werden herangezogen, um den größtmöglichen Nutzen aus den Daten zu ziehen. Dieses Zusammenspiel der beiden Systeme unterstützt auf vielfältige Weise den Leitsatz einer offenen Informationsinfrastruktur: „enter once, reuse often“.

Sarah Hartmann

Deutsche Nationalbibliothek
Arbeitsstelle für Standardisierung
Gemeinsame Normdatei (GND)
Adickesallee 1
60322 Frankfurt am Main
Deutschland
E-Mail: s.hartmann@dnb.de
<http://orcid.org/0000-0002-5765-6153>

Heinz Pampel

Helmholtz-Gemeinschaft
Helmholtz Open Science Koordinationsbüro
Deutsches GeoForschungsZentrum GFZ
Telegrafenberg
14473 Potsdam
Deutschland
E-Mail: heinz.pampel@os.helmholtz.de
<http://orcid.org/0000-0003-3334-2771>

⁴⁴ <http://www.orcid-de.org/save-the-date-2-orcid-de-workshop/> [Zugriff: 18.04.2017].