

## tagadata - Digitalisierung von Workflows zur Erstellung von Leistungskennziffern: Anwendungsfall Umfragebibliographien

Jakowatz, Stefan; Steinberg, Jan; Gebert, Agathe

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

### Empfohlene Zitierung / Suggested Citation:

Jakowatz, S., Steinberg, J., & Gebert, A. (2020). tagadata - Digitalisierung von Workflows zur Erstellung von Leistungskennziffern: Anwendungsfall Umfragebibliographien. *B.I.T. online: Bibliothek, Information, Technologie*, 23(3), 299-309. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-69259-2>

### Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

# tagadata – Digitalisierung von Workflows zur Erstellung von Leistungskennziffern

## Anwendungsfall Umfragebibliographien

Stefan Jakowatz, Jan Steinberg, Agathe Gebert

### Leistungsmessung im deutschen Wissenschaftsbetrieb

Leistungsmessung und Leistungsmonitoring sind zentrale Informations- und Steuerungsinstrumente für die leistungsorientierte Mittelvergabe (LOM) in der deutschen Wissenschaftslandschaft<sup>1</sup> mit dem Ziel, Effizienz- und Effektivitätsgewinne durch Ressourcenverteilung zu bewirken<sup>2</sup>. Voraussetzung dafür ist eine verlässliche Leistungsmessung durch verschiedene Evaluationssysteme sowohl auf der Makroebene, bspw. durch Länder oder Forschungsförderer, als auch auf der Mikroebene innerhalb von Hochschulen und Wissenschaftseinrichtungen<sup>3</sup>.

Im Rahmen einer umfassenden Verwaltungsmodernisierung seit den 2000er Jahren wurden Steuerungselemente des *New Public Management (NPM)*<sup>4</sup> an Hochschulen eingeführt, um Aufbau- und Ablaufstrukturen des Forschungs- und Lehrbetriebes<sup>5</sup> auf eine leistungsorientierte Ressourcensteuerung<sup>6</sup> auszurichten. Allen dabei zugrunde liegenden Modellen<sup>7</sup> des NPM ist ein Paradigmenwechsel von einer Input- zu einer Output-orientierten Steuerung von Hochschulaktivitäten gemein<sup>8</sup>. Die staatliche Hochschulfinanzierung wurde somit nicht mehr über die klassische Kameralistik mit ihrem Ressourcenverbrauchskonzept<sup>9</sup> gesteuert, sondern durch eine ergebnis-

*Die Einführung des New Public Management (NPM) an deutschen Wissenschaftseinrichtungen verbunden mit dem Wechsel von einer Input- zu einer Output-orientierten Ressourcensteuerung zog die Notwendigkeit von Leistungserfassung nach sich. Bei GESIS – Leibniz-Institut für Sozialwissenschaften wurde die Erhebung von Leistungskennzahlen zur Forschungsdatennutzung durch ein neues integriertes Verfahren zur Erstellung von Umfragebibliographien zentralisiert und optimiert. Das zentrale Management von Umfragebibliographien induziert jedoch ganz neue Anforderungen hinsichtlich der Steuerung von Arbeitsprozessen und des integrierten Metadatenmanagements über unterschiedliche Arbeitsbereiche hinweg. Da klassische Literaturverwaltungsprogramme in diesem Kontext an organisatorische und technische Anwendungsgrenzen stoßen, konzipiert und entwickelt GESIS ein bedarfsorientiertes, innovatives und nachnutzbares Management-Tool tagadata.*

1 Siehe Wissenschaftsrat: Empfehlungen zur Bewertung und Steuerung von Forschungsleistungen, Halle 2011, S. 24ff. [https://www.wissenschaftsrat.de/download/archiv/1656-11.pdf?\\_\\_blob=publicationFile&v=3](https://www.wissenschaftsrat.de/download/archiv/1656-11.pdf?__blob=publicationFile&v=3) [21. November 2019].

2 Schubert, Torben: *New Public Management und Leistungsmessung im deutschen Forschungssektor: Theorie, Umsetzung und Wirkungsanalyse*, Diss. Universität Erlangen-Nürnberg 2008, S. 50. <https://nbn-resolving.org/urn:nbn:de:bvb:29-opus-8701> [21. November 2019].

3 Vgl. ebda., S. 42; S. 80.

4 Im deutschsprachigen Raum auch als „Neues Steuerungsmodell“ (NSM) bezeichnet.

5 Siehe bspw. Hener, Yorck/ Giebisch, Petra/ Roessler, Isabel: *Entwicklung geeigneter Indikatoren und Kennzahlen für die Steuerung der Universität Leipzig – Benchmarking von Fakultäten*, Arbeitspapier Nr. 103, Gütersloh: CHE Centrum für Hochschulentwicklung 2008. [https://www.che.de/downloads/Benchmarking\\_Fakultaeten\\_AP103.pdf](https://www.che.de/downloads/Benchmarking_Fakultaeten_AP103.pdf) [21. November 2019].

6 Flemming, Arend: *Ursachen der Verwaltungsmodernisierung in Bibliotheken*, in: Umlauf, Konrad/ Vonhof, Cornelia (Hrsg.): *Erfolgreiches Management von Bibliotheken und Informationseinrichtungen*, Hamburg: Dashöfer 2011, S. 3.

7 Vgl. Schubert (2008), S. 16ff.

8 Ebda., S. 41.

9 Budäus, Dietrich: *Aktuelle Bestrebungen um Leistungserfassung und leistungsorientierte Ressourcensteuerung in öffentlichen Verwaltungen*, in: Budäus, Dietrich (Hrsg.): *Leistungserfassung und Leistungsmessung in Hochschule. 2. Norddeutsche Fachtagung zum New Public Management*, Wiesbaden: Springer Gabler 2000, S. 12f.

und wirkungsorientierte<sup>10</sup> sowie auf Indikatoren basierte Leistungserfassung<sup>11</sup>. Der im Hochschulrahmengesetz (HRG §5)<sup>12</sup> und in den Landeshochschulgesetzen festgeschriebene Leistungsbezug wirkt dabei als zentrales Steuerungsinstrument<sup>13</sup> auf Grundlage von Verfahren quantifizierbarer Leistungsmessung<sup>14</sup>. Erreicht wurde dies durch die Einführung betriebswirtschaftlicher Instrumente der Kosten- und Leistungsrechnungen (KLR)<sup>15</sup> sowie durch Monitoring bzw. Controlling von Forschungs- und Lehraktivitäten und nicht zuletzt aufgrund der Etablierung eines damit einhergehenden Berichtswesens<sup>16</sup>.

#### Leistungsmessung in außeruniversitären Wissenschaftseinrichtungen (AuW)

Leistungstransparenz ist auch in außeruniversitären Wissenschaftseinrichtungen (AuW) eine unabdingbare Voraussetzung für Effektivitäts- und Effizienzmessung<sup>17</sup>. Im Zuge des 2012 in Kraft getretenen Wissenschaftsfreiheitsgesetzes (WissFG)<sup>18</sup> wurden betriebswirtschaftliche Controlling-Maßnahmen nun auch in AuW initiiert. Vor diesem Hintergrund wurde die Einführung von Verfahren zur indikatorbasierten Leistungsmessung etabliert, um weitgehende Leistungstransparenz herzustellen<sup>19</sup>. Im Pakt für Forschung und Innovation<sup>20</sup> wurde den deutschen Wissenschaftsorganisationen schließlich auferlegt, ein wissenschaftsadäquates sowie transparentes Controlling durchzuführen und die Ergebnisse jährlich an die Gemeinsame Wissenschaftskonferenz (GWK)<sup>21</sup> zu berichten<sup>22</sup>.

#### Indikatorgestützte Leistungserfassung- und Messung in AuW

Indikatorgestützte Leistungserfassung ist eng verbunden mit dem Verfahren der Leistungserhebung, dessen Ergebnisse in einem vordefinierten Interpretationsrahmen möglichst eindeutig nachprüfbar und bewertbar sein sollen<sup>23</sup>. Die Auswahl an aussagekräftigen Indikatoren zur Leistungsmessung hängt also unmittelbar mit dem Interpretationsrahmen und der damit verbundenen Zielsetzung<sup>24</sup> zusammen. Wissenschaftscontrolling setzt dabei auf strategische sowie operative Maßnahmen und Kennzahlen<sup>25</sup>, die je nach Forschungseinrichtung sehr verschieden aussehen und unterschiedliche Relevanz besitzen können<sup>26</sup>. Für AuW hat erstmals Pfeffer (2016) die wichtigsten Leistungsindikatoren empirisch ermittelt und zu Indikatorenkategorien<sup>27</sup> zusammengefasst:

- Qualifikationen (Berufungen, wissenschaftliche Weiterqualifikation, Promotionen)
- Forschung und Innovation (Publikationen, Zitationen, Patente)
- Außenperspektive im Allgemeinen (Forschungsk Kooperationen, Wissenstransfer in die Öffentlichkeit)
- Scientific Community (Preise, eingeladenen Vorträge, Verhältnis zwischen Accepted und Submitted Papers)
- Finanzierung (Forschungsvolumen in Akquise, Kostendeckung, Anteil Wirtschaftserträge, Anteil Drittmittelträge, Erfolgsquote von Anträgen)
- Strategie (Internationalisierung, Wachstum, strategische Qualität der Projekte)

10 Vgl. Gülay (2014), S. 54.

11 Eine kritische Einordnung der Eignung von NPM als Mittel zur Hochschulsteuerung gibt es bspw. bei Pritchard, R.: Humboldtian Values in a Changing World: Staff and Students in German Universities, in: Oxford Review of Education 30 (2004) S. 509–528; siehe Schubert (2018), S. 98.

12 Siehe [https://www.gesetze-im-internet.de/hrg/\\_5.html](https://www.gesetze-im-internet.de/hrg/_5.html) [21. November 2019].

13 Leszczensky, Michael: Paradigmenwechsel in der Hochschulfinanzierung, in: Aus Politik und Zeitgeschichte (25/2004) S. 19.

14 Vgl. Mellewig, Thomas/ Decker, Carolin: Ansätze zur Messung des Organisationserfolgs, in: Controlling 19 (2007) S. 434.

15 Dabei spielt im Rahmen von Wirtschaftlichkeitsanalysen der Faktor Kosten eine zentrale Rolle als Kontroll- und Steuerungsgröße, vgl. ebda., S. 61.

16 Vgl. ebda., S. 181.

17 Klingebiel, Lasse: Steuerungsanforderungen in Non-Profit-Organisationen, in: krp-Kostenrechnungspraxis 43 (1999) S. 372–379.

18 Siehe <https://www.bmbf.de/de/das-wissenschaftsfreiheitsgesetz-466.html> [22. November 2019].

19 Vgl. ebda., S. 212; Hinze, Sybille: Forschungsförderung in Deutschland, in: Simon, Dagmar/ Knie, Andreas/ Hornbostel, Stefan (Hrsg.): Handbuch Wissenschaftspolitik, Wiesbaden: VS Verlag 2010, S.171.

20 Siehe <https://www.gwk-bonn.de/themen/foerderung-von-ausseruniversitaeren-wissenschaftseinrichtungen/pakt-fuer-forschung-und-innovation/> [22. November 2019].

21 Vgl. <https://www.gwk-bonn.de> [22. November 2019].

22 Gemeinsame Wissenschaftskonferenz (GWK): Pakt für Forschung und Innovation. Monitoring-Bericht 2018, Bonn 2018, S. 5. [https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/GWK-Heft-58\\_Monitoring-Bericht-2018.pdf](https://www.gwk-bonn.de/fileadmin/Redaktion/Dokumente/Papers/GWK-Heft-58_Monitoring-Bericht-2018.pdf) [22. November 2019].

23 Vgl. Brüggemeier, Martin: Leistungserfassung und Leistungsmessung in Hochschulen, in: Dietrich Budäus (Hrsg.): Leistungserfassung und Leistungsmessung in öffentlichen Verwaltungen. 2. Norddeutsche Fachtagung zum New Public Management, Wiesbaden: Gabler 2000, S. 240.

24 Schubert (2008), S. 39.

25 Vgl. Pfeffer, Markus: Leistungsmessung in der außeruniversitären Forschung. Performance Measurement mit der Balanced Scorecard in Non-Profit-Organisationen, Wiesbaden: Springer Gabler 2016, S. 203ff.

26 Vgl. ebda., S. 261.

27 Vgl. ebda., S. 267f.

### Repräsentation von Serviceleistungen durch forschungs- und lehrbezogene Leistungsindikatoren – ein unzureichender Ansatz

GESIS – Leibniz-Institut für Sozialwissenschaften<sup>28</sup> ist eine wissenschaftliche Infrastruktureinrichtung der Leibniz-Gemeinschaft (WGL)<sup>29</sup>, die forschungsbasierte Dienstleistungen anbietet<sup>30</sup>. Die zentralen Services von GESIS befassen sich mit Informationsinfrastrukturen rund um sozialwissenschaftliche Umfrageprogramme, „bei denen GESIS teilweise an der Datenerhebung teilnimmt und/oder dauerhaft die Aufgaben der Datenaufbereitung, -archivierung und -bereitstellung übernimmt.“<sup>31</sup> Im Rahmen seiner Datenservices<sup>32</sup> steht die Archivierung von empirischen Forschungsdaten im Datenbestandskatalog (DBK)<sup>33</sup>, die Datenregistrierung bei da|ra<sup>34</sup> und DataCite<sup>35</sup> sowie die Datenaufbereitung<sup>36</sup> für die Sekundärnutzung im Fokus. Der Outcome dieser Serviceaktivitäten wird periodisch gemessen, sei es als Zählung der aufgenommenen Studien im DBK oder der Bereitstellung von archivierten Forschungsdaten für die wissenschaftliche Community. Es liegt jedoch auf der Hand, dass der Nachweis der konkreten Nutzung zur Herausforderung wird, da bereitgestellte Forschungsdaten nicht zwangsläufig in eine unmittelbare Verwertung im Forschungsprozess einmünden.

Vor diesem Hintergrund ist das an der Dualität von Forschung und Lehre ausgerichtete Wissenschaftscontrolling<sup>37</sup> auf Einrichtungen wie GESIS nur bedingt übertragbar, da die Leistungen von Infrastruktureinrichtungen<sup>38</sup> zwar forschungsbasiert erbracht werden, jedoch keine Forschungsleistungen sui generis sind. Es ist an dieser Stelle die zentrale Frage zu stel-

len, wie Serviceleistungen von wissenschaftlichen Forschungsinfrastruktureinrichtungen erfasst werden können, wenn die üblichen Indikatorensysteme auf Forschung und Lehre abzielen und keine adäquate Abbildung von Serviceleistungen liefern können. Da die Bewertung von Serviceleistungen über gängige Indikatorensysteme nur unzureichend als Methode der Leistungsmessung taugt, bietet die Quantifizierung von Forschungsdatennutzung einen wichtigen Lösungsansatz, der dazu beiträgt, Effektivität und Effizienz von Datenservices und Umfrageprogrammen bei GESIS adäquat bestimmen zu können<sup>39</sup>.

### Forschungsdatennutzung als relevanter Leistungsindikator für Umfrageprogramme und Datenservices von GESIS

Für die Leistungserfassung bei Umfrageprogrammen und Datenservices von GESIS ist die Verwertung von Forschungsdaten eine ganz entscheidende Variable. Dabei gilt es, Zitationen von Daten in Forschungspublikationen ausfindig zu machen und die daraus resultierende Zählung von Publikationen als Proxy-Variable<sup>40</sup> für die Leistungsbeschreibung von Umfrageprogrammen und Datenservices heranzuziehen. Da die „Leistung“ einer Infrastruktur nicht mit einer direkten objektiven Messung erfasst werden kann, bedarf es einer näherungsweise messbaren Variable wie „Publikationen mit genutzten Daten“ als Proxy-Variable. Der Leistungsparameter „Forschungsdatennutzung“ ist somit ein Schlüsselindikator, der in landläufigen Indikatorenkategorien bislang nicht prioritär adressiert wurde<sup>41</sup>. Vor dem Hintergrund der fünf institutionell verankerten Forschungsdatenzent-

28 <https://www.gesis.org/home> [22. November 2019].

29 <https://www.leibniz-gemeinschaft.de> [22. November 2019].

30 Siehe <https://www.leibniz-gemeinschaft.de/ueber-uns/ueber-die-leibniz-gemeinschaft.html> [22. November 2019].

31 Vgl. <https://www.gesis.org/institut/integrierte-erhebungs-und-dateninfrastruktur> [22. November 2019].

32 Vgl. <https://www.gesis.org/datenservices/home> [22. November 2019].

33 <https://dbk.gesis.org/dbksearch/index.asp?db=d> [22. November 2019].

34 <http://www.da-ra.de/home/> [22. November 2019].

35 <https://datacite.org> [22. November 2019].

36 Siehe <https://www.gesis.org/datenservices/ueberblick-aller-datenservices> [22. November 2019].

37 Slunder, Sascha: Kennzahlenbasierte Leistungsmessung und darauf aufbauende Mittelallokation in Universitäten – eine empirische und implikationstheoretische Analyse, Diss. Universität Duisburg/Essen 2008, S. 35.

38 Für WGL siehe <https://www.leibniz-gemeinschaft.de/infrastrukturen/forschungsinfrastrukturen-in-der-leibniz-gemeinschaft/informationsinfrastrukturen.html> [22. November 2019].

39 Vgl. Klingebiel, Norbert: Steuerungserfordernisse in Non-Profit-Organisationen, in: krp-Kostenrechnungspraxis 43 (1999) S. 374.

40 Schubert betrachtet Zitationsindikatoren als Proxy-Variable für Qualität von Forschungsleistungen, vgl. Schubert (2008), S. 57ff. Das Prinzip der Proxy-Variable wird an dieser Stelle auf die Messung der Leistung von Umfrageprogrammen und Datenservices übertragen.

41 Die zur Anwendung in AuW kommenden Kategorien sind zumeist Bibliometrie-zentriert und lassen Datenbezüge außen vor, vgl. Pfeffer (2016), S. 267. Im internationalen Rahmen, bspw. im European Strategy Forum on Research Infrastructures (ESFRI), werden Monitoring-Indikatoren für Forschungsinfrastrukturen durchaus appliziert, siehe <https://www.esfri.eu/esfri-events/esfri-validation-workshop-monitoring-research-infrastructures-methodology-and-key> [21. November 2019].

**Tabelle 1:**  
Forschungs-  
datenzentren,  
Umfragepro-  
gramme und  
Umfragebiblio-  
graphien

FDZ ALLBUS	Allgemeine Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS)
FDZ IUP	Comparative Study of Electoral Systems (CSES)
	Eurobarometer Surveys
	European Values Study (EVS)
	International Social Survey Programme (ISSP)
FDZ Wahlen	German Longitudinal Election Study (GLES)
	Politbarometer
FDZ PIAAC	Programme for the International Assessment of Adult Competencies (PIAAC)
GESIS	GESIS Panel

ren (FDZ)<sup>42</sup> – FDZ ALLBUS<sup>43</sup>, FDZ Internationale Umfrageprogramme (IUP)<sup>44</sup>, FDZ Wahlen<sup>45</sup>, FDZ PIAAC<sup>46</sup> und FDZ German Microdata Lab (GML)<sup>47</sup> – hat sich bei GESIS ein dezentrales System der Erhebung von Leistungskennzahlen zur Forschungsdatennutzung etabliert, das in der AuW-Landschaft in diesem Umfang einmalig ist. Fast allen<sup>48</sup> FDZ ist gemein, dass sie Umfragebibliographien mit referenzierten Publikationen erstellen, in denen Umfragedaten verwertet und entsprechend zitiert wurden. Dieses System und dessen digitale Optimierung durch eine nachnutzbare Softwareentwicklung soll im Folgenden näher betrachtet werden.

### Umfragebibliographien als Instrument zur Erhebung von Forschungsdatennutzung

#### Forschungsdatennutzung im Kontext des Berichtswesens von GESIS

Wie bei allen AuW steht GESIS bei verschiedenen Stakeholdern in der Berichtspflicht, so auch hinsichtlich der Forschungsdatennutzung. Kennzahl für die Datennutzung ist die Anzahl veröffentlichter Publikationen auf Basis der darin verwerteten Forschungsdaten

aus den Umfrageprogrammen.

Das Berichtswesen von GESIS zur Forschungsdatennutzung erstreckt sich auf drei unterschiedliche Bereiche. Leistungskennzahlen werden einerseits an institutionelle Gremien (1) berichtet, bspw. an das Kuratorium, den wissenschaftlichen Beirat und den Nutzerbeirat<sup>49</sup>. Darüber hinaus sind Kennzahlen zur Datennutzung wichtig für die regelmäßige Berichterstattung an externe Stakeholder (2), bspw. für den Evaluationsbericht<sup>50</sup> der Leibniz-Gemeinschaft und für den Rat für Sozial- und Wirtschaftsdaten (RatSWD)<sup>51</sup>, bei dem die Forschungsdatenzentren von GESIS akkreditiert<sup>52</sup> sind<sup>53</sup>. Letztlich melden die Umfrageprogramme ihre Kennzahlen an ihre jeweiligen Programmbeiräte (3), bspw. an das GESIS Panel Scientific Quality Assurance Board<sup>54</sup> und an die Koordinationsgruppe GLE<sup>55</sup>. Neben den zahlreichen Publikationsformaten bei GESIS werden Umfragebibliographien auch an die assoziierten Umfrageprogramme weitergegeben, bspw. an das NORC at the University of Chicago<sup>56</sup> zur Einarbeitung in die offizielle ISSP-Bibliographie. Nicht zuletzt für das Institut selbst ist das Reporting relevant, um Kennzahlen zur

42 <https://www.gesis.org/institut/forschungsdatenzentren> [22. November 2019].

43 <https://www.gesis.org/institut/forschungsdatenzentren/fdz-allbus> [22. November 2019].

44 <https://www.gesis.org/institut/forschungsdatenzentren/fdz-internationale-umfrageprogramme> [22. November 2019].

45 <https://www.gesis.org/institut/forschungsdatenzentren/fdz-wahlen> [22. November 2019].

46 <https://www.gesis.org/institut/forschungsdatenzentren/fdz-piaac> [22. November 2019].

47 <https://www.gesis.org/institut/forschungsdatenzentren/fdz-german-microdata-lab> [22. November 2019].

48 Bibliographien sind für das German Microdata Lab von GESIS geplant, aber bislang noch nicht umgesetzt (Stand Oktober 2019).

49 Vgl. <https://www.gesis.org/institut/der-verein> [22. November 2019].

50 Siehe Leibniz Gemeinschaft: Stellungnahme zu GESIS – Leibniz-Institut für Sozialwissenschaften e.V., Mannheim/Köln 2019, S. A-10ff.

[https://www.leibniz-gemeinschaft.de/fileadmin/user\\_upload/Bilder\\_und\\_Downloads/Über\\_uns/Evaluierung/Senatsstimmungen/GESIS\\_-\\_Senatsstimmungen\\_\\_26-03-2019\\_\\_mit\\_Anlagen.pdf](https://www.leibniz-gemeinschaft.de/fileadmin/user_upload/Bilder_und_Downloads/Über_uns/Evaluierung/Senatsstimmungen/GESIS_-_Senatsstimmungen__26-03-2019__mit_Anlagen.pdf) [22. November 2019].

51 <https://www.ratswd.de> [22. November 2019].

52 Vgl. <https://www.ratswd.de/forschungsdaten/fdz> [22. November 2019].

53 Siehe u. a. die Tätigkeitsberichte der WGL, <https://www.ratswd.de/publikationen/taetigkeitsberichte> [22. November 2019].

54 <https://www.gesis.org/gesis-panel-home/gesis-panel-team/> [22. November 2019].

55 Hier auch bis Abschluss des Projektrahmens 2017 an die Deutsche Forschungsgemeinschaft (DFG), <https://www.dfg.de> [22. November 2019].

56 Vgl. <https://www.norc.org/Pages/default.aspx> [22. November 2019].

Datennutzung für die interne Evaluation des GESIS-Angebotskatalogs<sup>57</sup> und für die haushälterische Kosten-Leistungsrechnung (KLR) heranzuziehen.

## Erstellung von Umfragebibliographien bei GESIS

### Status Quo einer dezentralen Erhebung von Publikationen mit Forschungsdatenbezug

Umfragebibliographien wurden bei GESIS für die jährlichen Berichtszeiträume bislang durch die jeweiligen Teams der Umfrageprogramme<sup>58</sup> erstellt. Die dezentralen Erstellungsprozesse lassen sich survey-übergreifend in vier Phasen unterteilen:

Die Recherche (1) nach relevanten Publikationen auf Grundlage verwendeter Forschungsdaten wurde dezentral in jeweils für die Surveys maßgeblichen Quellen durchgeführt, bspw. in den von GESIS lizenzierten Discovery-Systemen EBSCO Discovery-Service<sup>59</sup>, PsycINFO<sup>60</sup> oder Web of Science<sup>61</sup>, in frei zugänglichen digitalen Online-Bibliotheken wie ERIC<sup>62</sup>, in Bibliothekskatalogen, in relevanten Zeitschriften und vor allen Dingen in Google Scholar<sup>63</sup>.

Die Rechercheergebnisse wurden durch die Survey-Teams einer Relevanzbewertung (2) unterzogen. Dazu bedurfte es einer intellektuellen Identifizierung von Datenzitationen anhand von verfügbarem Textmaterial<sup>64</sup>, um Rückschlüsse auf die Art der Datenverwertung ziehen zu können<sup>65</sup>. Bei relevanten Publikationen wurden die vorhandenen Datenzitationen extrahiert und auf Studien-, Daten- und Variablen-Ebene referenziert. Sowohl die Relevanzbeurteilung als auch die Identifizierung von Datenzitationen führten i. d. R. wissenschaftliche Mitarbeiter/-innen der Survey-Teams mit spezifischen Kenntnissen im Umgang mit Umfragedaten durch.

Die daran anschließende bibliographische Aufbereitung und Dokumentation (3) von Publikationen wurde dezentral in jeweils eigenen, zumeist proprietären Literaturverwaltungssystemen<sup>66</sup> der Survey-Teams vorgenommen. Bis dato wurden Publikationen auf

Grundlage von Forschungsdaten ausschließlich von den Teams der jeweiligen Umfrageprogramme recherchiert, hinsichtlich ihrer Bedeutung und Relevanz bewertet und in ihren jeweils eigenen Literaturverwaltungssystemen dokumentiert.

Nach Abschluss des Erstellungsprozesses oblag es den jeweiligen Survey-Teams, Kennzahlen zur Datennutzung (4) für das Reporting an relevante Stakeholder in verschiedenen Ausgabeformaten<sup>67</sup> aus den lokalen Datenbanken heraus zu generieren.

### Zentralisierung und Standardisierung – Synergien durch die Reorganisation der Bibliographienerstellung

Die ausschließlich dezentrale Erstellung von Umfragebibliographien führte zu redundanten Arbeitsprozessen, die sich auch in den dafür verausgabten Ressourcen widerspiegelten. Als Konsequenz daraus hat GESIS eine weitgehende Reorganisation der Bibliographienerstellung initiiert, um potentielle Synergien und damit einhergehende Effektivitäts- und Effizienzgewinne zu erzielen. Die vormalige Einheit der Erstellung von Bibliographien wurde in zwei Arbeitsbereiche unterteilt.

Mit der Zentralisierung (1) von Dokumentation und Recherche wurde eine funktionale Arbeitsteilung entlang der fachlichen Qualifikation der beteiligten Mitarbeiter/-innen umgesetzt.

Aufgaben der Recherche und Dokumentation sowie der Herstellung von Ausgabeformaten wurden auf ein zentrales Team mit vorwiegend dokumentarischen Kompetenzen übertragen. Aufgaben der Relevanzbeurteilung und die Identifizierung von Datenbezügen verblieben dagegen bei den Umfrageprogrammen.

Die bislang bestehende Heterogenität der Datenhaltung und der bibliographischen Erschließung konnte durch Standardisierung (2) egalisiert werden. Die unterschiedlichen lokalen Literaturverwaltungsprogramme wurden zugunsten einer zentralen, webba-

57 Zum differenzierten Angebot von GESIS siehe <https://www.gesis.org/angebot> [22. November 2019].

58 Vgl. Tabelle 1.

59 <https://www.ebsco.com/de-de/produkte/ebsco-discovery-service> [22. November 2019].

60 <https://www.apa.org/pubs/databases/psycinfo/> [22. November 2019].

61 <https://clarivate.com/webofsciencegroup/solutions/web-of-science/> [22. November 2019].

62 <https://eric.ed.gov> [22. November 2019].

63 <https://scholar.google.de> [22. November 2019].

64 Bspw. PDF, Google Books Textsnippets, Abstracts etc.

65 Datenverwertung wird als Verarbeitung von Forschungsdaten im engeren Sinne definiert, d.h. über Datenberechnung, komparative Datennutzung etc. Nicht berichtet werden Publikationen, die nur über Daten – bspw. als Variable Report – oder Analyseergebnisse der Daten durch Dritte Auskunft geben oder Daten nur benennen.

66 Bspw. Reference Manager, EndNote, Citavi, Zotero etc.

67 Bspw. Jahresberichte als GESIS Papers <https://www.gesis.org/angebot/publikationen/gesis-papers>, Web-Datenbanken, Programmformate etc. [22. November 2019].

sierten Literaturverwaltung in Bibsonomy<sup>68</sup> abgelöst. Gleichzeitig konnten die bislang verteilten Metadaten auf ein neu geschaffenes, zentrales Metadatenformat harmonisiert und ein einheitliches dokumentarisches Reglement für die weiteren Erschließungsarbeiten eingeführt werden.

Wesentliche Synergieeffekte ließen sich durch Vermeidung von Redundanzen (3) bei der zentralen Recherche nach relevanten Publikationen erzielen. Durch die nun systematisch und weitgehend einheitlich durchgeführten Recherchen an zentraler Stelle und vor allem durch die Vermeidung von Mehrfachaufnahmen (Dubletten) konnte der Ressourcenaufwand signifikant reduziert werden. Zudem ließ sich die erhebliche Verdichtung der Arbeitsbelastung durch den bisherigen Recherchezyklus zum Ende eines Berichtsjahres durch unterjährig verteilte Recherchen deutlich reduzieren.

### Herausforderungen durch Reorganisation der Bibliographienerstellung – ein digitaler Lösungsansatz

Effizienzsteigerung, Ressourcenoptimierung und eine höhere dokumentarische Qualität der Bibliographien sind einerseits positive Aspekte der Reorganisation. Andererseits zieht die funktionale Arbeitsteilung einen deutlich höheren Koordinations- und Kommunikationsbedarf zwischen dem neuen zentralen Team und den jeweiligen Survey-Teams und vor allem zunehmende Informations- und Datenbewegungen zwischen den Teams nach sich.

Eine zentrale Literaturverwaltung über Bibsonomy bietet – dies gilt auch für andere Literaturverwaltungssysteme<sup>69</sup> – keine adäquate Lösung für eine effiziente Massen- und Mehrfachbearbeitung von Rechercheergebnissen, die aufgrund von unterjährig wiederholten Recherchen mit wiederkehrenden Recherchemengen erzeugt werden. Insofern ist das Verfahren zur

68 <https://www.bibsonomy.org> [22. November 2019].

69 Bspw. Zotero, Endnote oder Reference Manager, aber auch DSpace (<https://duraspace.org/dspace/>). DSpace ist die Open Source-Software der Social Science Open Access Repository (SSOAR) (<https://www.gesis.org/ssoar/home>) von GESIS [22. November 2019].

# BIS-C 2020

<4th. generation>

Archiv- und Bibliotheks-Informationssystem

DABIS.eu

Gesellschaft für Datenbank-Informationssysteme

DABIS.eu - alle Aufgaben - ein Team

Synergien: WB-Qualität und ÖB-Kompetenz

Modell: FRBR . FRAD . RDA Szenario 1 + 2

Regelkonform RDA.RAK.RSWK.Marc21.MAB

Web . SSL . Integration & Benutzeraccount

Verbundaufbau.Cloud/Outsourcing-Betrieb

Archiv Bibliothek Dokumentation

<b>singleUser</b>	<b>System</b>	<b>multiUser</b>
<b>Lokalsystem</b>	<b>und</b>	<b>Verbund</b>
<b>multiDatenbank</b>		<b>multiServer</b>
<b>multiProcessing</b>		<b>multiThreading</b>
<b>skalierbar</b>	<b>performance</b>	<b>stufenlos</b>
<b>Unicode</b>	<b>DSGVO-konform</b>	<b>multiLingual</b>
<b>Normdaten</b>	<b>GND RVK</b>	<b>redundanzfrei</b>
<b>multiMedia</b>	<b>eMedia</b>	<b>Integration</b>

Software - State of the art - flexible

31 Jahre Erfahrung Wissen Kompetenz

Leistung Sicherheit Datenschutz

Standards Offenheit Individualität

Stabilität Partner Verlässlichkeit

Service Erfahrung Support

Generierung Customizing Selfservice

Outsourcing Cloudbetrieb SaaS

Dienstleistung Zufriedenheit

GUI.Web.XML.Z39.50/SRU.OAI-METS

Portale mit weit über 17 Mio Beständen

<https://Landesbibliothek.eu>
<https://bmnt.at>  
<https://OeNDV.org>
<https://VThK.eu>  
<https://VolksLiedWerk.org>
<https://bmdw.at>  
<https://Behoerdenweb.net>
<https://wkweb.at>

DABIS GmbH

Heiligenstädter Straße 213, 1190 Wien, Austria  
 Tel. +43-1-318 9777-10 Fax +43-1-318 9777-15  
 eMail: [support@dabis.eu](mailto:support@dabis.eu) <https://www.dabis.eu>

Zweigstellen: 61350 - Bad Homburg vdH, Germany / 1147 - Budapest, Hungary / 39042 - Brixen, Italy

Ihr Partner für Archiv-, Bibliotheks- und DokumentationsSysteme

Dublettenbehandlung in Bibsonomy inadäquat, da es immer nur auf den Gesamtbestand ausgerichtet ist und keine Möglichkeit der differenzierten Dublettenbehandlung bei Neuzugängen bietet. Darüber hinaus erfordern (mehrmals) wiederkehrende Publikationen einen deutlich höheren Aufwand, da Informationen zum vorliegenden Umfrageprogramm sowie zur Relevanzbeurteilung dokumentiert und nicht-relevante Publikationen ebenfalls archiviert werden müssen, um bei Bedarf einen Abgleich mit späteren Importen vornehmen zu können.

### tagadata – Ein digitales Management-Tool zur effizienten Bibliographienerstellung

Die Herausforderungen infolge der Reorganisation der Bibliographienerstellung liegt darin, Daten- und Informationsflüsse sowie Dokumentationsprozesse zwischen den beteiligten Teams weitgehend automatisiert und in einem Tool integriert zu steuern. Zu diesem Zweck wurde das webbasierte Dokumentations- und Management-Tool tagadata entwickelt, das verteilte Aufgaben in einem zentralen System abbildet und Workflows digital steuert. Folgende Hauptfunktionen sind in tagadata bislang maßgeblich:

- Bibliographisches Dokumentationssystem mit Möglichkeiten zum Import von Rechercheergebnissen über BibTeX<sup>70</sup>. Unterstützung der dokumentarischen Aufbereitung von Metadaten durch formale Validierung der Erschließung.
- Verfahren der automatischen Identifizierung von Dubletten über bibliographische Entitäten. Bei wiederkehrenden Publikationen erfolgt ein Abgleich von Informationen zur (positiven bzw. negativen) Relevanz und zu den bereits prozessierten Surveys.
- Ein Tagging-strukturierter Metadatenpool für alle Publikationen, differenziert nach automatisch zugewiesenen „Tags“. Je nach Umfrageprogramm werden alle Publikationen mit Survey- und Relevanz-Tags (relevant vs. irrelevant) im tagadata-Pool archiviert. Referenzen „relevant“ beurteilter Publikationen werden zudem über eine Schnittstelle nach Bibsonomy automatisch exportiert. Die existierende Tagging-Struktur ist letztendlich eine grundlegende Voraussetzung für die Erhebung von Leistungskennzahlen.
- Rollengestütztes Workflow-Management sequenzi-

eller Arbeitsabläufe über teamspezifische Arbeitslisten.

- Automatische Generierung von Statistiken und Kennzahlen für jedes Umfrageprogramm.

Mit tagadata wurde eine Open-Source-Software für ein rollengestütztes Workflow-Management entwickelt, das äußerst flexibel an bestehende Use Cases angepasst werden kann, seien es Arbeitsaufgaben und -abläufe, Entitäten, Metadatenschemata und die Art und Weise der automatisierten Verarbeitung. Im GESIS Use Case (Abbildung 1) werden die Workflows zur Erstellung von Umfragebibliographien mit Hilfe von tagadata durch weitgehende Automatisierung der Verfahrensabläufe – bspw. der Koordination von Arbeitsschritten zwischen verteilten Teams – und der Informationsverarbeitung – bspw. die Dublettenkontrolle – kohärent prozessiert.

### Konzeption und softwaretechnische Entwicklung von tagadata

Die Konzeption für tagadata folgt den Anforderungen der bestehenden Workflows zur Bibliographienerstellung bei GESIS. Da es sich bei sämtlichen zur Anwendung kommenden Technologien um Open-Source-Produkte handelt, ist die freie Nachnutzung garantiert und eine Adaption von tagadata an andere spezifische Arbeitsumgebungen jederzeit möglich.

Bei tagadata handelt es sich um eine Java-Anwendung (Framework Spring Boot<sup>71</sup>) mit einem React-Redux-Frontend<sup>72</sup> und einer SQL-Datenbank der Wahl<sup>73</sup>. Hintergrund für dieses Setup ist die Plattformunabhängigkeit, der weite Verbreitungsgrad sowie die Nachhaltigkeit durch eine aktive Entwickler-Community. Wie im agilen Projektmanagement<sup>74</sup> üblich, entstand tagadata in iterativen Zyklen. Der dafür notwendige und stetige Review-Prozess wurde durch ein Continuous Delivery<sup>75</sup> realisiert, wobei das entstehende Produkt jederzeit in einer lauffähigen Version zusätzlich zu den automatisierten Testverfahren – auch zum manuellen Testen – zugänglich blieb<sup>76</sup>.

Innerhalb der Softwarearchitektur ist tagadata mit seinen Funktionalitäten um die Hauptentität „Publication“ aufgebaut. Direkt an sie angehängt finden sich neben der Referenz bzw. Object-ID auch die für die

70 Die meisten Recherchequellen – Datenbanken, Google Scholar, Verlagsangebote etc. – verwenden als Exportformat BibTeX (<http://www.bibtex.org/de/>) [22. November 2019].

71 Siehe <https://spring.io/projects/spring-boot> [22. November 2019].

72 Siehe <https://react-redux.js.org/> [22. November 2019].

73 Im GESIS Use Case handelt es sich um eine MySQL-Datenbank.

74 Erläuterung der Grundprinzipien des Agilen Projektmanagements siehe <https://www.cio.com/article/3156998/agile-project-management-a-beginners-guide.html> [22. November 2019].

75 Siehe <https://searchitoperations.techtarget.com/definition/continuous-delivery-CD> [22. November 2019].

76 Der weitere Technologiestack u. a.: Maven, JUnit, Cucumber, GitLab, Jenkins.

Publikation relevanten Bezeichnungen der Umfrageprogramme, die jeweilige Tag-Auszeichnung der Surveys, diverse Status wie bspw. der Relevanz-Status sowie der optional zugehörige Volltext<sup>77</sup>. Jeder „Publication“ ist ein Metadatensatz angehängt, der die einzelnen Einträge in Key-Value-Paaren<sup>78</sup> bündelt. Das Vorhalten der Metadaten in Key-Value-Paaren ermöglicht eine hohe Flexibilität beim Einrichten von Internformaten<sup>79</sup>, so dass zusätzliche Felder ohne aufwändige Eingriffe in den Programmcode angelegt werden können. Um eine umfassende, optimale Nachnutzung zu ermöglichen, wurde in tagadata „DublinCore qualified<sup>80</sup>“ als Ausgangsformat vorkonfiguriert. Dabei bleibt dem Systemanwender die Möglichkeit erhalten, Beschreibungsfelder innerhalb oder jenseits des Schemas nach eigenen Vorstellungen hinzuzufügen bzw. ein komplett eigenes Internformat zu verwenden.

Die Vorgabe eines solchen Ausgangsformats bringt die Möglichkeiten der vorkonfigurierten automatischen Metadaten-Validierung vorgegebener Felder mit sich, bspw. für standardisierte Inhalte wie Erscheinungsjahr oder URLs und des schon vorweg integrierten Mappings auf externe Datenhaltungssysteme wie Bibsonomy im GESIS Use Case. Die einzelnen Workflow-Schritte werden in tagadata durch Java-Sets<sup>81</sup> (Listenobjekte) abgebildet. Als feste Sets sind die Eingangsliste für neue Publikationen und die Archivliste mit fertig bearbeiteten Referenzen definiert, die als Ausgangs- bzw. als Endpunkt festgelegt sind, wobei die Konfiguration des Workflows dazwischen absolut flexibel gestaltbar ist.

Die Steuerung des Workflows in tagadata erfolgt über die rollengestützte Zuweisung von Rechten und Funktionalitäten für die Anwender bzw. Anwendergruppen

77 I. d. R. als PDF-Dokument.

78 So bildet der Feldname den Key (bspw. „dc.date.issued“) und der Feldinhalt den Value (bspw. „2019“).

79 Als Internformat bezeichnet man die Gesamtheit der Metadaten, die innerhalb der Applikation Verwendung findet. Es entspricht häufig einem bestimmten Metadatenschema (bspw. DublinCore, BibTeX, RIS, MARC21, UNIMARC u. ä.). Der hohe Freiheitsgrad an dieser Stelle ermöglicht sämtliche denkbaren Feldname-Feldinhalt-Kombinationen auch ohne Schemafestlegung.

80 Siehe <https://guides.library.ucsc.edu/c.php?g=618773&p=4306386> und <https://www.dublincore.org/specifications/dublin-core/dces/> [22. November 2019].

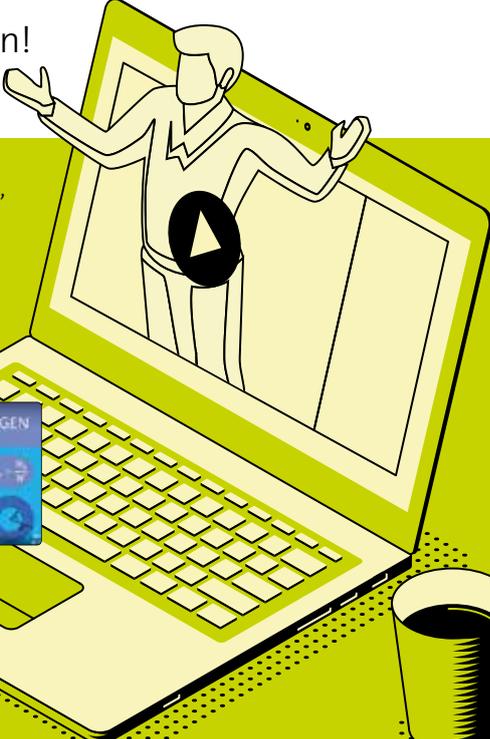
81 <https://docs.oracle.com/javase/9/docs/api/java/util/Set.html> [22. November 2019].



# E-Learning mit wiso

Jetzt die neuen Lern- und Erklärvideos entdecken!

Der Bedarf an digitalen Inhalten, die die eigene Lehre unterstützen, ist gerade in der aktuellen Zeit hoch. Die vielseitigen Lern- und Erklärvideos von wiso vermitteln Ihren Studierenden komplexes Wissen verständlich und kurzweilig. Klausurrelevante Lerninhalte werden dauerhaft verankert – und das ganz nach individuellem Lerntempo.



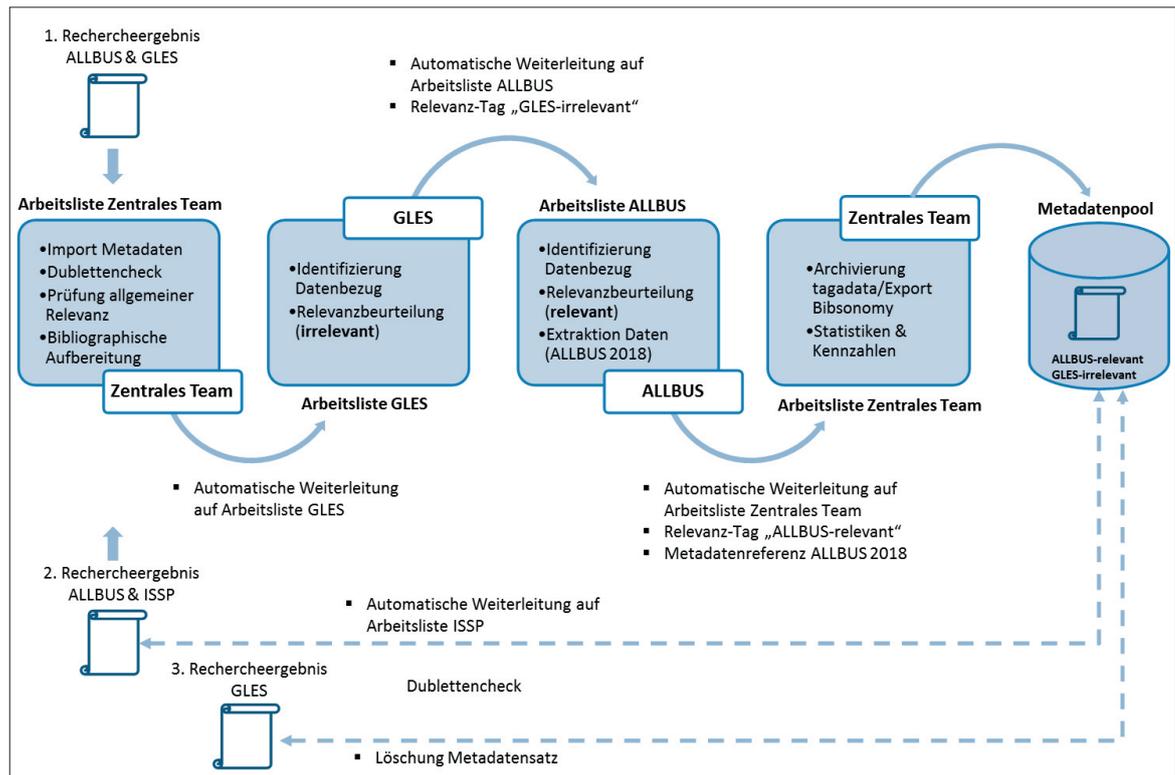




Jetzt kostenlos und unverbindlich testen!  
Schreiben Sie uns: [valerie.lode@genios.de](mailto:valerie.lode@genios.de)

Ein Angebot von GBI-Genios Deutsche Wirtschaftsdatenbank GmbH

Abbildung 1:  
Darstellung des  
rollengestütz-  
ten Workflow-  
Managements  
von tagadata  
im Beispiel-Use  
Case



der verschiedenen Arbeitsbereiche. Technisch gesehen ist die Definition von Rollen und die damit verbundene Konfiguration von Funktionalitäten frei wählbar. Über die administrative Klammer der Listenobjekte bzw. Arbeitslisten werden die verschiedenen Rollen und Funktionalitäten zu einem stringenten Workflow miteinander verbunden.

Im GESIS Use Case nimmt das zentrale Team die Rolle der Dokumentation und Archivierung ein, die verschiedenen Surveys bekommen dagegen die Rolle der Relevanzbeurteilung zugewiesen. Durch die Rollen ist wiederum festgelegt, welche Funktionalitäten für die jeweiligen Anwendergruppen zugelassen sind, bspw. erhalten Survey-Teams nur die Möglichkeit des Zugriffs auf ihre eigenen Arbeitslisten und nicht auf die Arbeitslisten anderer Teams. Für die Rolle des zentralen Teams stehen alle Funktionalitäten der Dokumentation, d. h. des Imports von Metadaten<sup>82</sup>, der Dublettenkontrolle und der bibliographischen Erschließung sowie der abschließenden Qualitätssicherung und Archivierung zur Verfügung, demgegenüber umfasst die Rolle der Surveys die Funktionalitäten der Relevanzbeurteilung und Dokumentation der Datenbezüge im tagadata-Metadatenchema.

Die beliebige Konfigurierbarkeit von Java-Sets in tagadata erlaubt es, den Koordinationsbedarf zwischen den Anwenderrollen als kohärenten Arbeitsprozess

passgenau zu modellieren. Dabei wird der Datenfluss („Flow“) durch eine vorab festgelegte Abfolge von rollenbasierten Arbeitsschritten gesteuert, in dem der Abschluss einer Arbeitsaufgabe den jeweils nächsten rollenbasierten Arbeitsschritt adressiert und die Hauptentität „Publication“ von einer Anwenderrolle (Arbeitsliste) zur nächsten sequentiell „fließen“ lässt (siehe Abb. 1).

In der Softwarearchitektur von tagadata nimmt das Verfahren der automatisierten Dublettenkontrolle eine zentrale Rolle ein. Dieses Verfahren ist insofern von Bedeutung, da die wiederholt unterjährigen Recherchen einerseits eine mehrmalige Verarbeitung von teilweise ein und denselben Publikationen notwendig macht. Andererseits müssen auch von den Survey-Teams bereits begutachtete und dementsprechend gekennzeichnete Publikationen im Archivbestand mit neu importierten Referenzen abgeglichen werden, um sie gegebenenfalls als bereits begutachtet identifizieren zu können. Das Verfahren der Dublettenprüfung wird direkt beim Import von Metadaten automatisch in Gang gesetzt, zudem kann der Vorgang nach Ergänzung von Metadaten jederzeit manuell und wiederholt angestoßen werden. Eineindeutige Dubletten – sowohl bibliographisch als auch dublett bei der Survey-Kennzeichnung – werden automatisch aus dem Verarbeitungsprozess gelöscht, so

82 Vier Möglichkeiten, Metadaten zu Publikationen in tagadata einzugeben: Manuelle Eingabe über ein Webformular, copy & paste von BibTeX-Snippets, Import (auch mehrerer Datensätze) über eine BibTeX-Datei, Einlesen eines Bibsonomy-Repository.

dass kein weiterer Arbeitsaufwand entsteht. Das Verfahren der Dublettenkontrolle in tagadata trägt ganz entscheidend dazu bei, Mehrfacharbeiten im Vorfeld der Relevanzbeurteilung in erheblichem Umfang zu vermeiden.

### Zusammenfassung und Ausblick

Die Einführung des New Public Management (NPM) an Forschungs- und Lehreinrichtung und der damit einhergehenden leistungsorientierten Mittelvergabe führte zu neuen Formen der indikatorbasierten Leistungsmessung. Für GESIS ist vor diesem Hintergrund die Verwertung von Forschungsdaten seiner Umfrageprogramme in Publikationen ein zentraler Leistungsparameter.

Mit der Zentralisierung der Bibliographienerstellung konnte eine weitgehende Standardisierung und Qualitätsverbesserung der – bis dato uneinheitlichen – Metadaten- und Ausgabeformate erreicht werden. Erhebliche Synergien wurden dadurch geschaffen, so dass die jährlichen Berichte und Reports nunmehr auf einer gleichförmigen Datengrundlage basieren und quasi auf „Knopfdruck“ weitgehend einheitlich und vergleichbar generiert werden können.

Die Reorganisation der Bibliographienerstellung bei GESIS führte zu einer stärkeren Arbeitsteilung zwischen Recherche und Datenverarbeitung einerseits und Relevanzbeurteilung andererseits. Um dem deutlich höheren Koordinationsbedarf zwischen dem zentralen Team und den jeweiligen Survey-Teams und dem zusätzlichen Bedarf an Informations- und Datenbewegungen zwischen den Teams gerecht zu werden, wurde ein digitales Management-Tool zur Steuerung der Workflows entwickelt. Mit tagadata ist eine Open-Source-Software entstanden, die sehr flexibel an bestehende arbeitsteilige Workflows zur Erstellung von Umfragebibliographien angepasst werden kann. Sie leistet für GESIS einen wesentlichen Beitrag zur Erhebung von Leistungskennzahlen im Bereich der Forschungsdatennutzung.

Über den konkret beschriebenen Use Case hinaus eignet sich tagadata als Tool für ein effizientes Organisations- und Metadatenmanagement zur Nachnutzung in anderen Forschungsinfrastruktureinrichtungen. Interessant ist die Nachnutzung von tagadata nicht nur im Rahmen der Leistungsmessung von Umfrage- und Datenservices, sondern auch vor dem Hintergrund ei-

ner offenen Wissenschaft als innovative Methode, die Text-Fakten-Integration in Informationssystemen<sup>83</sup> durch neue digitale Verfahren, frei verfügbarer Software und Textmining entscheidend voranzutreiben<sup>84</sup>. Dieser innovative Aspekt wird mittelfristig durch die geplante Integration von InFoLis<sup>85</sup> in den tagadata-Workflow noch exponierter zutage treten. InFoLis ist ein technisches Verfahren zur automatischen Identifizierung von Datenzitationen in Volltexten, das die bislang aufwändige intellektuelle Identifizierung ablösen wird. Einerseits wird der zukünftige Aufwand für die Erstellung von Umfragebibliographien durch die Automatisierung der Verknüpfung von Publikationen und Forschungsdaten erheblich reduziert, andererseits wird sich der Erstellungsprozess erheblich beschleunigen und zudem validere Ergebnisse bei der Erhebung von Leistungskennzahlen hervorbringen als bisher. **I**



#### Dipl.-Soz. Stefan Jakowatz

Wissenschaftlicher Mitarbeiter bei GESIS und derzeit stellvertretender Teamleiter Team Bibliothek und Open Access. Im Rahmen verschiedener institutionell- und drittmittelgeförderter Projekte begleitete er die Entwicklung zahlreicher Fachinformationssysteme (Sowiport, SOLIS, SOFISwiki, SSOAR etc.). GESIS Leibniz-Institut für Sozialwissenschaften Team Open Access, Abt. Wissenstransfer Unter Sachsenhausen 6-8, 50667 Köln stefan.jakowatz@gesis.org



#### Dipl.-Bibl. Jan Steinberg (M.A. LIS)

Wissenschaftlicher Mitarbeiter und Softwareentwickler im Team Bibliothek und Open Access sowie Leiter der GESIS-Bibliothek. jan.steinberg@gesis.org



#### Dr. Agathe Gebert

Wissenschaftliche Mitarbeiterin und Leitung des Teams Bibliothek und Open Access bei GESIS. Dozentin am Zentrum für Bibliotheks- und Informationswissenschaftliche Weiterbildung (ZBIW) der TH Köln. agathe.gebert@gesis.org

<sup>83</sup> Krause, Jürgen/ Stempfhuber, Maximilian: Nutzerseitige Integration sozialwissenschaftlicher Text- und Dateninformationen aus verteilten Quellen, in: König, C./ Stahl, M./ Wiegand, E. (Hrsg.): Datenfusion und Datenintegration: 6. wissenschaftliche Tagung (Tagungsberichte / Informationszentrum Sozialwissenschaften, 10) Bonn: Informationszentrum Sozialwissenschaften 2005, S. 141-158. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-261147> [22. November 2019].

<sup>84</sup> Siehe <https://ag-openscience.de/open-science/> [22. November 2019].

<sup>85</sup> Siehe <https://www.gesis.org/?id=8948> [22. November 2019].