

Überlieferungsbildung und die Intelligenz im Web.

Digitale Hilfsmittel bei der Bewertung personenbezogener Unterlagen

Dr. Franz-Josef Ziwes



Die Bewertung personenbezogener Unterlagen, insbesondere von Personalakten, erfolgte im Landesarchiv Baden-Württemberg bis 2010 offiziell auf der Grundlage eines Bewertungsmodells von 1991. Danach galten als archivwürdig:

- alle Hauptpersonalakten der Geburtsjahrgänge älter/gleich 1875;
- ab 1876 alle Hauptpersonalakten mit den „repräsentativ ausgewählten Buchstaben D, O und T“
- Hauptpersonalakten der Geburtsjahrgänge mit der Endziffer 5,
- Hauptpersonalakten der Amtsleiter, besondere Einzelfälle (die Auswahl soll in der Regel durch die Behörde erfolgen).

Der qualitative Teil der Bewertung von Personalunterlagen konnte und kann allerdings nicht über die Vorgaben eines Bewertungsmodells geregelt werden. Um die Akten zu finden, die „das Salz in der Suppe ausmachen“, ist das „berufliche Ethos des Archivars besonders gefragt“, so haben es Kurt Hochstuhl und Alfred Straub formuliert, als sie 1997 das baden-württembergische Bewertungsmodell der archivarischen Öffentlichkeit vorstellten.¹

Die Auswahl nach der Dicke der Akten oder nach berühmten und berüchtigten Zeitgenossen mag bei kleineren behördlichen Aussonderungen ja noch angehen. Je größer der Umfang der angebotenen Unterlagen jedoch wird, umso schwieriger gestaltet sich eine effiziente Bewertung am Regal und auch die Bereitschaft der aussondernden Behörde zur inhaltlichen Mitarbeit dürfte sich in den meisten Fällen in reziproker Relation zum Umfang entwickeln.

Eine Qualitätsauslese von Einzelfällen jenseits schematischer Vorgaben bleibt also auch mit Bewertungsmodell eine archivische Herausforderung. Vor einer solchen sah sich die Abteilung des Staatsarchivs Sigmaringen des Landesarchivs Baden-Württemberg, als ihr vom Regierungspräsidium Tübingen zur Aussonderung anstehende Personalakten im Umfang von 27.428 Einheiten, hochgerechnet etwa 330 laufende Meter, zur Übernahme angeboten wurden.

Elektronische Hilfsmittel erleichtern die Auswahl der quantitativen Merkmale durch die mechanische Selektion von Geburtsdaten, Anfangsbuchstaben, Zufallszahlen etc. sofern man natürlich entsprechende Metadaten von den aussondernden Einrichtungen erhalten hat. Glücklicherweise stellte das Regierungspräsidium Tübingen dem Staatsarchiv Sigmaringen eine Aussonderungsliste mit Angaben zu Namen, Vornamen, akademischen Titeln und Geburtsdatum zur Verfügung.

¹ Kurt HOCHSTUHL unter Mitwirkung von Alfred STRAUB: Bewertung von Personalakten. Das baden-württembergische Modell. In: Historische Überlieferungsbildung aus Verwaltungsunterlagen. Zur Praxis der archivischen Bewertung in Baden-Württemberg, hrsg. von Robert KRETZSCHMAR (Werkhefte der Staatlichen Archivverwaltung Baden Württemberg A 7). Stuttgart 1997, S. 227–234.

Ref12_Personalakten_Teil5.xls [Kompatibilitätsmodus] - Microsoft Excel

Start Menü Einfügen Seitenlayout Formeln Daten Überprüfen Ansicht Entwicklertools

Einfügen Zwischenablage Schriftart Ausrichtung Zahl Bedingte Formatierung Als Tabelle formatieren Zellenformatvorlagen Einfügen Löschen Format Zellen Sortieren und Filtern Suchen und Auswählen Bearbeiten

E1 15.3.1910

	A	B	C	D	E	F	G	H	I	J	K
1	4120	Mi	Walter		15.3.1910						
2	4121	Mi	Egon		7.12.1910						
3	4122	Mi	Maximilian		12.8.1949						
4	4123	Ri	Eberhard		17.10.1919						
5	4124	Ri	Hermann		07.11.1896						
6	4125	Se	Hermann		7.3.1910						
7	4126	Se	Hermann		7.7.1947						
8	4127	Se	Gerd		27.1.1954						
9	4128	St	Roland		14.8.1943						
10	4129	St	Alfons		3.1.1910						
11	4130	Va	Annemarie		3.11.1914						
12	4131	W	Gustav		14.5.1933						
13	4132	St	Anton	Dr.	20.11.1912						
14	4133	W	Burkhard		20.1.1928						
15	4134	W	Pius		7.6.1943						
16	4135	Zi	Hermann		19.10.1913						
17	4136	Le	Karel		13.2.1945						
18	4137	Mi	Hans		15.2.1965						
19	4138	Mi	Josef		10.12.1925						
20	4139	Mi	Rüdiger		27.1.1945						
21	4140	Mi	Monika		30.12.1945						
22	4141	Mi	Elzbieta		23.11.1955						
23	4142	Mi	Axel		28.9.1955						
24	4143	Mi	Roland		30.9.1945						
25	4144	Li	Roland		7.10.1945						
26	4145	Mi	Willi		19.7.1945						
27	4146	Ne	Manfred		28.4.1935						
28	4147	Mi	Liselotte		22.3.1915						
29	4148	Ni	Margarete		20.3.1925						
30	4149	Ni	Siegfried		22.6.1935						
31	4150	Ni	Erhard		29.8.1935						
32	4151	Ni	Elisabeth		16.11.1955						
33	4152	Ni	Walter		5.4.1945						
34	4153	Ni	Rainer		21.3.1955						
35	4154	Ni	Gerhard		6.3.1955						

Abbildung 1: Aussonderungsliste des Regierungspräsidiums Tübingen

Die Auswahl der Personalakten nach den schematischen Vorgaben des baden-württembergischen Bewertungsmodells war damit kein großes Problem. Für die qualitative Auslese lagen die Hürden jedoch etwas höher. Hier wurde schließlich auf der Grundlage eines Beschlusses der 48. Sitzung der Arbeitsgruppe Überlieferungsbildung vom 13. Juli 2009 ein Weg gefunden, der aus einer Kombination von detaillierter Einzelanalyse und automatisiertem Abgleich bestand.² Im Folgenden sollen hier die einzelnen Schritte des automatisierten Verfahrens näher beleuchtet werden und damit im Sinne des Workshops als praktische Handreichung für interessierte Kolleginnen und Kollegen dienen.

Es gibt durchaus elektronische Werkzeuge, die eine qualitative, also nach inhaltlichen Kriterien zu treffende Auswahl von Personalakten unterstützen können. Zum einen sind dies Fachverfahren von Landeseinrichtungen, wie die im Internet frei nutzbare Personendatenbank der Landesbibliographie Baden-Württemberg.³ Allerdings können die dort angebotenen Personendaten nur über das Webinterface für die gezielte Recherche nach einzelnen Personen genutzt werden. Eine freie Verwendung der Datenbank für automatisierte Verfahren ist aus Gründen des Datenschutzes nicht möglich.

² An dieser Stelle sei Herrn Markus WOHLBOLD und den Mitarbeitern der Personalaktenregistratur vom Regierungspräsidium Tübingen für das tatkräftige Engagement bei der Aussonderungsaktion gedankt. Zur detaillierten Darstellung des Bewertungsprojekts vgl. Franz-Josef ZIWES: Wikipedia und Co. statt Sisyphus? Konventionelle und digitale Hilfsmittel zur qualitativen Bewertung von Personalakten, in: Archivar. Zeitschrift für Archivwesen 63 (2010), S. 175-178 (online abrufbar unter http://www.archive.nrw.de/archivar/hefte/2010/ausgabe2/Archivar_2_10.pdf).

³ Vgl. www.statistik.baden-wuerttemberg.de/LABI/PersDB.asp; www.wlb-stuttgart.de/literatursuche/baden-wuerttemberg/landesbibliographie/personendatenbank-baden-wuerttemberg/.

[Datei](#) [Bearbeiten](#) [Ansicht](#) [Chronik](#) [Lesezeichen](#) [Extras](#) [Hilfe](#)
[Meistbesuchte Seiten](#) [OPAC](#) [BelWue \(belwue\) on T...](#) [Desktop](#) [LA Home](#) [SixCMS](#) [Bestellverw](#) [Findbuchgen.](#) [Intranet](#) [Suche Findb](#) [Kundenport.](#) [SWR](#) [Google](#) [iGoogle](#) [Stas](#) [Maps](#)
[Intranet-Homepag...](#) [Landesarchiv Bade...](#) [Landesarchiv Bade...](#) [Landesarchiv Bade...](#) [Landesarchiv Bade...](#) [Index of /dewiki/la...](#) [Bearbeiten von „H...](#) [Landesbibliogra...](#)



Landesbibliographie Baden-Württemberg Personendatenbank

online

[Literatursuche](#)

[Personensuche](#)

[Impressum/Information](#)

[Bedienungshinweise](#)

[Ihre Meinung](#)

[Trefferliste](#) - [vorige Person](#) - [nächste Person](#)

Person 38 von 177

ID-Nr.: 36776 >> [Literatur über diese Person](#)
 Name: Feger, Otto
 Kurzbiographie: Archivar, Historiker, 1905-1968
 Geburtsdatum: 10.11.1905
 Geburtsort: Mühlhausen
 Sterbedatum: 26.04.1968
 Sterbeort: Konstanz
 Berufe: Archivar
 Historiker
 Nachweis: Lt 8; BadBi NF 1
 Zuletzt bearbeitet: 25.08.1989

Abbildung 2: Fachverfahren Personendatenbank der Landesbibliographie Baden-Württemberg

Zum anderen finden sich auch frei zugängliche Datenbanken mit Listen von berühmten, öffentlich in Erscheinung getretenen und auch von berüchtigten Personen. Sie liegen freilich etwas versteckt im World Wide Web und ihre Verwendung erfordert einige Handarbeit, doch ist der Zugriff problemlos möglich: Wohl kaum ein Internet-Nutzer kennt nicht die Online Enzyklopädie Wikipedia. Wer von uns hat nicht schon einen Personenartikel darin gelesen? Wer einen solchen Artikel sogar schon einmal bearbeitet oder neu angelegt hat, kennt auch die am Ende des Personenartikels zu berücksichtigenden Personendaten. Dies sind nach der Definition von Wikipedia selbst „spezielle Metadaten, die in Artikeln über Personen aufgenommen werden, damit sie daraus automatisch extrahiert und weiterverarbeitet werden können. Sie bestehen aus einer Reihe von Datenfeldern wie beispielsweise Name, Geburtstag und Kurzbeschreibung einer Person“.⁴

⁴ Stand des Zitats: 19.1.2011; vgl. <http://de.wikipedia.org/wiki/Hilfe:Personendaten>.



Hauptseite
Über Wikipedia
Themenportale
Von A bis Z
Zufälliger Artikel

Mitmachen
Hilfe
Autorenportal
Letzte Änderungen
Kontakt
Spenden

Werkzeuge

In anderen Sprachen
Afrikaans
Alemannisch
Dansk
English
Français
Nederlands
Slovenščina
Tiếng Việt

Hilfeseite **Diskussion**

Lesen

Quelltext anzeigen

Versionsgeschichte

Suche

Hilfe:Personendaten

Abkürzung: Hilfe:PD, WP:PD

Hilfe > Bausteine und Vorlagen > Personendaten

Personendaten sind spezielle Metadaten, die in Artikeln über Personen aufgenommen werden, damit sie daraus automatisch extrahiert und weiterverarbeitet werden können. Sie bestehen aus einer Reihe von Datenfeldern wie beispielsweise Name, Geburtstag und Kurzbeschreibung einer Person.

Bearbeitungsstand: Am 10. August 2009 gab es in 278.677 Artikeln der deutschsprachigen Wikipedia Personendaten. In der englischsprachigen Ausgabe vom 20. Juli 2009 waren in über 43.000 Artikeln Personendaten enthalten.

Inhaltsverzeichnis [Verbergen]

- 1 Kopiervorlage
- 2 Verwendung
- 3 Ansicht einschalten
- 4 Datenfelder
 - 4.1 Name
 - 4.2 Datum
 - 4.3 Alternativnamen
 - 4.4 Kurzbeschreibung
 - 4.5 Geburtsort, Sterbeort
- 5 Zur Geschichte der Personendaten
- 6 Arbeiten mit den Personendaten
 - 6.1 Eintragen neuer Personendaten
 - 6.2 Extraktion der Personendaten
 - 6.3 Herunterladen der Personendaten
 - 6.4 Abfragen und Auswerten von Personendaten
 - 6.5 Wartung der Personendaten
 - 6.6 weitere Verwendungen
- 7 Statistik
- 8 Siehe auch

Hilfe

- Navigation
- Benutzerkonto
- Bearbeiten
- Seitenverwaltung
- Bausteine und Vorlagen
 - Textbausteine
 - Personendaten
 - PND (Namennr.)
 - Navigationsleisten
 - Vorlagen
 - Vorlagenprogrammierung
 - Variablen
 - Eingabefelder
- Bild und Ton
- Software
- FAQ
- Index – Hilfe von A bis Z
- Sitemap
- Wörterliste

Kopiervorlage

```

{{Personendaten
|NAME=
|ALTERNATIVNAMEN=
|KURZBESCHREIBUNG=
|GEBURTSDATUM=
|GEBURTSORT=
|STERBEDATUM=
|STERBEORT=
}}
```

1. Kopiervorlage ans Ende (zwischen die Kategorie- und Interwiki-Links) eines Personenartikels kopieren
2. Datenfelder ausfüllen (siehe Datenfelder). Dabei beachten: Leere Datenfelder (etwa bei lebenden Personen Sterbedatum und Sterbeort) nicht löschen, nur unausgefüllt lassen! Auch wenn nur ein (oder gar kein) Alternativname vorhanden ist, muss das Datenfeld dennoch ALTERNATIVNAMEN im Plural heißen!
3. Fertig



{{Personendaten}}

Die Vorlage sollte am Ende des Artikels stehen, nach allen Weblinks, Navigationsleisten und den Kategorieangaben. Hinter den Personendaten folgen nur noch die Sprachlinks („Interwiki-Links“).

Abbildung 3: Hilfeartikel "Personendaten" in Wikipedia



WIKIPEDIA
Die freie Enzyklopädie

Hauptseite
Über Wikipedia
Themenportale
Von A bis Z
Zufälliger Artikel

Mitmachen
Hilfe
Autorenportal
Letzte Änderungen
Kontakt
Spenden

Drucken/exportieren

Werkzeuge

In anderen Sprachen
Niederlands
Polski

Neue Funktionen Anmelden / Benutzerkonto erstellen

Artikel **Diskussion**

Lesen

Bearbeiten

Versionsgeschichte

Suche



Hermann Grotefend

(Ernst Heinrich) **Hermann Grotefend** (* 18. Januar 1845 in Hannover; † 26. Mai 1931 in Schwerin) war ein deutscher Archivar und Historiker.

Leben und Leistungen [Bearbeiten]

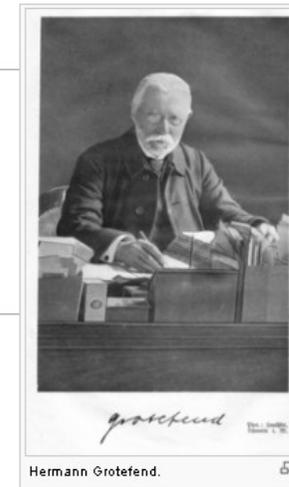
Hermann Grotefend ist der Enkel des Orientalisten und Sprachwissenschaftlers Georg Friedrich Grotefend und der Sohn des Historikers Karl Ludwig Grotefend. Er studierte von 1864 bis 1870 Medizin und anschließend Geschichtswissenschaften. Nach der Promotion arbeitete er am Königlichen Staatsarchiv in Breslau. Von dort wechselte er 1874 nach Aurich, 1876 an das Stadtarchiv Frankfurt am Main und wurde 1887 Leiter des Geheimen und Hauptarchivs Schwerin.

Er veröffentlichte 1891 in Hannover sein *Taschenbuch der Zeitrechnung des deutschen Mittelalters und der Neuzeit*, das bis heute ein unverzichtbares Hilfsmittel ist, historische Datumsangaben in moderne umzurechnen (z. B. aus dem französischen Revolutionskalender, aus dem mittelalterlichen Festkalender oder aus dem römischen Kalender). Dieser Klassiker der historischen Hilfswissenschaften, nach seinem Verfasser meist schlicht „der Grotefend“ genannt, erschien 2007 in 14. Auflage.

Weblinks [Bearbeiten]

Wikisource: Hermann Grotefend – Quellen und Volltexte

- Literatur von und über Hermann Grotefend im Katalog der Deutschen Nationalbibliothek
- Literatur über Hermann Grotefend in der Landesbibliographie MV
- Onlineversion der Zeitrechnung des Deutschen Mittelalters und der Neuzeit



Hermann Grotefend.

Normdaten: PND: 116874732 | LCCN: n85144882 | VIAF: 47524868 | WP-Personeninfo

Kategorien: Archivar | Historiker | Person (Mecklenburg) | Kalender | Deutscher | Geboren 1845 | Gestorben 1931 | Mann

Diese Seite wurde zuletzt am 20. April 2010 um 12:49 Uhr geändert.

Der Text ist unter der Lizenz „Creative Commons Attribution/Share Alike“ verfügbar; zusätzliche Bedingungen können anwendbar sein. Einzelheiten sind in den Nutzungsbedingungen beschrieben. Wikipedia® ist eine eingetragene Marke der Wikimedia Foundation Inc.

Datenschutz Über Wikipedia Impressum



Abbildung 4: Personenartikel der Wikipedia



WIKIPEDIA
Die freie Enzyklopädie

- Hauptseite
- Über Wikipedia
- Themenportale
- Von A bis Z
- Zufälliger Artikel

- ▼ Mitmachen
- Hilfe
- Autorenportal
- Letzte Änderungen
- Kontakt
- Spenden
- Werkzeuge

http://127.0.0.1:4664/?s=8CZ-WJ3_PUedoU6nY1jCI3p3xik

Neue Funktionen Anmelden / Benutzerkonto erstellen

Artikel Diskussion

Lesen Bearbeiten Versionsgeschichte

Bearbeiten von „Hermann Grotefend“

Deine Änderungen werden angezeigt, sobald sie gesichtet wurden. (Hilfe)

Du bearbeitest diese Seite **unangemeldet**. Wenn du deine Änderung speicherst, wird deine aktuelle IP-Adresse in der **Versionsgeschichte** aufgezeichnet und ist damit öffentlich einsehbar. Wenn du ein **Benutzerkonto anlegst**, bleibt deine IP-Adresse verborgen.

Speichere hier bitte keine Textversuche ab. Dafür haben wir unsere [Spielwiese](#).

```

F K      
  ▶ Erweitert ▶ Sonderzeichen ▶ Hilfe

{{Normdaten|PND=116874732|LCCN=n/85/144882|VIAF=47524868}}

{{SORTIERUNG:Grotefend, Hermann}}
[[Kategorie:Archivar]]
[[Kategorie:Historiker]]
[[Kategorie:Person (Mecklenburg)]]
[[Kategorie:Kalender]]
[[Kategorie:Deutscher]]
[[Kategorie:Geboren 1845]]
[[Kategorie:Gestorben 1931]]
[[Kategorie:Mann]]

{{Personendaten
|NAME=Grotefend, Hermann
|ALTERNATIVNAMEN=Grotefend, Ernst Heinrich Hermann
|KURZBESCHREIBUNG=deutscher Archivar
|GEBURTSDATUM=18. Januar 1845
|GEBURTSORT=[[Hannover]]
|STERBEDATUM=26. Mai 1931
|STERBEORT=[[Schwerin]]
}}

[[nl:Hermann Grotefend]]
[[pl:Hermann Grotefend]]

```

Das Kopieren urheberrechtlich geschützter Werke ist verboten!

Bitte gib deine Quellen an!

Mit dem Speichern dieser Seite versichere ich, dass ich den Beitrag selbst verfasst habe bzw. dass er keine fremden Rechte verletzt, und willige ein, ihn unter der [Creative Commons Attribution/Share-Alike Lizenz 3.0](#) und der [GNU-Lizenz für freie Dokumentation](#) zu veröffentlichen. Ich stimme einer Autoren-Nennung mindestens durch URL oder Verweis auf den Artikel zu. Einzelheiten sind in den [Nutzungsbedingungen](#) beschrieben.

Wikipedia-Artikel sollen nur überprüfbare Informationen aus zuverlässigen Quellen enthalten.

Abbildung 5: Personenartikel der Wikipedia im Bearbeitungsmodus mit Metadaten

Dank dieser Personendaten und etlicher rühriger Wikipedia-Autoren⁵ gibt es zahlreiche Möglichkeiten, die Personenartikel der Wikipedia bereits online über das Webinterface, also mit dem eigenen Browser, nach diversen Kriterien abzufragen und auszuwerten.⁶ Weitaus mehr Möglichkeiten der Verwertung hat man jedoch, wenn die komplette Datenbasis der Wikipedia für die weitere freie Verarbeitung zur Verfügung steht. Zu diesem Zweck bietet die Wikimedia-Stiftung als Betreiberin der Wikipedia mehr oder weniger regelmäßig komplette „Dumps“, also Datenbankkopien der Wikipedia, im XML-Format an. Die deutschsprachigen Datenbankauszüge bzw. -kopien finden sich unter <http://dumps.wikimedia.org/dewiki/latest/>. Der unter <http://dumps.wikimedia.org/dewiki/latest/dewiki-latest-pages-articles.xml.bz2> abgelegte Dump hat zurzeit (Januar 2011) als komprimierte Archivdatei einen Umfang von etwa 2 GB.

⁵ Stellvertretend sei hier verwiesen auf Christian Thiele (Pseudonym: APPER), der maßgeblich an der Weiterentwicklung des Projekts „Personendaten“ in der Wikipedia beteiligt ist (<http://de.wikipedia.org/wiki/Benutzer:APPER>), und auf Stefan Kühn, der sich u.a. große Verdienste um die Pflege der Personendaten und der Auswertung von Artikelvorlagen erworben hat (http://de.wikipedia.org/wiki/Benutzer:Stefan_K%C3%BChn).

⁶ Zur Illustration sei hier nur auf einige wenige Seiten verwiesen. Auswertung: <http://de.wikipedia.org/wiki/Wikipedia:Personendaten/Auswertung>; Kategoriensuche: <http://de.wikipedia.org/w/index.php?title=Spezial:Exportieren&action=submit>; Personensuche: <http://toolserver.org/~apper/pd/>.

Index of /dewiki/latest/

Name	Last Modified	Size	Type
Parent Directory/		-	Directory
dewiki-latest-abstract.xml	2010-Sep-02 14:40:39	956.3M	application/xml
dewiki-latest-abstract.xml-rss.xml	2010-Sep-02 14:40:40	0.5K	application/xml
dewiki-latest-all-titles-in-ns0.gz	2010-Sep-02 04:43:02	10.3M	application/octet-stream
dewiki-latest-all-titles-in-ns0.gz-rss.xml	2010-Sep-02 04:43:02	0.6K	application/xml
dewiki-latest-category.sql.gz	2010-Sep-02 04:42:28	1.8M	application/octet-stream
dewiki-latest-category.sql.gz-rss.xml	2010-Sep-02 04:42:28	0.5K	application/xml
dewiki-latest-categorylinks.sql.gz	2010-Sep-02 04:40:37	75.1M	application/octet-stream
dewiki-latest-categorylinks.sql.gz-rss.xml	2010-Sep-02 04:40:37	0.6K	application/xml
dewiki-latest-externallinks.sql.gz	2010-Sep-02 04:41:44	149.7M	application/octet-stream
dewiki-latest-externallinks.sql.gz-rss.xml	2010-Sep-02 04:41:44	0.6K	application/xml
dewiki-latest-flaggedpages.sql.gz	2010-Sep-02 23:22:13	13.8M	application/octet-stream
dewiki-latest-flaggedpages.sql.gz-rss.xml	2010-Sep-02 23:22:13	0.6K	application/xml
dewiki-latest-flaggedrevs.sql.gz	2010-Sep-02 23:23:32	252.3M	application/octet-stream
dewiki-latest-flaggedrevs.sql.gz-rss.xml	2010-Sep-02 23:23:32	0.6K	application/xml
dewiki-latest-image.sql.gz	2010-Sep-02 04:38:12	26.8M	application/octet-stream
dewiki-latest-image.sql.gz-rss.xml	2010-Sep-02 04:38:12	0.5K	application/xml
dewiki-latest-imagelinks.sql.gz	2010-Sep-02 04:40:47	36.9M	application/octet-stream
dewiki-latest-imagelinks.sql.gz-rss.xml	2010-Sep-02 04:40:47	0.5K	application/xml
dewiki-latest-interwiki.sql.gz	2010-Sep-02 04:42:27	7.7K	application/octet-stream
dewiki-latest-interwiki.sql.gz-rss.xml	2010-Sep-02 04:42:27	0.5K	application/xml
dewiki-latest-langlinks.sql.gz	2010-Sep-02 04:42:27	69.1M	application/octet-stream
dewiki-latest-langlinks.sql.gz-rss.xml	2010-Sep-02 04:42:27	0.5K	application/xml
dewiki-latest-md5sums.txt	2010-Sep-15 03:48:28	1.9K	text/plain
dewiki-latest-oldimage.sql.gz	2010-Sep-02 04:38:13	3.5M	application/octet-stream
dewiki-latest-oldimage.sql.gz-rss.xml	2010-Sep-02 04:38:14	0.5K	application/xml
dewiki-latest-page.sql.gz	2010-Sep-02 04:42:55	103.1M	application/octet-stream
dewiki-latest-page.sql.gz-rss.xml	2010-Sep-02 04:42:55	0.5K	application/xml
dewiki-latest-page_props.sql.gz	2010-Sep-02 04:42:56	3.1K	application/octet-stream
dewiki-latest-page_props.sql.gz-rss.xml	2010-Sep-02 04:42:56	0.5K	application/xml
dewiki-latest-page_restrictions.sql.gz	2010-Sep-02 04:42:56	872.6K	application/octet-stream
dewiki-latest-page_restrictions.sql.gz-rss.xml	2010-Sep-02 04:42:56	0.6K	application/xml
dewiki-latest-pagelinks.sql.gz	2010-Sep-02 04:40:13	485.4M	application/octet-stream
dewiki-latest-pagelinks.sql.gz-rss.xml	2010-Sep-02 04:40:14	0.5K	application/xml
dewiki-latest-pages-articles.xml.bz2	2010-Sep-02 19:34:38	1.9G	application/octet-stream
dewiki-latest-pages-articles.xml.bz2-rss.xml	2010-Sep-02 19:34:38	0.6K	application/xml
dewiki-latest-pages-logging.xml.gz	2010-Sep-02 23:21:58	1.0G	application/octet-stream
dewiki-latest-pages-logging.xml.gz-rss.xml	2010-Sep-02 23:21:58	0.6K	application/xml
dewiki-latest-pages-meta-current.xml.bz2	2010-Sep-02 21:51:37	2.9G	application/octet-stream
dewiki-latest-pages-meta-current.xml.bz2-rss.xml	2010-Sep-02 21:51:37	0.6K	application/xml
dewiki-latest-pages-meta-history.xml.7z	2010-Sep-15 03:45:38	7.7G	application/x-7z-compressed
dewiki-latest-pages-meta-history.xml.7z-rss.xml	2010-Sep-15 03:45:38	0.6K	application/xml
dewiki-latest-pages-meta-history.xml.bz2	2010-Sep-09 13:55:21	87.3G	application/octet-stream
dewiki-latest-pages-meta-history.xml.bz2-rss.xml	2010-Sep-09 13:55:21	0.6K	application/xml
dewiki-latest-protected_titles.sql.gz	2010-Sep-02 04:42:56	251.3K	application/octet-stream
dewiki-latest-protected_titles.sql.gz-rss.xml	2010-Sep-02 04:42:56	0.6K	application/xml
dewiki-latest-redirect.sql.gz	2010-Sep-02 04:42:58	7.5M	application/octet-stream
dewiki-latest-redirect.sql.gz-rss.xml	2010-Sep-02 04:42:58	0.5K	application/xml
dewiki-latest-site_stats.sql.gz	2010-Sep-02 04:38:06	0.7K	application/octet-stream
dewiki-latest-site_stats.sql.gz-rss.xml	2010-Sep-02 04:38:06	0.5K	application/xml
dewiki-latest-stub-articles.xml.gz	2010-Sep-02 17:51:13	140.8M	application/octet-stream

<http://dumps.wikimedia.org/dewiki/latest/dewiki-latest-pages-articles.xml.bz2>

Ein aktiver Download (1 Stunde, 26 Minuten verbleibend)

Abbildung 6: Downloadseite für den Wikipedia Dump

Mit dem unter <http://www.7-zip.org/> frei zugänglichen Packprogramm 7-Zip kann die Archivdatei ausgepackt werden. Die extrahierte XML-Datei hat schließlich eine Größe von ca. 7,7 GB und beinhaltet den kompletten Text der deutschsprachigen Wikipedia in der zum Zeitpunkt der Dump-Erstellung aktuellen Fassung. Er beinhaltet indessen keine Versionsgeschichte.

```
- <mediawiki xmlns="http://www.mediawiki.org/xml/export-0.4/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.mediawiki.org/xml/export-0.4/ http://www.mediawiki.org/xml/export-0.4.xsd" version="0.4" xml:lang="de">
- <siteinfo>
  <sitename>Wikipedia</sitename>
  <base>http://de.wikipedia.org/wiki/Wikipedia:Hauptseite</base>
  <generator>MediaWiki 1.16wmf4</generator>
  <case>first-letter</case>
- <namespaces>
  <namespace key="-2" case="first-letter">Media</namespace>
  <namespace key="-1" case="first-letter">Spezial</namespace>
  <namespace key="0" case="first-letter" />
  <namespace key="1" case="first-letter">Diskussion</namespace>
  <namespace key="2" case="first-letter">Benutzer</namespace>
  <namespace key="3" case="first-letter">Benutzer Diskussion</namespace>
  <namespace key="4" case="first-letter">Wikipedia</namespace>
  <namespace key="5" case="first-letter">Wikipedia Diskussion</namespace>
  <namespace key="6" case="first-letter">Datei</namespace>
  <namespace key="7" case="first-letter">Datei Diskussion</namespace>
  <namespace key="8" case="first-letter">MediaWiki</namespace>
  <namespace key="9" case="first-letter">MediaWiki Diskussion</namespace>
  <namespace key="10" case="first-letter">Vorlage</namespace>
  <namespace key="11" case="first-letter">Vorlage Diskussion</namespace>
  <namespace key="12" case="first-letter">Hilfe</namespace>
  <namespace key="13" case="first-letter">Hilfe Diskussion</namespace>
  <namespace key="14" case="first-letter">Kategorie</namespace>
  <namespace key="15" case="first-letter">Kategorie Diskussion</namespace>
  <namespace key="100" case="first-letter">Portal</namespace>
  <namespace key="101" case="first-letter">Portal Diskussion</namespace>
</namespaces>
</siteinfo>
- <page>
  <title>Alan Smithee</title>
  <id>1</id>
- <revision>
  <id>77481336</id>
  <timestamp>2010-08-05T18:42:20Z</timestamp>
- <contributor>
  <username>PhChAK</username>
  <id>806501</id>
</contributor>
<comment>Eigenen IMDb-Eintrag als Quelle und weitere Schreibweisen sowie Anmerkung, daß nicht alles Anagramme ergänzt</comment>
<text xml:space="preserve">""Alan Smithee"" ist ein [[Anagramm]] von „The Alias Men“<ref name="IMDb">[http://www.imdb.com/name/nm0000647/
Eigener Eintrag für "Alan Smithee" in der IMDb]</ref> und wird als Name im Abspann von [[Film]]en verwendet, in denen der eigentliche [[Regisseur]]
seinen Namen nicht mit dem Werk in Verbindung gebracht haben möchte. Alternative Schreibweisen mit dem gleichen Ziel, die aber keine Anagramme
von „The Alias Men“ sind, sind unter anderem „Allen Smithee“, „Alan Smythee“, und „Adam Smithee“. Auch zwei teilweise asiatisch anmutende
Schreibweisen „Alan Smi Thee“ und „Sumishii Aran“ sollen - so die Internet Movie Database - dazu gehören.<ref name="IMDb" /> == Entstehung ==
Das Pseudonym "Alan Smithee" entstand 1969, als keiner der beiden Regisseure des Filmes [[Death of a Gunfighter]] ([[Robert Totten]]) wurde durch Don
```

Abbildung 7: Wikipedia im XML-Format

Das Auslesen der Metadaten aus den Personenartikeln geschieht mittels eines Perlskripts, das von der Seite [http://de.wikipedia.org/wiki/Hilfe:Personendaten/Datenextraktion#Extraktion aus einem XML-Dump mittels Perl-Skript](http://de.wikipedia.org/wiki/Hilfe:Personendaten/Datenextraktion#Extraktion_aus_einem_XML-Dump_mittels_Perl-Skript) kopiert und als einfache Textdatei abgespeichert werden kann.⁷ Leicht verständliche Erläuterungen zum Absetzen des Skriptbefehls finden sich ebenfalls auf dieser Seite.

```
printf "Name\tAlternativnamen\tkurzbeschreibung\tgeburtsdatum\tgeburtsort\tsterbedatum\tsterbeort\n"; #Tabellenkopf
while ($zeile = <>) #Die folgenden Aktionen laufen für jede Zeile des XML-Dumps
{
    if ($zeile =~ m/{Personendaten/}
    {
        if ($offen == 1)
        {
            printf "Syntaxfehler in der Datenquelle. Offene Personendaten nicht geschlossen. $zeile\n" ;
            $fehlerzaehler = $fehlerzaehler + 1;
            $zaehler = $zaehler + 1;
        }
        else
        {
            $offen = 1;
            $zaehler = $zaehler + 1;
        }
    }
    if ($offen == 1)
    {
        while($zeile =~ m/(.*)\[\[\[([+\\|]?(.*)\]\]\](.*)/) #entfernt wiki-Links, solange noch welche in der Zeile
        {
            $zeile = $1.$3.$4;
        }
        if($zeile =~ m/NAME=(.*)/)
        {
            $name = "$1";
        }
        if($zeile =~ m/ALTERNATIVNAMEN=(.*)/)
        {
            $alternativnamen = "$1";
        }
        if($zeile =~ m/KURZBESCHREIBUNG=(.*)/)
        {
            $beschreibung = "$1";
        }
        if($zeile =~ m/GEBURTSDATUM=(.*)/)
        {
            $geburtsdatum = "$1";
        }
        if($zeile =~ m/GEBURTSORT=(.*)/)
        {
            $geburtsort = "$1";
        }
        if($zeile =~ m/STERBEDATUM=(.*)/)
        {
            $sterbedatum = "$1";
        }
        if($zeile =~ m/STERBEORT=(.*)/)
        {
            $sterbeort = "$1";
        }
    }
}
if ($zeile =~ m/})/ && $offen == 1)
{
    $offen = 0;
    printf "$name\t$alternativnamen\t$beschreibung\t$geburtsdatum\t$geburtsort\t$sterbedatum\t$sterbeort\n"; #Datenzeile
}
printf "Personen: $zaehler\n";
if($fehlerzaehler > 0)
```

Abbildung 8: Perlskript zum Auslesen der Personendaten

⁷ Auf Windows Betriebssystemen muss zusätzlich eine Perl-Distribution wie beispielsweise „Strawberry Perl“ aus dem gleichnamigen Open-Source-Projekt installiert werden (vgl. <http://strawberryperl.com/support.html>).

Mit dem Auslesen der Metadaten erzeugt das Perl-Skript eine ca. 46 MB große datenbankgerechte Textdatei im UTF-8-Code mit Tabulatoren als Feldbegrenzungen. Stand 12. Januar 2011 enthält die Datei 356259 Datensätze mit Personendaten zu den Personenartikeln der Wikipedia.⁸

Name	Alternativnamen	Kurzbeschreibung	Geburtsdatum	Geburtsort	Sterbedatum	Sterbeort		
Bechstein, Ludwig		deutscher Schriftsteller, Bibliothekar und Archivar	17. Januar 1860	Esslingen am Neckar	24. November 1901	Weimar	14. Mai 1860	Meiningen
Rarisch, Klaus M.		deutscher Lyriker	17. Januar 1936	Berlin				
Vlk, Miloslav		tschechischer Geistlicher, Bischof von Budweis, Erzbischof von Prag, Kardinal	17. Mai 1932	Lisnice bei Milevsko				
Genz, Albert		deutscher Architekt	1877	Esslingen am Neckar	nach 4. April 1944			
Kaufmann, Alexander		deutscher Schriftsteller und Archivar	14. Mai 1817	Bonn	1. Mai 1893			Werth
Lisch, Georg Christian Friedrich		deutscher Prähistoriker, Archivar, Bibliothekar, Konservator, Heraldiker	29. März 1801					
Wahle, Julius		österreichischer Literaturwissenschaftler und Herausgeber der Briefe Johann Wolfgang von Goethes	15. Februar 1861	Wien				
Koldewey, Ehrenreich Gerhard		deutscher Gelehrter, Jurist, Poet und Kenner der ostfriesischen Geschichte	7. August 1702	Oldenburg				
Kestner, Johann Christian		deutscher Jurist und Archivar, Ehemann von Charlotte Buff	28. August 1741	Hannover				Lünebu
Stenzel, Gustav Adolf Harald		deutscher Geschichtsforscher	21. März 1792	Zerbst	3. Januar 1854	Breslau		
Barthel, Ludwig Friedrich		deutscher Erzähler und Essayist	12. Juni 1898	Marktbreit	14. Februar 1962			München
Vehse, Karl Eduard		deutscher Geschichtsschreiber	18. Dezember 1802	Freiberg (Sachsen)			18. Juni 1870	Striesen
Tekath, Karl-Heinz		deutscher Historiker, Archivar und Vorkämpfer der deutsch-niederländischen Freundschaft	23. August 1955	Labbeck	8. Dezember 2006			
Ledderhose, Konrad Wilhelm		deutscher Jurist	23. Dezember 1812	Hanau				
Stickler, Alfons Maria		österreichischer Kardinal der römisch-katholischen Kirche	23. August 1910	Neunkirchen (Niederösterreich)	12. Dezember 1988			
Delius, Christian Heinrich		deutscher Archivar und Historiker	24. Oktober 1778	wernigerode			14. April 1840	wernigerode
Montelius, Oscar		schwedischer Reichsarchivar und Prähistoriker	9. September 1843	Stockholm				
Goldwin, Sergius		Schweizer Publizist und Mythenforscher	31. Januar 1930	Prag, Tschechien			17. Juli 2006	Bern, Schweiz
Poggi, Luigi		italienischer Geistlicher, Kardinalprotodiakon der katholischen Kirche	25. November 1917	Piacenza, Italien			4. Mai 2010	
Beyer, Heinrich Beyer, Christian Heinrich		deutscher Archivar und Historiker	4. August 1806	Erfurt	16. September 1886			Stolberg (Harz)
Capriv, Julius Leopold von		deutscher Archivar, Historiker und Dichter von Kirchenliedern	29. August 1695	Landau, Schlesien			8. November 1771	
Meisner, Heinrich Otto		deutscher Archivar und Historiker	1. April 1890	Berlin	26. November 1976			Potsdam
Kretschmar, Hellmut		deutscher Archivar und Historiker	12. Juli 1893	Stolpen	2. Dezember 1965			Dresden
Beschoner, Hans		deutscher Archivar und Historiker	5. April 1872	Dresden	30. Januar 1956			Presden
Lieberich, Heinz		deutscher Rechtshistoriker und Archivar	29. Januar 1905	Kaiserslautern	24. Oktober 1999			Mosbach
Gerlich, Fritz Gerlich, Carl Albert Fritz		deutscher Journalist und Historiker	15. Februar 1883	Stettin	30. Juni 1934			KZ Dachau
Friedensburg, Walter		deutscher Archivar und Historiker	6. März 1855	Hamburg	19. Februar 1938			wernigerode
Grotefend, Ernst Heinrich Hermann		deutscher Archivar	28. Januar 1845	Hannover			26. Mai 1931	Schwerin
Stock, Ludwig Stock, Christian Ludwig		deutscher Archivar und Historiker	14. November 1778	Halle (Saale)			26. Mai 1861	wernigerode
Glafey, Adam Friedrich von		deutscher Archivar und Rechtshistoriker	14. Juli 1692	Reichenbach im Vogtland	14. Juli 1753			Dresden
Fries, Lorenz		Würzburger Fürstbischöflicher Rat, Geschichtsschreiber	24. Juni 1489	Bad Mergentheim	5. Dezember 1550			Würzburg
Pertz, Georg Heinrich		deutscher Historiker	28. März 1795	Hannover			7. Oktober 1876	München
Ausfeld, Eduard Ausfeld, Heinrich Eduard		deutscher Archivar und Historiker	27. Mai 1850	Schnepfenthal	4. April 1906			Magdeburg
Vollmer, Gisela		deutsche Historikerin und Archivarin	23. Oktober 1922	Düsseldorf	4. März 2005			Düsseldorf
Wahl, Volker		deutscher Archivar und Historiker	10. Juni 1943	Steinbach-Hallenberg				
Ermisch, Hubert Maximilian		deutscher Archivar und Historiker	23. Juni 1850	Torgau	6. April 1932			Dresden
Michelot, Jules		französischer Historiker	21. August 1798	Paris, Frankreich	9. Februar 1874			Hyères, Frankreich
Schmidt, Michael Ignaz		deutscher Historiker	30. Januar 1736	Arnstein (Unterfranken)	1. November 1794			Wien
Schwarzbaum, Haim		israelischer Orientalist und Erzählforscher	24. September 1911	Warschau			11. Oktober 1983	Kiron
Beck, Friedrich		deutscher Archivar und Historiker	20. Juni 1927	Greiz-Untergrochlitz				
Rasch, Manfred		deutscher Wirtschaftshistoriker und Leiter des Thyssen-Archivs in Duisburg	1955					
Brachmann, Bocho		deutscher Archivar und Historiker	2. Oktober 1930	Berlin-Pankow				
Winter, Georg Winter, Carl Georg (vollständiger Name)		deutscher Archivar und Historiker	3. Februar 1856	Breslau	1. September 1912			Magdeburg
Brichzin, Hans		deutscher Archivar und Autor	5. Juni 1938	Stargard				
Nedo, Michael		Direktor des Wittgensteinarchivs in Cambridge	1940					
Hempel, Wolfgang		deutscher Archivar	14. Oktober 1931	Minden				
Weber, Hartmut		deutscher Archivar, Präsident des Bundesarchivs	März 1945	Ohringen				
Augustin, Milan		tschechischer Archivar und Historiker	1960	Karlovy Vary				
Höbar, Hermann		deutscher katholischer Geistlicher, Kirchenhistoriker	11. Dezember 1907	Osnabrück			21. September 1992	Osnabrück
Beneke, Otto		Hamburger Archivar und Schriftsteller	5. Oktober 1812	Hamburg	9. Februar 1891			
Leverkus, Wilhelm		deutscher Archivar und Politiker	13. Februar 1808	Wermelskirchen	30. November 1870			Oldenburg (Oldenburg)
Mandyczewski, Eusebius		Mandicevski, Eusebie österreichischer Musikwissenschaftler und Komponist	18. August 1857	Czernowitz			13. Juli 1929	Wien
Holz, Walter K. B.		deutscher Archivar	11. Mai 1908	24. Juni 1995				
Schmidt, Heinrich		deutscher Archivar und Professor	18. Dezember 1874	Heubach	2. Mai 1935			Jena
Oberdorffer, Kurt		deutscher Historiker, nationalsozialistischer Historiker und Archivar	28. April 1900	Schluckenau	10. November 1980			
Schnack, Georg		deutscher Historiker	6. November 1898	Hannover				
Mötsch, Johannes		deutscher Archivar und Historiker	8. Juli 1949	Bonn				
Weber, Karl von		deutscher Historiker	1. Januar 1806	Dresden	18. Juli 1863			Dresden-Loschwitz
Mülverstedt, George Adalbert von		deutscher Archivar und Historiker	4. Juli 1825	Neufahrwasser	29. September 1914			Magdeburg
Decker-Hauff, Hansmartin		deutscher Historiker und Genealoge	29. Mai 1917	Zettlingen	31. März 1992			Stuttgart
Forstmann, Ernst		deutscher Archivar, Bibliothekar und Historiker	18. September 1822	Danzig	4. November 1906			
Langlois, Henri		französischer Filmarchivar	13. November 1914	Smyrna, heute Izmir, Türkei	13. Januar 1977			Paris, Frankreich
Huser, Fritz		deutscher Bibliothekar	4. Oktober 1908	Heißen (heute Mülheim an der Ruhr)	4. März 1979			Dortmund
Albinus, Petrus Albinus, Peter; weiße, Peter von (eigentlicher Name)		deutscher Historiker, Heimatforscher und Poet	18. Juni 1543	Schneeberg (Erzgebirg)				
KlÜbendorff, Niklot		deutscher Numismatiker	10. Februar 1944	Hamburg				
Wutke, Konrad		deutscher Archivar und Historiker	2. Oktober 1861	Berlin	6. April 1951			Uttenhofen bei Augsburg
Posse, Otto		deutscher Historiker und Archivar	29. Juli 1847	Weißensee	13. November 1921			Dresden
Schulte, Aloys		deutscher Historiker und Archivar	2. August 1857	Münster (Westfalen)	14. Februar 1941			Bonn
Lehmann, Rudolf Lehmann, Friedrich Adolf Rudolf		deutscher Historiker und Archivar	16. September 1891	Staßfurt			14. Januar 1984	Marburg
Ennen, Edith		deutsche Historikerin und Archivarin	28. Oktober 1907	Bonn				
Sonnleitner, Joseph		Sonnleitner, Joseph von; Sonnleitner, Joseph; Sonnleitner, Joseph Ferdinand österreichischer Librettist, Komponist, Theaterdirekt						

Abbildung 9: Ergebnis der Datenextraktion mittels Perlskript

⁸ Zur Veranschaulichung der Dynamik, mit der die Personenartikel in der deutschsprachigen Wikipedia anwachsen, seien zum Vergleich die Zahlen der vorangegangenen Dumps genannt: 26.3.2010: 315795, Stand 16.6.2010: 326734; Stand 2.9.2010: 341150, Stand 25.9.2010: 344855, Stand 13.10.2010: 347745.

Diese Textdatei lässt sich anschließend in jedes leistungsfähige Datenbank-Programm einer handelsüblichen Office-Anwendung wie Microsoft Access 2007 importieren.

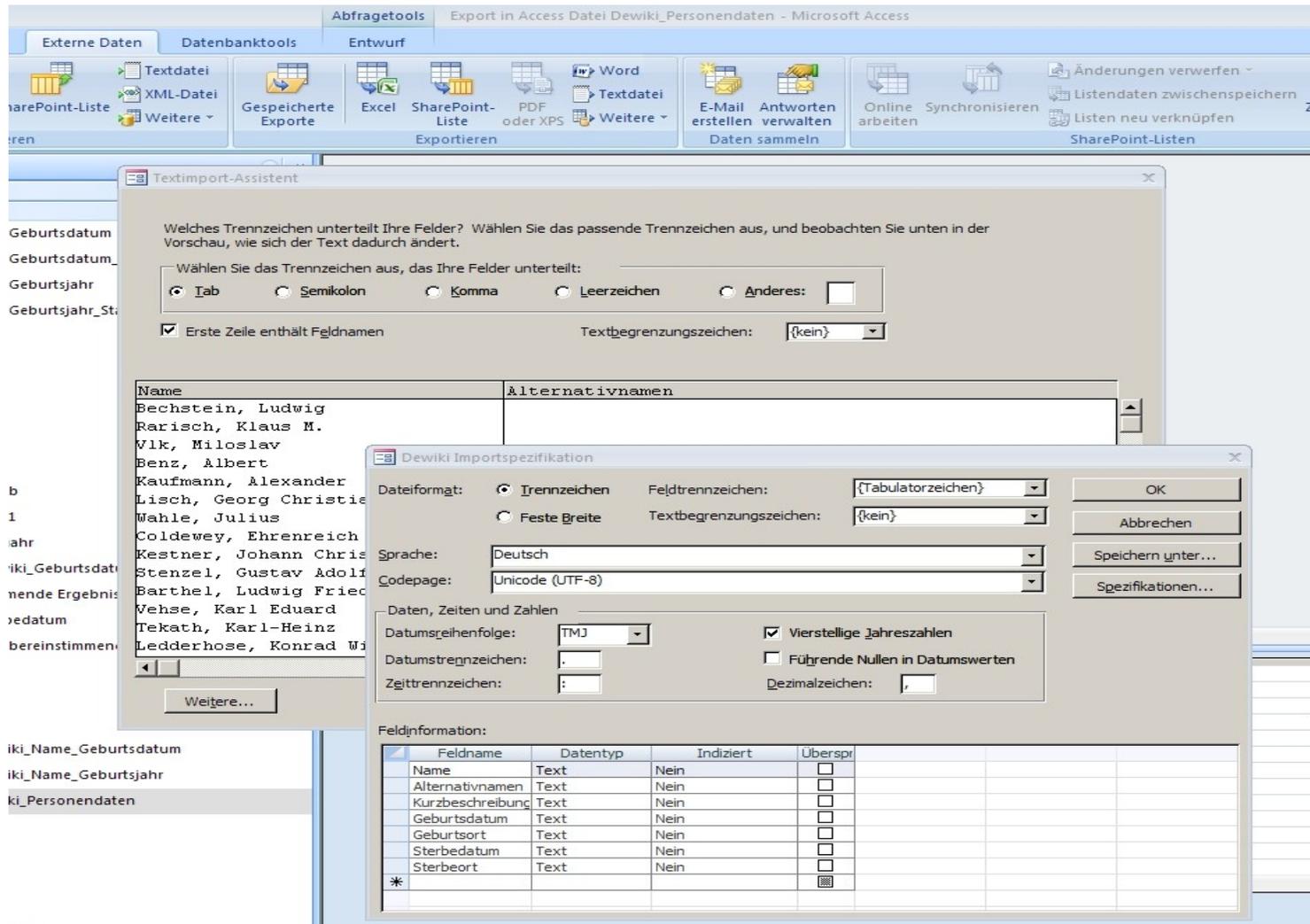


Abbildung 10: Import der Textdatei

Dabei sollte man darauf achten, alle Felder im Datentyp „Text“ zu importieren. Erst nach dem Import sollte man die Felder „Geburtsdatum“ und „Sterbedatum“ in den Datentyp „Datum“ ändern.

ID	Name	Alternativnamen	Kurzbeschreibung	Geburtsdatum	Geburtsort	Sterbedatum	Sterbeort
1	Lee, Ang		taiwanischer Regisseur und Drehbuchautor	23. Oktober 1954	Pingtung, Taiwan		
2	Minghella, Anthony		britischer Dramen- und Drehbuchautor, Filmregisseur	6. Januar 1954	Isle of Wight	18. März 2008	London
3	Hitchcock, Alfred	Hitchcock, Alfred Joseph (vollständiger Name)	britischer Filmregisseur und Produzent	13. August 1899	London, Großbritannien	29. April 1980	Los Angeles, Kalifornien
4	Kaurismäki, Aki		finnischer Regisseur	4. April 1957	Orimattila, Finnland		
5	Pacino, Al	Pacino, Alfredo James (vollständiger Name)	US-amerikanischer Schauspieler, Regisseur und Drehbuchautor	25. April 1940	New York City		
6	Alexander der Große	Alexander III.	makedonischer Feldherr und König	zwischen 20. Juli 356 v. Chr. und 13. Juni 323 v. Chr.	Pella	10. Juni 323 v. Chr.	Babylon
7	Hope, Anthony	Hawkins, Anthony Hope (wirklicher Name)	englischer Rechtsanwalt und Autor	9. Februar 1863	London	8. Juli 1933	
8	Turing, Alan	Turing, Alan Mathison (vollständiger Name)	britischer Logiker, Mathematiker und Kryptoanalytiker	23. Juni 1912	London	7. Juni 1954	Wilmslow, Cheshire
9	Harris, Arthur	Harris, Arthur Travers	britischer Luftmarschall, Oberkommandierender der RAF	13. April 1892	Cheltenham	5. April 1984	Henley-on-Thames
10	Wellesley, Arthur, 1. Duke of Wellington	Wellesley, Arthur; Wesley, Arthur	britischer Feldmarschall und Staatsmann	30. April 1769 oder 1. Mai 1770	Dublin, Irland	14. September 1852	Walmer Castle, Kent
11	Jolie, Angelina	Voight, Angelina Jolie (Geburtsname)	US-amerikanische Schauspielerin	4. Juni 1975	Los Angeles, Kalifornien		
12	Archimedes	Archimedes von Syrakus	griechischer Mathematiker, Physiker und Ingenieur	um 287 v. Chr.	unsicher: Syrakus auf Sizilien	212 v. Chr.	Syrakus auf Sizilien
13	Aristoteles	Der Stagirit; Der Philosoph	griechischer Philosoph und Naturforscher	384 v. Chr.	Stageira, Chalkidike	322 v. Chr.	Chalkis, Euböa
14	Lincoln, Abraham		16. Präsident der Vereinigten Staaten von Amerika	12. Februar 1809	bei Hodgenville, Kentucky	15. April 1865	Washington D.C.
15	Merkel, Angela	Merkel, Angela Dorothea (vollständiger Name)	deutsche Politikerin (CDU), Bundeskanzlerin der BRD	17. Juli 1954	Hamburg		
16	Silverstone, Alicia		US-amerikanische Schauspielerin	4. Oktober 1976	San Francisco		
17	Al-Biruni	al-Bīrūnī, Abu al-Raiḥān Muḥammad ibn Ahmad	persischer Mathematiker, Wissenschaftler und Geograph	4. September 973	Kath, Zentralasien (deutlich)	9. Dezember 1048	Ghazni (damals)
18	Seghers, Anna	Radványi, Netty (eigentlicher Name); Reiling, Netty	deutsche Schriftstellerin	19. November 1900	Mainz	1. Juni 1983	Berlin
19	Pärt, Arvo		estnischer Komponist	11. September 1935	Paide		
20	Crowley, Aleister	Crowley, Edward Alexander (wirklicher Name); Hecateus	Okkultist, Mystiker, Poet, Künstler, Künstler-Manager	12. Oktober 1875	Leamington Spa	1. Dezember 1947	Hastings
21	Haigis, Anne		deutsche Musikerin und Sängerin	9. Dezember 1955	Rottweil		
22	Augustus	Octavius Thurinus, Gaius; Octavian; Oktavian	gilt als erster römischer Kaiser	23. September 63 v. Chr.	Rom oder Velletri	19. August 14 v. Chr.	Nola bei Neapel
23	Schopenhauer, Arthur		deutscher Philosoph	22. Februar 1788	Danzig	21. September 1860	Frankfurt am Main
24	Aemilianus	Aemilius Aemilianus, Marcus	römischer Kaiser	207 oder 214	Djerba	253	
25	Aurelian	Aurelianus, Lucius Domitius	römischer Kaiser (270–275)	9. September 214	Moesia oder Sirmium	275	bei Konstantinopel
26	Linde, Andrei Dmitrijewitsch	Линде, Андрей Дмитриевич	russischer Kosmologe	2. März 1948	Moskau, Russland		
27	Hitler, Adolf		deutscher Reichskanzler, Diktator und Kriegsverbrecher	20. April 1889	Braunau am Inn	30. April 1945	Berlin
28	Arcadius, Flavius	Arkadios	Kaiser des Öströmischen Reiches (395–408)	um 377	Spanien	408	Konstantinopel
29	Avitus	Eparchius Avitus	weströmischer Kaiser	um 385	Auvergne	457	
30	Christie, Agatha	Christie, Agatha Mary Clarissa; Mary Westmacott (Pseudonym)	englische Krimi-Schriftstellerin	15. September 1890	Torquay, England	12. Januar 1976	Wallingford, England
31	Nobel, Alfred	Nobel, Alfred Bernhard	schwedischer Chemiker und Erfinder	21. Oktober 1833	Stockholm	10. Dezember 1896	Sanremo
32	Ries, Adam	Riese, Adam; Ris, Adam; Rise, Adam; Ryse, Adam	deutscher Rechenmeister	1492 oder 1493	Staffelstein, Oberfranken	30. März 1559 oder 1560	unsicher: Annaberg-Buchholz
33	Ampère, André-Marie		französischer Physiker und Mathematiker	20. Januar 1775	Poleymieux-au-Mont-d'Or	10. Juni 1836	Marseille
34	Celsius, Anders		schwedischer Astronom, Mathematiker und Physiker	7. Dezember 1701	Uppsala	6. Mai 1744	Uppsala
35	Droste-Hülshoff, Annette	Droste-Hülshoff, Annette Elisabeth Freilin von (vollständiger Name)	deutsche Schriftstellerin	10. Januar 1797	Burg Hülshoff bei Münster	24. Mai 1848	Meersburg
36	Albrecht von Brandenburg	Albrecht von Hohenzollern	Hochmeister des Deutschen Ordens, erster Herzog von Preußen	17. Mai 1490	Ansbach	20. März 1568	Burg Tapiau
37	Fleming, Alexander		schottischer Bakteriologe und Nobelpreisträger, erfinder des Penicillins	6. August 1881	Lochfield, Darvel, East Lothian	11. März 1955	London
38	Lovelace, Ada	Augusta Ada King Byron, Countess of Lovelace	britische Mathematikerin	10. Dezember 1815	London	27. November 1852	London
39	Dumas, Alexandre der Jüngere		französischer Schriftsteller	27. Juli 1824	Paris	27. November 1895	Marly-le-Roi
40	Vesalius, Andreas	Vesal, Andreas; Witinck, Andreas	flämischer Anatom; gilt als Begründer der neuzeitlichen Anatomie	31. Dezember 1514	Brüssel	15. Oktober 1564	auf Zakynthos
41	Scharon, Ariel	Scheinermann, Ariel (Geburtsname); Arik; Sharon, Arieh	israelischer General und Politiker	27. Februar 1928	Kfar Malal, Scharonbezirk		
42	Arnim, Achim von	Arnim, Carl Joachim Friedrich Ludwig von (vollständiger Name)	deutscher Dichter der Heidelberger Romantik	26. Januar 1781	Berlin	21. Januar 1831	auf Schloss Wilmersdorf
43	Volta, Alessandro	Volta, Alessandro Giuseppe Antonio Anastasio Grazioplene	italienischer Physiker, Begründer der Elektrizitätslehre	18. Februar 1745	Como	5. März 1827	Camnago bei Varese
44	Leeuwenhoek, Antoni van	Leeuwenhoek, Antonie van; Leeuwenhoek, Anton	niederländischer Naturforscher und Mikroskopbauer	24. Oktober 1632	Delft, Niederlande	27. August 1723	Delft, Niederlande
45	Gide, André	Gide, André Paul Guillaume	französischer Schriftsteller; Literaturnobelpreisträger	22. November 1869	Paris	19. Februar 1951	Paris
46	Anaximander	Anaximandros; Αναξίμανδρος (altgriechisch)	vorokzratlicher griechischer Philosoph	um 610 v. Chr.	Milet	nach 547 v. Chr.	Milet
47	Demetrius I. Poliorketes	Δημήτριος Α΄ ο Πολιορκητής	seleukidischer König	337 v. Chr.		854 v. Chr.	

Abbildung 11: Die Personendaten aus Wikipedia nach dem Import in der Datenbank

Die importierte Tabelle kann anschließend mit ebenfalls in die Datenbank zu importierenden Tabellen (z.B. Ablieferungsverzeichnissen von Behörden) abgeglichen werden. In der Regel sind zuvor jedoch wegen der erforderlichen Datenkongruenz kleinere Anpassungen bei den Namensfeldern und den Geburtsdaten in der jeweiligen Vergleichstabelle erforderlich. Für den Abgleich ist eine einfache Abfrage zu erstellen, in der nach der Kongruenz von Namen, Vornamen und Geburtsdatum gefragt wird. In Zweifelsfällen können die zusätzlichen Angaben wie Kurzbeschreibung, Geburtsort, Sterbeort oder Sterbedatum zur weiteren Verifizierung beitragen.

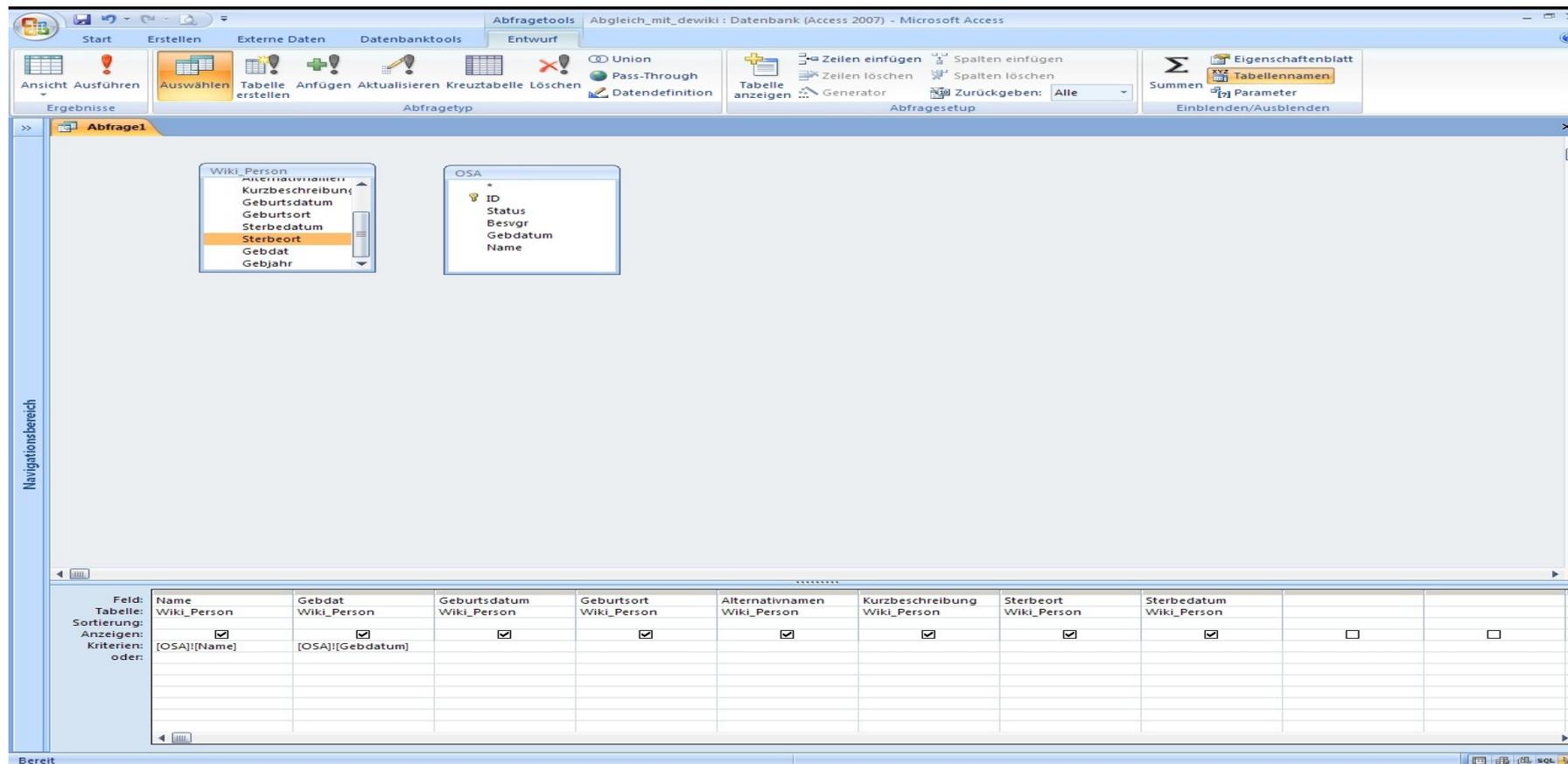


Abbildung 12: Abfrage für den Datenabgleich

OSA	Familienname	Geburtsdatum	Kurzbeschreibung
Ki		8.6.1951	deutscher Jazzmusiker, Autor und Fotograf
Le		30.4.1972	deutscher Maler, Zeichner, Illustrator
Le		24.7.1968	deutscher Filmkomponist
M		13.12.1937	deutscher Kinderbuchautor, Illustrator, Übersetzer, Drehbuch- und Theaterautor
M		27.8.1927	deutscher Schriftsteller und Gymnasialprofessor
M		6.3.1924	deutscher Politiker der FDP/DVP und Sport- und Vertriebenenfunktionär
M		2.2.1948	deutscher Degenfechter
M		25.12.1942	deutscher Realschullehrer und Mundartlyriker
N		29.9.1926	deutscher Agrarwissenschaftler und Politiker (SPD), MdB
N		16.2.1940	deutscher Pädagoge und Theologe
P		18.8.1967	deutscher Fußballspieler
R		20.5.1965	baden-württembergische Politikerin (CDU)
R		25.9.1940	deutscher Politiker (SPD)
R		18.12.1949	deutscher Fußballtrainer und Fußballprofi
S		30.12.1960	deutsche Volleyball-Nationalspielerin
S		8.3.1951	deutscher Politiker (SPD)
S		9.12.1939	deutscher Jazzmusiker
Si		22.6.1934	deutscher Pädagoge und Politiker (SPD), MdB
St		28.7.1947	deutscher Politiker (SPD)
St		19.7.1946	deutscher Philosoph
St		17.8.1952	deutscher Pädagoge
Ti		27.11.1941	deutscher Schriftsteller
V		7.2.1940	deutscher Sonderpädagoge
W		14.12.1971	deutscher Schriftsteller
2			
A		22.6.1944	deutsche Politikerin (SPD), Baden-Württembergische Landtags- und Bundestagsabgeordnete
A		6.6.1952	deutscher Ornithologe
B		1.10.1952	deutsche Politikerin (CDU), MdL
D		9.1.1940	deutscher Politiker (NPD)
D		22.9.1944	deutscher Mundartautor und -dichter
D		4.5.1975	deutscher Fußballspieler
E		17.1.1948	deutsche Schriftstellerin

Donnerstag, 15. Juli 2010

SEITE 2 VON 6

Abbildung 13: Ergebnisliste eines Abgleichs mit Personalakten

Die Erfahrungen im Staatsarchiv Sigmaringen mit dem Abgleich elektronischer Ablieferungsverzeichnisse waren durchweg positiv. Die anfänglichen Investitionen in die Datengewinnung und –bearbeitung haben sich gelohnt. Mittlerweile ist mit den Lehrerpersonalakten des ehemaligen Oberschulamts Tübingen das nächste große Bewertungsprojekt in Angriff genommen, das elektronisch unterstützt wird. Selbst in der Nachbewertung von Beständen, über deren Archivwürdigkeit vor Jahrzehnten noch kein abschließendes Urteil gefällt werden konnte, hat sich das Verfahren bewährt.

The screenshot displays the 'Intranet-Findbuch' interface for the Landesarchiv Baden-Württemberg, Staatsarchiv Sigmaringen. The main content area shows the 'Bestand Wü 42 T 127' in 'Strukturansicht' mode, displaying a list of 5 entries under the heading 'Personalakten: 1. Qualitative Auswahl'. The entries are:

- Wü 42 T 127 Nr. 3971**
Enders, Wilhelm (Regierungsdirektor und Ministerstellvertreter im Landwirtschaftsministerium
 Württemberg-Hohenzollern)
 geb. am 26. Mai 1901
 Bem.: Bewertungsgrundlage: Autopsie
- Wü 42 T 127 Nr. 4027**
Ei [redacted] (Ministerialdirektor und Amtschef des Ministeriums für Wissenschaft und Kunst
 Baden-Württemberg; Senator für Wissenschaft und Forschung Berlin)
 geb. am [redacted]
 Bem.: Bewertungsgrundlage: Wikipedia und Personendatenbank Landesbibliographie Baden-
 Württemberg
- Wü 42 T 127 Nr. 1**
Erlar, Fritz (Vorsitzender der SPD-Fraktion im Deutschen Bundestag Landrat)
 geb. am 14. Juli 1913 gest. am 22. Februar 1967
 Bem.: Bewertungsgrundlage: Autopsie

The interface includes a navigation menu on the left with folders for letters A through Z, and a footer with links: Anfang, Einführung, Stichwortlisten, Erweiterte Suche, Hilfe, Beständeübersicht, Übergreifende Suche, Signatursuche, Bestellkorb, and Lesezeichen.

Abbildung 14: Intranet-Findbuch mit automatisiert ermittelten Personalakten

Selbstverständlich ist der automatisierte Abgleich von behördlichen Aussonderungslisten nicht der Königsweg in der Bewertung personenbezogener Unterlagen. Er befreit nicht von der Verpflichtung, die Methoden sowohl der quantitativen als auch der qualitativen Bewertung bei jeder Anbietung kritisch abzuwägen und nach einer angemessenen Lösung im Einzelfall zu suchen, er kann die sorgfältige archivfachliche Vorbereitung und Umsetzung der Bewertung schon gar nicht ersetzen, er kann sie aber wirkungsvoll unterstützen.

Unter diesen Prämissen könnte es für die deutschsprachige Archivlandschaft durchaus lohnenswert sein, sich um die Entwicklung eines gemeinsamen standardisierten Verfahrens zu bemühen, mit dessen Hilfe einschlägige Rumpfdaten aus frei zugänglichen, dynamischen, nicht kommerziellen Datenverzeichnissen über Personen von öffentlichem Interesse in die archivische Bewertungs- und Erschließungspraxis routinemäßig eingebunden und genutzt werden können. Die potenzielle, täglich wachsende Datengrundlage ist bereits jetzt beachtlich. Neben der deutschsprachigen **Wikipedia** (ca. 350000 Einträge) und der bereits erwähnten **Personendatenbank der Landesbibliographie Baden-Württemberg** (<http://www.statistik.baden-wuerttemberg.de/LABI/PersDB.asp>; 40000 Einträge) wären diese die

- **Datenbank Niedersächsische Personen** (http://www.gwlb.de/nis/niedersaechsische_personen/, ca. 25.000 Einträge für Niedersachsen und Bremen), die
- **Rheinland-Pfälzische Personendatenbank** (<http://www.rlb.de/cgi-bin/wwwalleg/maskrnam.pl?db=rnam>; ca. 10000 Einträge) und nicht zuletzt das
- **Biographie-Portal** (<http://www.biographie-portal.eu/>), das als Kooperation zwischen der Bayerischen Staatsbibliothek, der Historischen Kommission bei der Bayerischen Akademie der Wissenschaften, der Österreichischen Akademie der Wissenschaften und der Stiftung Historisches Lexikon der Schweiz über ein gemeinsames Register zunächst die [Allgemeine Deutsche Biographie \(ADB\)](#) [Neue Deutsche Biographie \(NDB\)](#) [Österreichisches Biographisches Lexikon 1815-1950 \(ÖBL\)](#) [Historisches Lexikon der Schweiz \(HLS\)](#) erschließen und so „weit über 100.000 wissenschaftlich fundierte Biographien aus allen gesellschaftlichen Bereichen und fast allen Epochen der deutschen, österreichischen und der schweizerischen Geschichte“ erfassen will. Danach sollen weitere nationale und regionale biographische Nachschlagewerke in das Biographie-Portal einbezogen werden.

Die technische Realisierung dürfte das geringste Problem darstellen. Die Kooperationsbereitschaft der zahlreichen Partner zu finden und vor allem das Erarbeiten einer datenschutzgerechten Regelung dürften weitaus anspruchsvollere, aber trotz allem keine unlösbaren Aufgaben sein. Einen Versuch, auf diese Weise die Intelligenz im Netz für archivische Zwecke sinnvoll zu nutzen, wäre es allemal wert.