

# Beschreibung der XML-Dateien der Retrodigitalisierung von NDB und ADB

Version	Datum	Änderung	Verantwortlich
0.1	2018-06-06	Erstfassung	Matthias Reinert

Die unter [data.deutsche-biographie.de](http://data.deutsche-biographie.de) ausgelieferte XML-Dateistruktur orientiert sich an der TEI-P5 Empfehlung: es werden *vorwiegend* Elemente aus dem core Modul genutzt.

## Abschnitte

Die Hauptabschnitte im XML korrespondieren mit dem Inhalt

Inhaltstyp	Abschnitt	Inhalte
Allgemeines	<teiHeader>	Rechte, Änderungen, Teilnehmer des Projekts
Registereintrag	<particDesc> <profileDesc>	Personen, Relationen Orte, Relationen
wenn vorhanden:	<body>	
NDB-Artikel	<div subtype="ndb">	Artikelinhalte ergänzt um Angaben zum Vorgänger und Nachfolger im NDB Band
ADB-Artikel	<div subtype="adb">	Artikelinhalte ergänzt um Angaben zum Vorgänger und Nachfolger im ADB Band

## Der Abschnitt <teiHeader>

Statischer Inhalt, absichtlich in allen XML-Dateien redundant aufgeführt, da diese einzeln ausgeliefert werden

## Der Abschnitt <fileDesc>

<respStmt> listet die Teilnehmenden in verschiedenen Projektabschnitten auf, sofern sie identifizierbare Änderungen an den XML-Tags vorgenommen haben, ist ihr Kürzel verzeichnet, das – in einer internen XML-Fassung – unter @resp mitgeführt wird

<publicationStmt> listet unter <availability> die drei Lizenzmöglichkeiten auf

- CC-BY-NC-SA für die ADB Artikel
- CC-BY-NC-ND für die NDB Artikel
- CC-0 für die Metadaten / alle anderen Daten

## Der Abschnitt <profileDesc>

für <particDesc> und <settingsDesc>

die genannten Personen – bis auf biographierte Person/Familie – werden nicht redundant aufgeführt, da sie eigene Instanzen in XML haben, dagegen werden die Orte redundant aufgeführt, sie haben noch keine eigenen Instanzen

## Der Abschnitt <particDesc>

Unter <listPerson> wird bisher nur die biographierte Person aufgeführt, da die anderen Personen/Familien in ihren eigenen XML-Dateien beschrieben werden.

< <a href="#">person@role</a> =entry	Eigenschaften der biographierten Person/Familie
@sex	1 = Mann, 2 = Frau, 9 = Familie, 0 = unbekannt
@n	Schlüsselfeldzahl „sfz[0-9][0-9_-]+“
@xml:id	GND Nummer „ppn[0-9-X]+“
< <a href="#">persName@default</a> >	Namensansetzung
< <a href="#">persName@search</a> >	Suchvarianten des Namens
< <a href="#">persName@sort</a> >	Sortiervariante des Namens
<floruit>	Angabe zum Wirkungszeitraum
<occupation>	Semikolon separierte Liste; Angabe einer Berufsklassifikation im Attribut @
<faith>	Semikolon separierte Liste
<note type="adb ndb"> <rs> @type @ref <name	Registerangaben zu einer Person – Referenz auf die @type=adb ndb mit dem Ziel @ref=sfz...

<pre> subtype="author" type="adb ndb"&gt;   &lt;bibl type="adb ndb"&gt;   &lt;list@type     &lt;item@n       &lt;list&gt;         &lt;ref target="URL"&gt;Seitenanga be ggf. Ergänzungen&lt;/ref&gt; </pre>	<ul style="list-style-type: none"> <li>- Name des Autors @type=adb ndb</li> <li>- Zitierangabe mit Bandangaben je nach @type=adb ndb -</li> <li>- geschachtelte Liste um die Struktur NDB &gt; Band &gt; Seitereferenz abzubilden als <a href="#">list@type</a>=(a n)db &gt; <a href="#">item@n</a>=BANDNR &gt; list &gt; item &gt; ref+</li> </ul>
<pre> &lt;birth&gt; &lt;death&gt;   &lt;date&gt;   &lt;placeName&gt; </pre>	
<pre> &lt;listRelation type="spatial"&gt;   &lt;relation name="" active="Biographierte r" passive="OrtsID"&gt; </pre>	
<pre> &lt;listRelation type="genealogical"&gt;   &lt;relation name="" active="Biographierte r" passive="SFZ"&gt; </pre>	

### Der Abschnitt <stettingsDesc>

<pre> &lt;listPlace&gt;   &lt;place@xml:id&gt; </pre>	<p>Die Liste der genannten Orte</p> <ul style="list-style-type: none"> <li>- die ID ist intern und bezieht sich auf im Aufbau befindliche Ortsdatenbank</li> </ul>
<pre> &lt;location&gt;&lt;geo&gt; </pre>	<p>Koordinaten</p>
<pre> &lt;placeName&gt; </pre>	<p>Ortsname, ggf. mit normalisierter Fassung unter @n wenn bekannt</p>
<pre> &lt;idno type="gnd osm odb"&gt; </pre>	<p>ID-Nummer die auf Normdaten/Datenbanken referenziert</p> <p>GND</p> <p>OpenStreetmap (kein persitenter Identifier)</p>

**Der Abschnitt** `<div type="entry" subtype="adb|ndb">`

... hat weitere Attribute @n und @xml:id

<div type="prev"> <head>	Vorgänger im Band, dessen Name und @sfz
<div type="prev"> <head>	Nachfolger im Band, dessen Name und @sfz
<div type="citation">	Zitiernamen dieses Biographierten
<div type="kopf"> <p> <seg type="notname"> <birth@n> <date@type,@n>  <placeName@type> <death@n> <date@type,@n>  <placeName@type> <placeName type="tomb"> <faith> </seg>	<ul style="list-style-type: none"> <li>– Kopfzeile der Biographie, in einem &lt;p&gt; der Namen enthält und ggf. ein &lt;seg&gt; für den Rest (Beruf, Geburt, Tod, Konfession) = <i>alle Daten sind bereits in den Metadaten kodiert</i></li> <li>– Geburt und Tod mit @n als normalisierte Datumsangabe, tw. @calendar=Julian (nicht interpretiert)</li> <li>– Ortsnamen wenn identifiziert @key = interne OrtsID, @resp = Zuordnungskürzel, @corresp = lat,lon Angabe, @osmname = wenn bekannt</li> <li>–</li> </ul>
<div type="geneal">	
<div type="leben">	
<div type="A W L P Qu">	
	– Verweise auf Familienmitglieder – nur in sog. Familienartikeln möglich

## Elemente unterhalb von <p>

<seg>	bisher nur im <div@kopf>, sonstungenutzt
<choice><abbr type="acronym"><expan> <choice><abbr type="lemma"><expan>	Abkürzung, die als Akronym nicht durch <expan> ersetzt werden sollte  Das Lemma ist abgekürzt und i.d.R. vereinfacht die Kurzform das Lesen.
<choice><abbr><expan>	Abkürzung, die durch <expan> erklärt werden sollte, die Ersetzung ist nur in den Bänden NDB 1-12 realisiert, in den übrigen Bänden sind die <expan> automatisch eingesetzt und tw. nicht eindeutig [Psychologe, Psychologie]
<choice><orig><reg>	Trennungen von Worten an <pb> und <milestone> werden durch <reg> zusammengesetzt angegeben
<name type="(person place org job)">	
<ref target="">	Link auf <ul style="list-style-type: none"> <li>- Personenartikel</li> <li>- Literaturtitel</li> <li>- Projekte wie GW, dMGH</li> </ul>
<hi rend="(italics antiqua sup sub)">	kann überall unterhalb von <p> kann vorkommen:
<lb/> <milestone /> <pb />	Zeilenumbruch Spaltenwecheln Seitenwechsel