

Cornelia Diebel

# Sammlung von Online-Publikationen an der Deutschen Nationalbibliothek (DNB)



[CC-BY-SA 3.0 DE](https://creativecommons.org/licenses/by-sa/3.0/de/)

## Inhaltsverzeichnis

- 1. Auftrag der Deutschen Nationalbibliothek – Welche Ziele verfolgt die DNB mit der Sammlung von Online-Publikationen?**
- 2. Voraussetzungen für die Sammlung von Online-Publikationen**
- 3. Umsetzung der Ziele**

# Auftrag der Deutschen Nationalbibliothek

"Die Deutsche Nationalbibliothek hat die Aufgabe, lückenlos alle deutschen und deutschsprachigen Publikationen ab 1913, (...) dauerhaft zu archivieren, bibliografisch zu verzeichnen sowie der Öffentlichkeit zur Verfügung zu stellen."

Gesetzlich geregelt wird dies im "[Gesetz über die Deutsche Nationalbibliothek](#)" (DNBG) vom 22.06.2006.

## Was bedeutet der gesetzliche Auftrag?

Neben der schon lange bestehenden Ablieferungspflicht für körperliche Medienwerke, sind seit 2006 auch sogenannte *unkörperliche Medienwerke* ablieferungspflichtig.

"Medienwerke in unkörperlicher Form sind alle Darstellungen in öffentlichen Netzen." (§3 Nr. 3 DNBG)

Konsequenz: Druckwerke und zusätzlich Online-Publikationen sind ablieferungspflichtig.

## Fokus auf Einzelobjekte

Die DNB geht bei der Sammlung schrittweise vor.

Startpunkt und erster Fokus für die Sammlung nach der Gesetzesnovellierung waren Einzelobjekte, die ein Äquivalent in der körperlichen Welt haben (E-Books, E-Journals, digitale Noten, Hörbücher), inkl. Veröffentlichungen, die nur noch online erscheinen.

Webharvesting als weiterer Schritt. Hier werden zwei Wege der Sammlung beschrrieben: selektives Webharvesting und ein Domainharvesting in einem vergebenen Speichervolumen.

# Online-Publikationen: was sammeln wir?

Online-Publikationen, die wir aktiv sammeln:

- E-Books/monografische Veröffentlichungen, incl. Online-Hochschulschriften
- E-Journals
- E-Paper (PDF-Ausgaben von Tageszeitungen)
- Noten
- E-Hörbücher
- Webseiten

Online-Publikationen, die wir zurzeit noch nicht sammeln:

- Musiknetzpublikationen
- Datenbanken

## Welche Ziele verfolgt die DNB mit der Sammlung der Online-Publikationen?

- Möglichst vollständige Sammlung ohne manuellen Aufwand für Erschließung und Archivierung (Entwicklung eines automatisierten Workflows)
- Gute Präsentation im [Katalog](#), gute Recherchemöglichkeiten mit möglichst guten (bereits abgelieferten) Metadaten
- Persistente Adressierung und Zitierfähigkeit ([URN](#))
- [Langzeitarchivierung](#) und [Langzeitverfügbarkeit](#)

# Voraussetzungen für die Sammlung

Die Sammlung von Online-Publikationen basiert in der DNB ausschließlich auf automatischen Prozessen.

Hierfür müssen vier Voraussetzungen erfüllt werden:

1. Sammlung von Objekt und Metadaten
2. Nutzung von standardisierten Metadatenformaten
3. Bereitstellung von Standarddateiformaten
4. Nutzung von standardisierten [Schnittstellen](#)



# Sammlung von Objekt und Metadaten

DNB ist überzeugt, dass für alle Online-Publikationen auch Metadaten vorliegen (sollten) und nutzt diese für die eigenen Prozesse nach, indem sie in die Nachweissysteme der DNB transformiert werden. Dies geschieht in unterschiedlicher Granularität.

Für E-Books und monografisches Material in *einer* Beschreibung. Für Inhalte von E-Journals in einer *mehrstufigen* Beschreibung (Titel der E-Zeitschrift + Angaben zum E-Journalheft/bzw. Journalartikel).

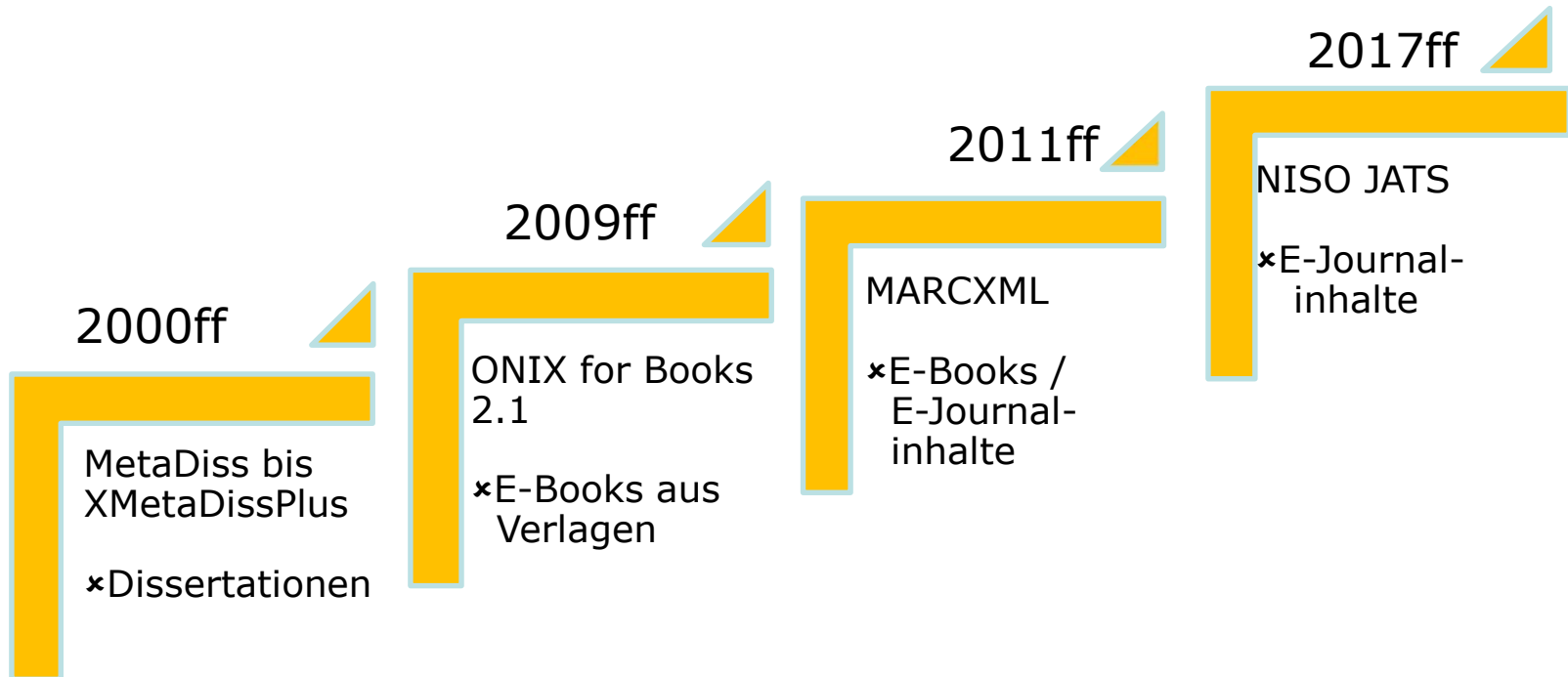
# Standardisierte Metadatenformate

DNB orientiert sich an Metadatenformaten, die in verschiedenen Communities verbreitet sind. Für diese Metadatenformate wurden sogenannte Kernsets erarbeitet, die jeweils für eine Ausprägung gelten.

Auch für die Ablieferung über das Webformular gilt das Kernset und ist so abgebildet, dass die Eingabe direkt erfolgt.

Hier wird zwischen obligatorischen und fakultativen Angaben unterschieden.

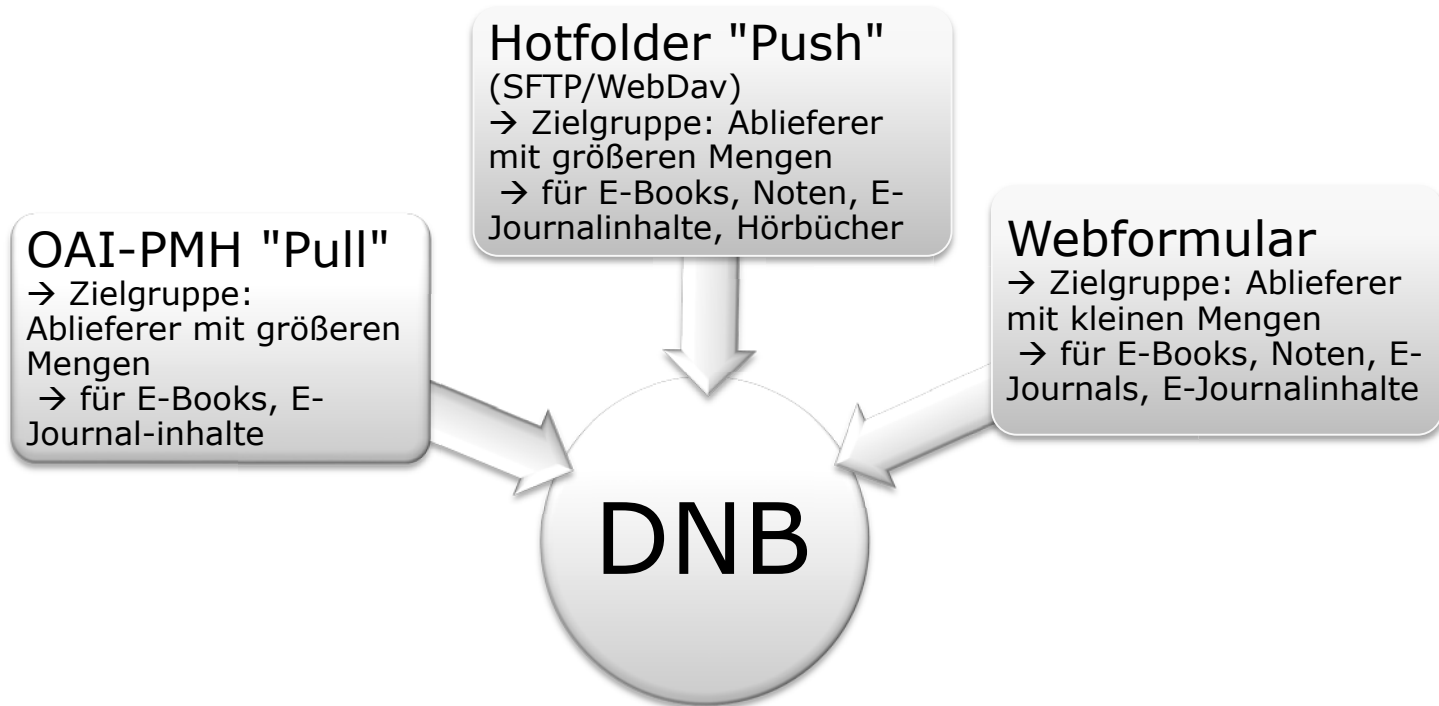
# Metadatenstandards



## Akzeptierte Dateiformate in DNB



# Standardisierte Schnittstellen



## Umsetzung der Ziele

- Möglichst vollständige Sammlung
- Gute Präsentation und gute Recherchemöglichkeiten im Katalog
- Persistente Adressierung und Zitierfähigkeit (URN)
- Langzeitarchivierung und Langzeitverfügbarkeit

# Zahlen zur aktuellen Sammlung - Stand 31.03.2019

Im Jahr 2018 wurden durchschnittlich 3.834 Netzpublikationen pro Tag importiert.

Gesamtzahl Ende März 2019: **6.703.422**

|  |           |                                   |
|--|-----------|-----------------------------------|
| E-Books:                                     | 2.110.541 |                                   |
| E-Paper-Ausgaben:                            | 2.206.000 | (1.200 pro Tag)                   |
| E-Journalinhalte:                            | 2.453.530 |                                   |
|  |           | von 7.847 laufenden Zeitschriften |
| Hörbücher, Noten, Zeitschnitte von Websites: | 66.652    |                                   |

# Darstellung im Katalog der DNB

## [Katalog der Deutschen Nationalbibliothek](#)

Außerdem Verzeichnung der Daten in der Deutschen Nationalbibliografie.

Suche nach [Online-Publikationen mit "Parlament" im Titel](#)

- Standort Online (frei zugänglich)
- Zugriff auf das Archivobjekt



## Darstellung im Katalog der DNB

Beispiele für Zeitschriftentitel:

- <http://d-nb.info/1001247639>
- <http://d-nb.info/1078169217>

Zugriff auf die dazugehörigen Artikel oder Hefte über  
„zugehörige Ausgaben/Hefte“

<http://d-nb.info/1113593202>

Zeitschriftenheft

<http://d-nb.info/1101845449>

Zeitschriftenartikel

# Nutzung von Online-Publikationen

Die Bereitstellung der Online-Publikationen erfolgt nach dem was die Abliefernden uns erlauben:

- Weltweit ohne Einschränkung
- Nutzung nur in den Lesesälen der DNB in Frankfurt und Leipzig

Urheberrechtliche Einschränkungen werden von DNB selbstverständlich beachtet.

# Persistente Adressierung und Zitierfähigkeit (URN)

Informationen zum Thema URN in der Präsentation von Frau Ackermann „Persistent identifizieren mit dem URN-Service der Deutschen Nationalbibliothek“

und unter

<http://www.dnb.de/urnservice>

# Langzeitarchivierung von Online-Publikationen

Langzeitarchivierung (LZA) bedeutet die Sicherstellung des Erhalts der digitalen Daten und die Gewährleistung der Langzeitverfügbarkeit. Gemeint ist hierbei der Erhalt der Interpretierbarkeit der Inhalte der digitalen Daten.

Publikationen in neuen Dateiformaten müssen auf aktuellen Computersystemen genauso genutzt werden können, wie solche in längst obsoleten.

## Analyse technischer Metadaten - „Vollständigkeit“ und Validität

- Metadaten-Kernset (pro Dateiformat):  
 Enthält eine Definition von technischen Metadaten, die aus DNB Sicht zwingend erforderlich sind um mit hoher Wahrscheinlichkeit langzeitarchiviert werden zu können.
- Prüfung ob ein erkanntes Dateiformat gemäß seiner Spezifikation korrekt erstellt worden ist
- Digital Rights Management vorhanden? Passwort-Schutz, Kopierschutz,... = für die Langzeitarchivierung untauglich

# Vielen Dank für Ihre Aufmerksamkeit

Wir laden Sie herzlich ein, mit uns über die Ablieferung von Online-Publikationen ins Gespräch zu kommen.

Weitere Informationen unter:

<http://www.dnb.de/netzpublikationen>

Claudia Germer [c.germer@dnb.de](mailto:c.germer@dnb.de)

Cornelia Diebel [c.diebel@dnb.de](mailto:c.diebel@dnb.de)