



DA-NRW Software Suite (DNS): „SIP-Builder“

Version 0.6.5

Redaktionelle Änderung: 10.08.15

Anleitung

1. Überblick

Der SIP-Builder ist ein Tool, mit dem sich zur Einlieferung in die DNS vorgesehene Daten in eine geeignete Paketstruktur überführen lassen. Das Programm kann in zwei Modi ausgeführt werden:

- Im **GUI-Modus** werden Sie in einer grafischen Benutzeroberfläche (GUI) durch mehrere

aufeinander folgende Schritte geführt, in deren Verlauf Sie die nötigen Einstellungen vornehmen können. Sie können dabei jederzeit über eine Reihe von Schaltflächen am linken Rand zwischen den einzelnen Schritten hin- und herwechseln. Am Ende des Verfahrens erhalten Sie eine bzw. mehrere Archivdateien (wahlweise im tgz- oder tar-Format), die zur Einlieferung in das DNS geeignet sind. Dieser Modus wird für die meisten Anwender empfohlen.

- Fortgeschrittene Benutzer können den SIP-Builder alternativ im **CLI-Modus** auf der Kommandozeile (CLI) aufrufen und die gewünschten Einstellungen als Parameter angeben. Diese Vorgehensweise bietet sich beispielsweise an, um den SIP-Erstellungsvorgang in automatisierte Prozesse einzubinden.

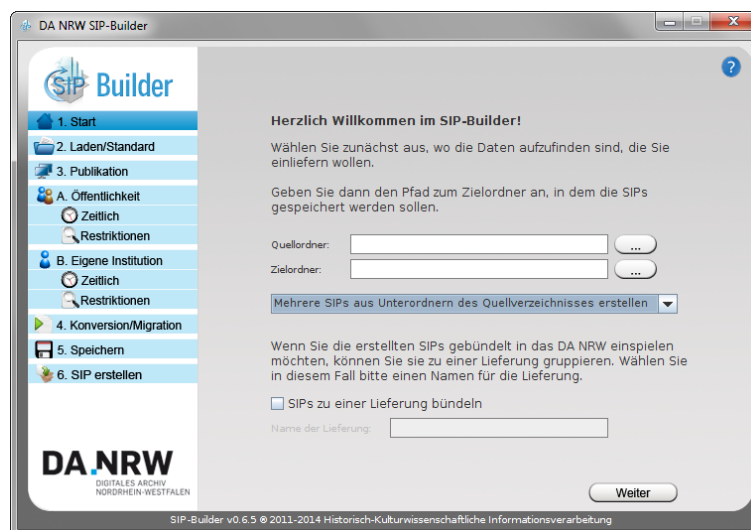
2. GUI-Modus

2.1 Ausführen des SIP-Builders im GUI-Modus

Starten Sie den SIP-Builder, indem Sie je nach Betriebssystem eine der folgenden Dateien ausführen:

- SipBuilder-Windows.exe (Windows)
- SipBuilder-Unix.sh (Unix/Mac OS)

2.2 Start



Quellordner

Geben Sie hier den Pfad zum Verzeichnis an, das die Quelldaten enthält, aus denen das SIP generiert werden soll.

Zielordner

Im hier angegebenen Verzeichnis wird das fertige SIP am Ende des Generierungsprozesses in Form einer tgz- oder tar-Archivdatei hinterlegt.

Generierungsmodus

Sie haben die Wahl zwischen zwei verschiedenen Modi der SIP-Generierung:

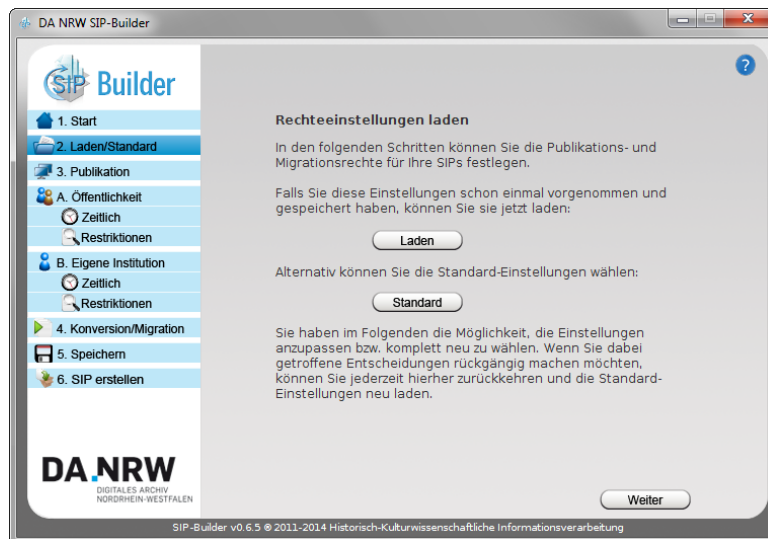
- *Einzelnes SIP aus dem Quellverzeichnis erstellen:*

Wählen Sie diese Option, wenn Sie die Daten des Quellordners zu einem einzelnen SIP bündeln möchten. Auch eventuell existierende Unterordner werden dem Paket hinzugefügt.

- *Mehrere SIPs aus Unterordnern des Quellverzeichnisses erstellen:*

Wählen Sie diese Option, um mehrere SIPs mit den gleichen Rechteeinstellungen auf einmal zu generieren. Dazu müssen sich die Daten für jedes Paket in einem eigenen Unterordner des Quellverzeichnisses befinden. Im Generierungsprozess werden die Daten jedes einzelnen Verzeichnisses zu jeweils einem eigenen SIP gebündelt.

2.3 Rechteeinstellungen laden



Laden

Wenn Sie die Rechteinstellungen bei einem früheren SIP-Generierungsvorgang schon einmal festgelegt und gespeichert haben, können Sie die dabei erstellte Datei mit dieser Option laden. Die geladenen Einstellungen können in den nachfolgenden Schritten überprüft und gegebenenfalls angepasst werden.

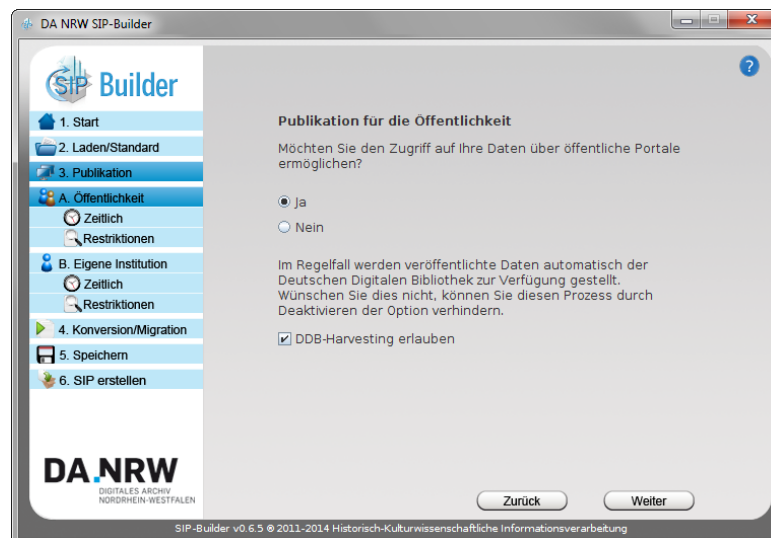
Standard

Wählen Sie diese Option, um die Standard-Rechteinstellungen zu wählen:

- Generierung von Publikationsdaten für die Öffentlichkeit, keine Restriktionen
- Keine Generierung von zusätzlichen Publikationsdaten für die eigene Institution
- Keine Migrationsbedingungen

Bei Betätigung des Buttons werden die zuvor gewählten Einstellungen verworfen und durch die Standardwerte ersetzt.

2.4 Publikation



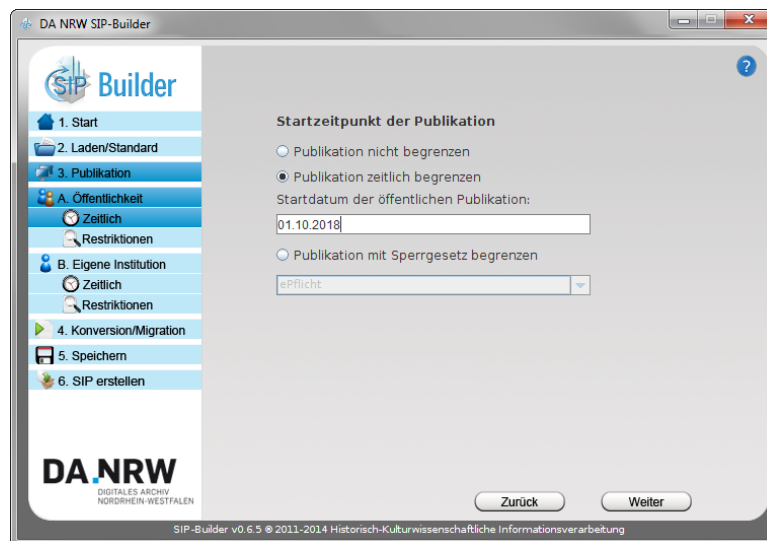
Die Publikationseinstellungen können in zwei Kategorien festgelegt werden:

- Erstellung von Publikationsderivaten zur Anzeige in öffentlichen Portalen
- Erstellung von Publikationsderivaten, die nur der eigenen Institution über die Schnittstellen des Presentation Repository zugänglich sind

Sie können jeweils wählen, ob die Erstellung der Publikationsdaten grundsätzlich stattfinden soll und ob in diesem Fall zusätzliche Restriktionen in Kraft treten sollen. Diese Restriktionen können für jede der beide Kategorien separat festgelegt werden.

Daten, die für die Anzeige in öffentlichen Portalen erstellt wurden, werden automatisch auch der Deutschen Digitalen Bibliothek zur Verfügung gestellt. Möchten Sie diesen Vorgang vermeiden, können Sie die Checkbox „DDB-Harvesting erlauben“ deaktivieren. In diesem Fall werden der Deutschen Digitalen Bibliothek keine Daten zugänglich gemacht.

2.4.1 Startzeitpunkt der Publikation



Publikation zeitlich begrenzen

Wählen Sie diese Option, um ein Datum zu bestimmen, ab dem die Publikation stattfinden soll. Vor diesem Datum findet keine Publikation statt. Das Datum muss dazu in der Form [Tag].[Monat].[Jahr] angegeben werden.

Beispiel: 14.07.2020

Publikation mit Sperrgesetz begrenzen

Alternativ zum Startdatum können Sie den Beginn der Publikation auch durch die Wahl eines regulierenden Gesetzes festlegen. Wählen Sie den Namen des gewünschten Gesetzes dazu aus der Drop-Down-Box aus.

Zur Wahl stehen derzeit:

- ePflcht
- UrhG DE

2.4.2 Vorschaurestriktionen



In diesem Schritt lassen sich Restriktionen wählen, die in den Rechteeinstellungen hinterlegt und während der Erstellung der Publikationsdaten angewendet werden.

Text: Einsehbare Seiten festlegen

Wenn Sie möchten, dass nur bestimmte Seiten eines PDF-Dokuments für den Benutzer einsehbar sind, können Sie die entsprechenden Seitenzahlen bei Aktivierung dieser Option in einem Textfeld angeben.

Reihen Sie die gewünschten Seitenzahlen dazu durch Kommas voneinander getrennt aneinander. Um mehrere aufeinander folgende Seiten festzulegen, genügt es, die erste und letzte Seitenzahl mit einem dazwischen liegenden Bindestrich anzugeben.

Beispiele:

- Eingabe: 5, 10, 27

Angezeigt werden die Seiten 5, 10 und 27.

- Eingabe: 1-5

Angezeigt werden die Seiten 1, 2, 3, 4 und 5.

- Eingabe: 1, 14-17, 24, 30

Angezeigt werden die Seiten 1, 14, 15, 16, 17, 24 und 30.

Bild: Bildqualität begrenzen

Bei Aktivierung dieser Option haben Sie die Möglichkeit, die Anzeigequalität von Bilddateien zu begrenzen, indem Sie die gewünschte Auflösung der Bilder angeben. Sie können dabei zwischen zwei Möglichkeiten wählen:

- Bestimmung der maximalen Auflösung durch absolute Werte in drei Abstufungen (niedrig, mittel, hoch)
- Bestimmung der Auflösung relativ zur Auflösung im Original (angegeben in Prozentwerten)

Bild: Fußzeile oder Wasserzeichen angeben

Verwenden Sie diese Option, um Ihren Bildern einen beliebigen Text hinzu-zufügen, der beispielsweise Copyright-Informationen beinhalten kann. Geben Sie den Text dazu in das Textfeld ein und legen Sie die gewünschten Einstellungen fest:

- *Texttyp:*
 - ☐ *Fußzeile:* Der Text wird im unteren Bereich des Bildes als weiße Schrift auf einem schwarzen Balken eingefügt. Ein kleiner Teil des Bildes wird dabei von der Fußzeile verdeckt.
 - ☐ *Wasserzeichen (oben):* Der Text wird im oberen Bereich des Bildes eingefügt.
 - ☐ *Wasserzeichen (mittig):* Der Text wird im mittleren Bereich des Bildes eingefügt.
 - ☐ *Wasserzeichen (unten):* Der Text wird im unteren Bereich des Bildes eingefügt.
- *Sichtbarkeit (nur bei Wasserzeichen):* Wasserzeichen können auf Wunsch transparent angezeigt werden. Zur Auswahl stehen Werte zwischen 5% (kaum sichtbar) und 100% (vollständig sichtbar).
- *Schriftgröße (nur bei Wasserzeichen):* Wählen Sie die gewünschte Schriftgröße, in der der Text angezeigt wird. Bitte beachten Sie, dass um-fangreiche Texte in höheren Schriftgrößen auf kleineren Bildern möglicherweise nicht mehr vollständig angezeigt werden können.

Bitte beachten Sie, dass die Länge der Texte beschränkt ist: Fußzeilen dürfen höchstens 65 Zeichen lang sein, bei Wasserzeichen beträgt die maximale Länge 20 Zeichen.

Audio: Länge von Audio-Dateien begrenzen

Um dem Benutzer lediglich den Beginn einer Audio-Datei zugänglich zu ma-chen, können Sie die maximale Abspieldauer von Audio-Dateien begrenzen; es werden dann etwa nur die ersten fünfzehn Sekunden abgespielt. Die genaue Sekundenanzahl können Sie hier festlegen.

Video: Bildqualität von Videos begrenzen

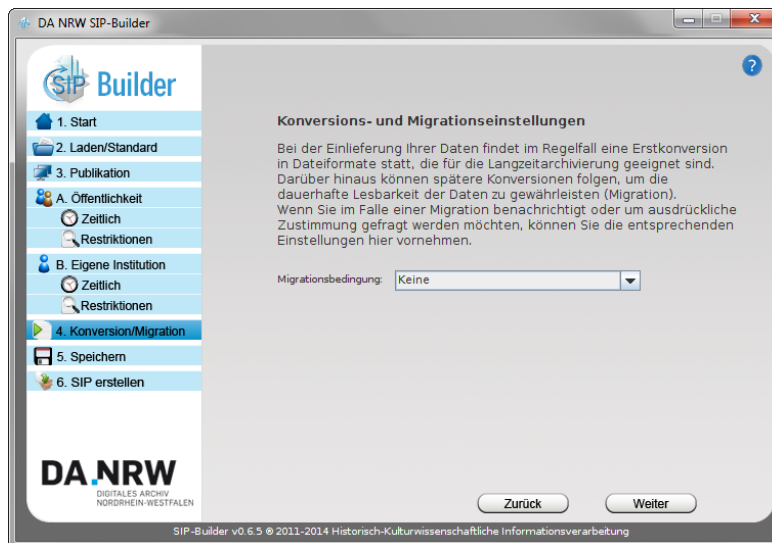
Bei Aktivierung dieser Option haben Sie die Möglichkeit, die Anzeigequalität von Videodateien zu begrenzen. Die Höhe der Auflösung kann dabei aus drei möglichen Einstellungen (niedrig, mittel, hoch) gewählt werden.

Video: Länge von Videos begrenzen

Neben Audio-Dateien lassen sich auch Videos in ihrer Länge begrenzen. Legen Sie hier fest,

wieviele Sekunden eines Videos abgespielt werden können.

2.5 Konversions- und Migrationseinstellungen

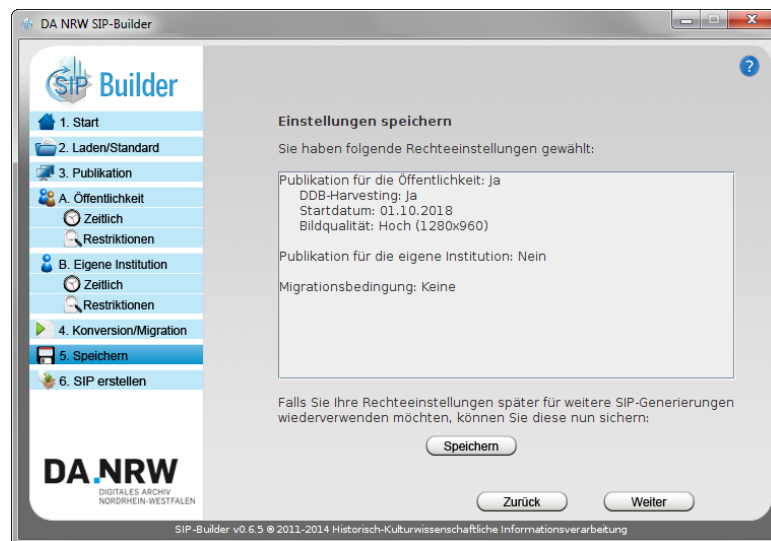


Bei der Einlieferung Ihrer Daten findet im Regelfall eine Erstkonversion in Dateiformate statt, die für die Langzeitarchivierung geeignet sind. Darüber hinaus können spätere Konversionen folgen, um die dauerhafte Lesbarkeit der Daten zu gewährleisten (Migration). Sie können die Durchführung der Migration dabei auf Wunsch an eine Bedingung knüpfen.

Mögliche Einstellungen sind:

- *Keine Migrationsbedingung:* Migrationen werden automatisch ohne Ihre Beteiligung durchgeführt (Standardeinstellung).
- *Über Migration informieren:* Im Falle einer Migration erhalten Sie eine Benachrichtung per Mail.
- *Zustimmung für Migration einholen:* Sie werden vor einer anstehenden Migration benachrichtigt und haben die Möglichkeit, der Migration zu widersprechen. Dies gilt für die Erstmigration und alle folgenden Migrationen. Nach einer Zeitfrist verfällt die Rückfrage, in diesem Fall wird Ihr Objekt nicht migriert.

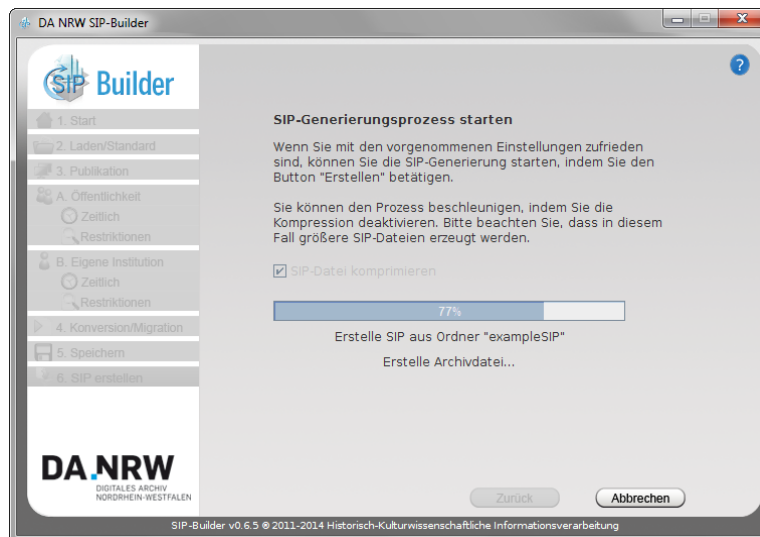
2.6 Einstellungen speichern



In diesem Schritt werden die zuvor getroffenen Rechteeinstellungen übersichtlich zusammengefasst, sodass Sie sie vor der SIP-Erstellung noch einmal überprüfen können.

Wenn Sie die Einstellungen bei einer späteren Verwendung des SIP-Builders noch einmal anwenden möchten, können Sie diese darüber hinaus sichern, indem Sie den Button „Speichern“ betätigen. Die dabei erstellte Datei können Sie später jederzeit im Schritt „Rechteeinstellungen laden“ wieder einlesen, sodass Sie die Einstellungen nicht erneut vornehmen müssen.

2.7 SIP-Generierungsprozess starten



Starten Sie den Prozess, indem Sie den Button „Erstellen“ betätigen. Je nach Menge der zu verarbeitenden Daten und der Leistungsfähigkeit Ihres Systems kann der Vorgang zwischen wenigen Sekunden und mehreren Stunden in Anspruch nehmen. Um den Prozess zu beschleunigen, können Sie die Option “SIP-Datei komprimieren” deaktivieren. In diesem Fall wird statt einer komprimierten tgz-Datei eine entsprechend größere, unkomprimierte tar-Datei erstellt.

3. CLI-Modus

3.1 Ausführen des SIP-Builders im CLI-Modus

Um den SIP-Builder im CLI-Modus zu starten, sind folgende Schritte notwendig:

- Rufen Sie die Kommandozeile auf und wechseln Sie in das Verzeichnis, in dem sich der SIP-Builder befindet.
- Starten Sie den SIP-Builder mit dem Kommando „java -jar SipBuilder.jar“; unter Unix können Sie alternativ auch das Kommando „./SipBuilder.sh“ verwenden.
- Geben Sie beim Aufruf zusätzlich eine Reihe von Parametern an, um den CLI-Modus zu aktivieren.
- Obligatorisch ist die Angabe von jeweils einem Parameter der Kategorien „Quelle“ und „Ziel“ (siehe unten). Darüber hinaus können Sie optional weitere Parameter angeben, um die gewünschten Einstellungen der SIP-Erstellung festzulegen.

3.2 Übersicht der Parameter

3.2.1 Quelle

-source="[Pfad]"

Geben Sie den Pfad zum Verzeichnis an, in dem sich die Quelldaten befinden. Aus diesem Quellordner werden anschließend die SIPs erstellt.

Beispiele:

```
-source="C:\Eigene Dateien\SIP-Source"
```

```
-source="/home/user/sipData"
```

-filelist="[Pfad]"

Diese Option kann verwendet werden, um ein einzelnes SIP aus Dateien zu erstellen, die sich in unterschiedlichen Verzeichnissen befinden. Erstellen Sie dazu eine Textdatei und listen Sie darin die Pfade zu den Dateien und/oder Verzeichnissen auf, aus denen das SIP erstellt werden soll. In jeder Zeile der Textdatei muss dabei genau eine Pfadangabe stehen. Geben Sie den Pfad eines Verzeichnisses an, wird der gesamte Inhalt des Verzeichnisses inklusive aller Unterordner in das SIP aufgenommen.

Bei Verwendung dieser Option muss gleichzeitig mit dem Parameter *-single= "[Name]"* ein Name für das SIP angegeben werden.

Beispiel:

```
-filelist="C:\Eigene Dateien\SIP-Source\filelist_sip1.txt"
```

Beispielinhalt einer Filelist:

```
D:\Foto\2010\foto_1.tif
```

```
D:\Foto\2010\foto_2.tif
```

```
D:\Foto\2013\
```

```
D:\Audio\audio_15_03_2012.wav
```

```
C:\Dokumente\dokument.pdf
```

-siplist="[Pfad]"

Diese Option ermöglicht es, die Inhalte eines oder mehrerer SIPs in einer XML-Datei festzulegen.

Erstellen Sie dazu eine XML-Datei mit dem Wurzelknoten „sipList“, dem Sie für jedes zu erzeugende SIP ein Element „sip“ hinzufügen. An jedes SIP müssen Sie über das Attribut „name“ einen Namen vergeben. Die einzelnen Dateien und/oder Verzeichnisse, aus denen das SIP erstellt werden soll, geben Sie dabei durch jeweils ein „file“-Element an.

Beispiel:

```
-siplist="C:\Eigene Dateien\SIP-Source\siplist.xml"
```

Beispielinhalt einer SIP-Liste:

```
<sipList>

  <sip name="SIP_Nr1">

    <file>D:\Foto\2010\foto_1.tif</file>

    <file>D:\Audio\audio_12_07_2011.wav</file>

  </sip>

  <sip name="SIP_Nr2">

    <file>D:\Foto\2003</file>

    <file>D:\Foto\2004</file>

    <file>D:\Foto\2005</file>

    <file>D:\Foto\2006</file>

    <file>D:\Foto\2007</file>

  </sip>

  <sip name="SIP_Nr3">
    <file>C:\Dokumente\Video_Dateiübersicht.pdf</file>
    <file>D:\Video\High_Resolution</file>
  </sip>

</sipList>
```

3.2.2 Ziel

-destination="[Pfad]"

Geben Sie den Pfad zum Verzeichnis an, in dem die SIPs erstellt werden sollen. Bitte beachten Sie bei gleichzeitiger Verwendung der Option *-source*, dass das gewählte Verzeichnis kein Unterverzeichnis des Quellordners sein darf.

Beispiel:

`-destination="D:\SIPs"`

3.2.3 Rechte

-default (Standardoption)

Die Standardrechte werden angewendet:

- Generierung von Publikationsdaten für die Öffentlichkeit, keine Restriktionen
- Keine Generierung von Publikationsdaten für die eigene Institution
- Keine Migrationsbedingungen

-premis="[Pfad]"

Geben Sie den Pfad einer Premis-Datei an, die eine Rights Section mit den von Ihnen gewünschten Rechten enthält. Die dort angegebenen Rechteeinstellungen werden in die Premis-Datei des neu erstellten SIPs übernommen.

Beispiel:

```
-premis="C:\Dokumente\SIPs\sip001\data\premis.xml"
```

-rights="[Pfad]"

Geben Sie den Pfad einer Rechte-Datei an, die zuvor im GUI-Modus des SIP-Builders erstellt wurde. Diese Option entspricht dem Laden von Rechteeinstellungen im GUI-Modus.

Beispiel:

```
-rights="C:\Dokumente\SIP-Source\contractRights.xml"
```

3.2.4 Generierungsmodus

-multiple (Standardoption)

Bei Anwendung dieser Option werden mehrere SIPs auf einmal generiert. Bei gleichzeitiger Angabe eines Quellordners mit dem Parameter *-source* werden die SIPs aus den Unterordnern dieses Quellordners erstellt, wobei aus jedem Unterordner jeweils ein SIP entsteht.

Die Option **muss** bei gleichzeitiger Verwendung von *-siplist* gewählt werden und kann **nicht** gleichzeitig mit *-filelist* benutzt werden.

-single

Bei Anwendung dieser Option wird ein einzelnes SIP generiert. Bei gleichzeitiger Angabe eines Quellordners mit *-source* wird das SIP aus dem gesamten Inhalt dieses Quellordners (inklusive möglicher Unterverzeichnisse) erstellt.

Die Option kann **nicht** gleichzeitig mit *-siplist* benutzt werden.

-single="[Name]"

Bei Anwendung dieser Option wird ebenfalls ein einzelnes SIP generiert. Zusätzlich wird der angegebene Text als Name des erstellten SIPs verwendet.

Die Option **muss** bei gleichzeitiger Verwendung von *-filelist* gewählt werden und kann **nicht** gleichzeitig mit *-siplist* benutzt werden.

Beispiel:

```
-single="Foto_SIP_2013"
```

3.2.6 Kompression

-compression (Standardoption)

SIPs werden als komprimierte tgz-Container gespeichert. Der SIP-Erstellungs-prozess verlängert sich dadurch; die erstellten Dateien nehmen im Gegenzug weniger Speicherplatz in Anspruch.

-noCompression

SIPs werden als unkomprimierte tar-Container gespeichert. Der SIP-Erstellungs-prozess verkürzt sich dadurch; die erstellten Dateien nehmen allerdings zusätz-lichen Speicherplatz in Anspruch.

3.2.7 Überschreiben von SIPs

-neverOverwrite (Standardoption)

SIPs werden nicht erstellt, wenn sich im Zielordner bereits ein SIP gleichen Namens befindet. Sie erhalten in diesem Fall im Anschluss an den SIP-Erstellungsprozess eine Mitteilung darüber, dass bereits existierende SIPs nicht neu erstellt wurden. Existiert bereits eine Lieferung gleichen Namens im Zielordner, wird der SIP-Erstellungsprozess nicht durchgeführt.

-alwaysOverwrite

Wenn sich bereits existierende SIPs oder Lieferungen gleichen Namens im Zielordner befinden, werden sie ohne weitere Nachfrage überschrieben, d. h. durch die neu erstellten Pakete bzw. die neu erstellte Lieferung ersetzt.

3.2.8 Sonstige Optionen

-ignoreExtensions="ext1;ext2;ext3..."

Wählen Sie diese Option, um Dateien mit bestimmten Dateiendungen nicht in die erzeugten SIPs aufnehmen zu lassen. Die entsprechenden Dateien werden bei der SIP-Erstellung ignoriert. Geben Sie die einzelnen Dateiendungen hintereinander durch Semikolons getrennt an.

Beispiel:

```
-ignoreExtensions="txt;doc;rtf;odt;pdf"
```

-help

Zeigt eine Übersicht der möglichen Optionen an.

3.3 Beispielaufrufe

```
java -jar SipBuilder.jar -source="D:\sipsource" -destination="D:\sips\" -single
```

Erstellt ein einziges SIP aus dem Ordner "D:\sipsource" im Verzeichnis "D:\sips\". Es werden die Standardrechte angewendet.

```
java -jar SipBuilder.jar -source="/home/user/sipData" -  
destination="/home/user/sips" -collection="ExampleCollection"
```

Erstellt mehrere SIPs aus den Unterordnern des Ordners "/home/user/sipData". Die SIPs werden anschließend zu einer Lieferung im Verzeichnis "/home/user/sips/ExampleCollection" gebündelt. Es werden die Standardrechte angewendet.

```
java -jar SipBuilder.jar -filelist="C:\Eigene Dateien\SIP-Source\filelist.txt"  
-destination="D:\sips\" -single="ExampleSIP" -  
rights="D:\sipRights\contractRights_001.xml" -alwaysOverwrite
```

Erstellt ein SIP gemäß den Angaben in der Datei "C:\Eigene Dateien\SIP-Source\filelist.txt" mit dem Namen "ExampleSIP" im Verzeichnis "D:\sips\". Dabei werden die in der Datei "D:\sipRights\contractRights_001.xml" hinterlegten Rechteeinstellungen verwendet.

Ein eventuell schon vorhandenes gleichnamiges SIP im Verzeichnis "D:\sips\" wird überschrieben.

```
java -jar SipBuilder.jar -siplist="C:\Eigene Dateien\SIP-Source\siplist.xml" -  
destination="D:\sips\" -premis="C:\Eigene Dateien\SIP-Source\premis.xml"
```

Erstellt mehrere SIPs gemäß den Angaben in der Datei "C:\Eigene Dateien\SIP-Source\siplist.xml". Die in der Premis-Datei "C:\Eigene Dateien\SIP-Source\premis.xml" angegebenen Rechte werden dabei übernommen.

3.4 Fehlercodes

Im CLI-Modus gibt das Programm beim Beenden einen Exit Code zurück, der im Falle einer erfolgreichen SIP-Generierung immer den Wert 0 beträgt. Wurde das Programm aufgrund eines Fehlers abgebrochen, gibt der Code Aufschluss über die Art des Fehlers. Durch Abfrage und Auswertung des Fehlercodes kann bei-spielsweise die Einbindung des SIP-Builders in automatisierte Prozesse erleichtert werden.

Die möglichen Codes und ihre Bedeutung können Sie der folgenden Übersicht entnehmen.

- 0 Erfolgreiche SIP-Erstellung
- 1 Ungültiger Parameter
- 2 Ungültige Kombination von Parametern
- 3 Angabe des Quellordners fehlt
- 4 Quellordner konnte nicht gefunden werden
- 5 Angabe des Zielordners fehlt
- 6 Quell- und Zielordner sind identisch
- 7 Zielordner ist ein Unterordner des Quellordners
- 8 Zielordner enthält Dateien, die keine Verzeichnisse sind (bei gleichzeiti- ger Erstellung
mehrerer SIPs)
- 9 Name des SIPs fehlt
- 10 Name des SIPs enthält ungültige Zeichen
- 11 Name der Lieferung fehlt

- 12 Name der Lieferung enthält ungültige Zeichen
- 13 Gleichnamige Lieferung existiert bereits
- 14 Dateiliste konnte nicht gefunden werden
- 15 Fehler beim Einlesen der Dateiliste
- 16 SIP-Liste konnte nicht gefunden werden
- 17 Fehler beim Einlesen der SIP-Liste
- 18 Premis-Datei konnte nicht gefunden werden
- 19 Rechte-Datei konnte nicht gefunden werden
- 20 Fehler beim Einlesen der Rechte-Datei
- 21 Fehler beim Einlesen der Standardrechte-Datei
- 22 Fehler beim Kopieren der Daten
- 23 Fehler beim Erstellen der Premis-Datei
- 24 Fehler beim Erstellen der BagIt-Metadaten
- 25 Fehler beim Erstellen der Archivdatei
- 26 Fehler beim Erstellen der Lieferung
- 27 Temporäre Daten konnten nicht entfernt werden