

Die Anbindung des DNB-Repositories an das DIAS-System

Erläuterungen zur Handhabung des LTP-Launcher und LTP-Submitters

Version 1.2

Stand: 07.08.2014

Redaktion: Stefan Hein, Karlheinz Schmitt, Andreas Schulz

Dokumenthistorie

Datum	Version	Bearbeitet von	Änderung
-	0.1	Karlheinz Schmitt	Dokumentenerstellung
19.06.2013	1.1	Stefan Hein	Update durch Inbetriebnahme von ltpLauncher v1.0.6
07.08.2014	1.2	Stefan Hein	bereinigt von Links und URLs

Inhalt

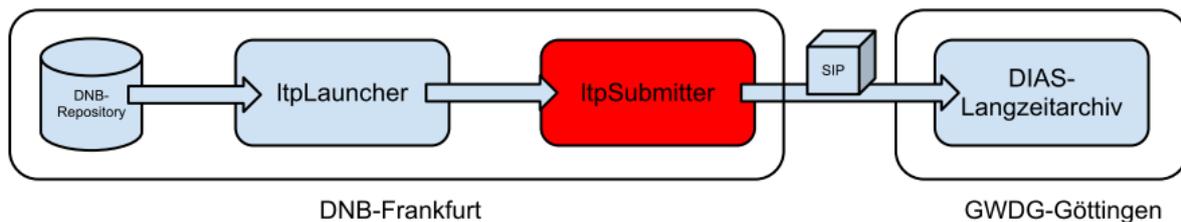
1	Einleitung	3
2	Aufgabe des LTP-Launcher's	3
3	Funktionen des LTP-Launchers	3
3.1	Anlegen und Löschen von Optionen für Harvesting-Aufträge	4
3.2	Anlegen von Harvesting-Aufträgen.....	8
3.3	Anhalten und Löschen von Harvesting-Vorgängen.....	10
3.4	Kontrolle von Harvesting-Vorgängen	12
3.5	Suche.....	14
3.6	Statistik und Konfiguration.....	17
4	Fehler-Handling	20

1 Einleitung

In diesem Dokument wird eine kurze Einführung in die Funktionalitäten des LTP-Launchers gegeben. Es wird beschrieben, wie (Harvesting-)Aufträge angelegt, kontrolliert und ggf. erneuert werden. Ebenso werden notwendige Konfigurationsdaten, die zur reibungslosen Kommunikation zwischen der DNB und dem DIAS-System notwendig sind, gelistet und erklärt.

Umgebung:

Der Workflow zur Übertragung an das Langzeitarchiv DIAS besteht aus 2 Applikationen, den LTP-Launcher (im Schaubild ItpLauncher) und den LTP-Submitter (im Schaubild ItpSubmitter).



Der ItpSubmitter empfängt Aufträge des ItpLauncher zur Langzeitarchivierung von im DNB-Repository abgelegten Netzpublikationen. Prinzipiell umfasst dieser Auftrag das Erstellen eines UOF-SIPs (inkl. einer Metadaten-Datei (mets.xml)), das Verpacken als TAR und Übertragen per FTP an DIAS.

2 Aufgabe des LTP-Launcher's

Die generelle Aufgabe des LTP-Launchers besteht darin, Objekte aus dem Repository der DNB an das DIAS-System zu übertragen und dort zu archivieren. Hierbei werden jedoch zur Zeit nur diejenigen Objekte im DNB-Repository berücksichtigt, die über den Import Service ab der Version 2.1 verarbeitet und ins Repository gelegt worden sind. Ältere Objekte, die nicht über diesen Workflow angenommen worden sind, verfügen zur Zeit noch nicht über die notwendige Struktur und müssen anderweitig nachbearbeitet werden.

3 Funktionen des LTP-Launchers

Im Folgenden werden die verschiedenen Funktionen des LTP-Launcher's beschrieben und erläutert wie diese innerhalb der grafischen Oberfläche gesteuert werden.

Es gibt insgesamt 6 Haupt-Funktionen

- 1.) Anlegen von Optionen für Harvesting-Aufträgen
- 2.) Anlegen von Harvesting-Aufträgen
- 3.) Anhalten und Löschen von Harvesting-Aufträgen
- 4.) Kontrolle von Harvesting-Vorgängen
- 5.) Suche

6.) Statistik und Konfiguration

3.1 Anlegen und Löschen von Optionen für Harvesting-Aufträge



Bevor ein Harvesting-Auftrag angelegt und gestartet werden kann muss eine Konfiguration für einen solchen Auftrag festgelegt werden. Ein Konfiguration wird angelegt, indem der Reiter: „Neue Harvesteroption einfügen“ (Offener Karteikasten) gewählt wird. Abbildung 1 zeigt die entsprechende Benutzeroberfläche im LTP-Launcher.



LTP - Launcher

DEUTSCHE
NATIONAL
BIBLIOTHEK



Neue Harvestingoption einfügen

Sie können hier eine neue Belegung für eine Harvesting-Option einfügen. Aktuell konfiguriertes Maximum für Anzahl ist 1000.

Eingabe

Option:

Wert:

Bei Fragen wenden Sie sich bitte an Stefan Hein (s.hein@dnb.de) oder Karlheinz Schmitt (k.schmitt@dnb.de)
Version 1.0.1 (07.03.2013)

Abbildung 1: Anlage einer neuen Option für einen Harvesting-Vorgang

Aktuell kann aus zwei verschiedenen Vorlagen eine neue „Option“ konfiguriert werden:

- 1.) Ingest eines konkreten Objekts mit spezifischer URN und
- 2.) Ingest von beliebigen Objekten in der Zahl von:

Die erste Option gibt dem Benutzer die Möglichkeit ein spezielles Objekt über seine spezifische URN für den durchzuführenden Harvesting-Auftrag zu bestimmen. Hierfür muss lediglich im Feld „Wert“ die spezifische URN des Dokumentes, welches archiviert werden soll, angegeben werden. Diese Option kann bspw. für einzelne Funktionstests verwendet werden – andere Szenarien werden weiter unten beschrieben.

Durch drücken des Knopfes „Option hinzufügen“ wird die neue Option gespeichert und steht sofort für die Erzeugung eines Harvesting-Auftrages zur Verfügung.

Die zweite Auswahloption bietet die Möglichkeit eine Anzahl von beliebigen Objekten für einen Harvesting-Auftrag im Vorfeld festzulegen. In diesem Fall wird im Feld „Wert“ eine Anzahl zu archivierenden Objekte eingetragen. Die maximale Anzahl von Objekten, die innerhalb eines Harvesting-Auftrages dem DIAS-System übergeben werden kann, ist momentan auf 1000 Objekte festgelegt. Die Auswahl der Objekte erfolgt zufällig über die Anzahl aller Objekte, die noch nicht langzeitarchiviert worden sind und sich im Status: „ToDo“, im DNB-Repository befinden.

Beispiele:

Zu Möglichkeit 1.)

Scenario: Rückspielung eines migrierten Objektes in das DIAS-Archiv

Migrationen werden außerhalb des DIAS-Systems und innerhalb der DNB durchgeführt. Hierfür wird ein Objekt aus dem DIAS-System herausgeholt und migriert. Daran anschließend muss dieses Objekt erneut dem DIAS-System übergeben werden.

Im Fall eines einzelnen oder sehr weniger migrierter Objekte sollte der Harvesting-Auftrag über die URN des Objektes konfiguriert werden. Hiermit kann gezielt das entsprechende Objekt dem DIAS-System übergeben werden.

Der Vorteil dieser Methode besteht darin, dass eine vollständige Kontrolle über die einzuspielenden Objekte besteht. Alle zu einem Migrationsauftrag gehörenden Objekte können über diesen Weg gezielt und Termingerech eingespült werden.

Zu Möglichkeit 2.)

Scenario: Bearbeitung der Tagesproduktion aus dem Netzpublikationsgeschäftsgang

Aus dem Netzpublikationsgeschäftsgang kommen am Tage zwischen 200 und mehreren Tausend neue Netzpublikationen in die DNB und das DNB-Repository. Alle Objekte werden aus dem Import-Prozess heraus mit dem Flag „ToDo“ innerhalb des Repositories abgelegt. Das Flag „ToDo“ zeigt an, dass ein digitales Objekt noch nicht an das DIAS-System übertragen worden ist und gehört somit zur Menge aller Objekte die noch zu übertragen sind.

Um die Objekte dem DIAS-System zu übergeben und somit der Langzeitarchivierung, werden im LTP-Launcher üblicherweise tägliche Harvesting-Aufträge angelegt, welche die

Tagesproduktion aus dem Netzpublikationsgeschäftsgang an das DIAS-System übertragen sollen.

Da die Reihenfolge der Übertragung in diesem Fall zweitrangig ist wurde im LTP-Launcher eine Option für Harvesting-Aufträge angelegt, die es ermöglicht die Anzahl an zu übertragenden Objekten zu bestimmen, nicht jedoch deren Reihenfolge oder nach bestimmten URNs.

Nachteil dieser Option ist jedoch, dass keine Kontrolle über die Rangfolge der einzuspielenden Objekte besteht. Übersteigt die Tagesproduktion aus dem Netzpublikationsgeschäftsgang die im den Harvester-Aufträgen angegebene Höchstzahl an zu übertragenden Objekten an das DIAS-System pro Tag, können Objekte nicht sofort an das DIAS-System übergeben werden. Die Gefahr einer „Halden-Bildung“ auf dem Repository besteht.

Löschen und Ändern von Optionen:



Eine Liste aller bisher angelegten Optionen für Harvesting-Aufträge kann über das Rechner-Symbol im Menue der Benutzeroberfläche ausgegeben werde. Ebenfalls besteht über dieses Menue die Möglichkeit existierende Optionen zu bearbeiten (Werkzeugschlüssel) oder zu löschen (Papierkorb)

Harvesting-Konfiguration
Sie sehen hier alle möglichen Optionen, einen DIAS-Ingest anzustoßen.

Nr	Name / Konfigurationswert	Aktion
1	URN urn:nbn:de:0072-298945	 
2	URN urn:nbn:de:0072-307611	 
3	URN urn:nbn:de:0072-298945	 

Abbildung 2: Optionen Anzeigen bearbeiten oder löschen

Dieser Menüpunkt gibt eine die Liste aller bisher angelegten Optionen für einen Harvestingauftrag an.

- Löschen funktioniert nur, wenn keine Aufträge mehr mit dieser Option verknüpft sind. Sonst wird derzeit noch eine systeminterne Fehlermeldung angezeigt.

3.2 Anlegen von Harvesting-Aufträgen



Zur Erzeugung eines neuen Harvesting-Auftrages und zu dessen Konfiguration muss das grüne Plus-Symbol in der Benutzeroberfläche ausgewählt werden.

Wie in Abbildung 3 zu sehen stehen dem Benutzer in der ersten Zeile sämtlich verfügbare Optionen für die Konfiguration eines neuen Harvesting-Auftrages zur Verfügung.

Wenn eine beliebige aber fest vorgegebene Anzahl an Objekte in einem Harvesting-Auftrag zusammengefasst und an das DIAS-System übertragen werden sollen, kann eine Option (Anzahl: XY) ausgewählt werden. Anderenfalls muss eine Option mit spezifischer URN herausgesucht werden.

Festlegen des Startzeitpunktes:

Anschließend muss über „Start“ sowohl das Datum als auch die Uhrzeit festgelegt werden an dem der Harvesting-Auftrag gestartet werden soll. Die zur Verfügung gestellte Kalender-Funktion in der Benutzeroberfläche (Kalendersymbol) soll die Festlegung erleichtern helfen.

Wiederholungen festlegen

Als nächstes muss festgelegt werden, ob der Harvesting-Auftrag einmalig (Keine Wiederholungen) oder regelmäßig durchgeführt werden soll. Hierzu wurden unter „Wiederholungen“ unterschiedliche Optionen zur Verfügung gestellt.

Zuweisung von Aufgaben:

Harvesting-Aufträge können für unterschiedliche Aufgaben konfiguriert werden. In der nächsten Zeile kann über die Wahlmöglichkeit: „Übergabe an:“ festgelegt werden, welche Aufgabe mit den geharvesteten Objekten erfüllt werden soll. Zur Zeit steht lediglich zur Auswahl, dass alle aus dem DNB-Repository geharvestete Objekte dem „LtpLauncher“ übergeben werden, dessen Aufgabe es ist, die geharvesteten Objekte an das DIAS-System zu übermitteln.

Zukünftig ist angedacht denselben Harvester auch für Migrationsaufgaben zu verwenden. Hier müsste dann lediglich die Übergabe an den entsprechenden Migrationsvorgang ausgewählt werden.

Option: Dateiliste speichern: Für statistische Zwecke.



LTP - Launcher

DEUTSCHE
NATIONAL
BIBLIOTHEK



Neuen Harvestingauftrag einfügen

Sie können hier einen neuen Harvestingauftrag anlegen und entsprechend konfigurieren.

Wählen Sie hierbei eine Harvesting-Option (Anzahl = eine bestimmte Anzahl an Paketen an DIAS übertragen).

Unabhängig von der gewählten Harvesting-Option können derzeit maximal 1000 Objekte pro Harvestingvorgang verarbeitet werden.

Option:	Anzahl: 5
Start:	2013-03-13T17:48:00Z
Wiederholung:	Keine Wiederholung
Übergabe an:	LtpLauncher
Dateiliste speichern:	<input checked="" type="checkbox"/>
<input type="button" value="Harvestingvorgang einfügen"/> <input type="button" value="Werte zurücksetzen"/>	

Bei Fragen wenden Sie sich bitte an Stefan Hein (s.hein@dnb.de) oder Karlheinz Schmitt (k.schmitt@dnb.de)
Version 1.0.1 (07.03.2013)

Abbildung 3: Anlagen und Konfiguration eine Harvesting-Vorganges

3.3 Anhalten und Löschen von Harvesting-Vorgängen



Zur Kontrolle aller existierenden Harvesting-Aufträge kann innerhalb des LTP-Launcher's eine Liste aller beauftragten Harvesting-Aufträge angezeigt werden. Wie in Abbildung 4 zu sehen, existiert in diesem Bild nur ein Harvesting-Auftrag der jeweils in einem Intervall von 3 Stunde (Spalte „Intervall“) erneut durchgeführt wird. Spalte „Beginn“ in der Tabelle gibt an, zu welchem Zeitpunkt der nächste Harvesting-Vorgang für diesen Auftrag durchgeführt wird. In diesem Fall am 13.03.2013 um 17:08 Uhr. Die in Spalte 2 gegebene Task-Beschreibung gibt die Konfiguration des Harvesting-Auftrages wieder. Im abgebildeten Fall besagt die Konfiguration, dass bei jedem Vorgang 10 beliebige Objekte aus dem Repository der DNB geharvestet werden und dem LTP-Launcher zur Archivierung im DIAS-System übergeben werden.

Anhalten und Löschen von Harvesting-Aufträgen:

In der Spalte „Aktion“ stehen zur Zeit zwei verschiedene Aktionen zur Steuerung des Auftrags zur Verfügung. Über den Pause-Knopf können Harvesting-Aufträge pausiert werden. Dies ist evtl. notwendig, wenn Wartungsarbeiten an Komponenten des Services, innerhalb des DIAS-Systems oder der GWDG selber eine Unterbrechung der Einspielung erfordern. Um jedoch den Harvesting-Auftrag nicht vollständig zu löschen und neu aufsetzen zu müssen kann die Pause-Funktion verwendet werden.

Über das Papierkorb-Symbol können Harvesting-Aufträge gelöscht werden.



LTP - Launcher

DEUTSCHE
NATIONAL
BIBLIOTHEK



Aktuelle Warteschlange

Sie sehen hier alle derzeit laufenden bzw. geplanten SIP-BUILDER- und DIAS-Transaktionen.

Nr	Task-Beschreibung	Beginn	Intervall	Aktion
1	Anzahl 10	2013-03-13 17:08:00.0 (in 0 Stunden 4 Minuten)	Alle 3 Stunden	 

Bei Fragen wenden Sie sich bitte an Stefan Hein (s.hein@dnb.de) oder Karlheinz Schmitt (k.schmitt@dnb.de)
Version 1.0.1 (07.03.2013)

Abbildung 4: Liste aller beauftragten Harvesting-Vorgänge; Pausieren und Löschen von Harvesting-Vorgängen

3.4 Kontrolle von Harvesting-Vorgängen



Zur Kontrolle aller Harvesting-Vorgänge, die für einen existierenden Harvesting-Auftrag bereits durchgeführt worden sind, kann in der Spalte 2 „Task-Beschreibung“ eine Liste aller Harvesting-Vorgänge bzgl. des Harvesting-Auftrags eingesehen werden. Hierfür ist ein Klick auf die jeweilige Task-Beschreibung ausreichend.



Historie für Task 589831

Anzahl

Durchgeführte Transaktion	Gestartet / Beendet	Status	Aktion	
1	Anzahl: 10 ExportHandlerId: 10010	2013-03-15 09:18:00.0 2013-03-15 09:19:56.0	1 min, 56 s	✓ 10 / 10
2	Anzahl: 10 ExportHandlerId: 10010	2013-03-15 12:18:00.0 2013-03-15 12:19:57.0	1 min, 57 s	✓ 10 / 10
3	Anzahl: 10 ExportHandlerId: 10010	2013-03-15 15:18:00.0 2013-03-15 15:20:45.0	2 min, 45 s	✓ 10 / 10

Abbildung 5: Detail-Liste aller Harvesting-Vorgänge zu einem Harvesting-Auftrag.

Die Abbildung zeigt eine Liste aller Vorgänge, die zu einem laufenden Harvesting-Auftrag gehören. Zu jedem Vorgang können entweder

- die zugehörigen Statistiken und Events eingesehen werden (Datenblatt-Symbol),
- die zugehörigen Fehlermeldungen gelesen (Käfer),
- der Vorgang neu aufgesetzt werden oder (grüner Pfeil)
- der Vorgang selber gelöscht werden (Papierkorb).

Für einen detaillierten Einblick, kann über das Datenblatt-Symbol ein Überblick über alle Transaktionen aufgerufen werden. Die in Abbildung 6 aufgezeigte Übersicht zeigt, dass im betrachteten Harvesting-Vorgang neun von zehn Transaktionen erfolgreich durchgeführt worden sind. Der genaue Startzeitpunkt und der Zeitpunkt des Abschlusses des Harvesting-Auftrages werden angegeben als auch die gesamte Verarbeitungsdauer.

Auf DIAS-wartende Aktionen

In der Abbildung ist weiterhin ein Abschnitt zu sehen, der die Anzahl an Transaktionen wiedergibt, die vom DIAS-System noch nicht vollständig bearbeitet worden sind. Die Übertragung an das DIAS-System wurde auf Seiten der DNB vollständig abgeschlossen aber eine Antwort (DIAS-Ticket) über die erfolgreiche oder nicht erfolgreiche Aufnahme wurde noch nicht erhalten.

Dies kann der Fall sein, wenn es sich bei der Transaktion um sehr große Datenmengen handelt die bei der Aufnahme ins DIAS-System Zeit beanspruchen. Sollte sich in diesem Abschnitt Transaktionen befinden stellt dies nicht sofort ein Fehler dar. Sollte jedoch auch nach 24 Stunden diese Transaktion sich noch in diesem Bereich befinden, muss von

einem Fehler ausgegangen werden. Die Oberfläche gibt in diesem Fall dem Nutzer eine Möglichkeit diese spezielle Transaktion noch einmal durchzuführen.

Aktuell laufende Transaktionen

Alle Transaktionen, die in diesem Abschnitt aufgeführt werden, befinden sich noch in der DNB internen Verarbeitung (Status „RUNNING“). Im Normalfall sollten hier keine Transaktionen aufgelistet sein. Sollte dies jedoch der Fall sein, ist ein internes Problem aufgetaucht, dass dazu geführt hat, dass der Status der Transaktion von „RUNNING“ nicht auf „DONE“ gesetzt worden ist. In diesem Fall gibt die Oberfläche dem Nutzer die Möglichkeit den Status der Transaktion zurückzusetzen (Von RUNNING auf ToDo), so dass das Objekt erneut zur Aufnahme in das DIAS-System bereitsteht.

Auswertung von Task 589831



Zusammenfassung

Verarbeitungszeitraum: 2013-03-15 12:18:00.0 bis
2013-03-15 12:19:57.0
Verarbeitungsdauer: 1 min, 57 s
ExportHandler-Id: 10010
Harvesting-Option: Anzahl: 10

9 archivierte(s) Objekt(e)
0 auf DIAS wartende Transaktion(en)
1 abgebrochene Transaktion(en)
0 aktuell laufende Transaktion(en)
10 Gesamtzahl aller geplanten Vorgänge

Ausgabe

9 archivierte(s) Objekt(e)

Nr.	urn	Eingangsdatum	letzte Änderung	Status
1	urn:nbn:de:101:1-201212051914	05.12.2012 03:37:21	15.03.2013 12:20:00	✓
2	urn:nbn:de:101:1-201212051959	05.12.2012 03:37:47	15.03.2013 12:20:02	✓
3	urn:nbn:de:101:1-201212051968	05.12.2012 03:38:27	15.03.2013 12:20:03	✓
4	urn:nbn:de:101:1-201212051986	05.12.2012 03:39:20	15.03.2013 12:20:04	✓
5	urn:nbn:de:101:1-201212051995	05.12.2012 03:39:57	15.03.2013 12:20:05	✓
6	urn:nbn:de:101:1-201212052012	05.12.2012 03:40:17	15.03.2013 12:20:06	✓
7	urn:nbn:de:101:1-201212052025	05.12.2012 03:40:35	15.03.2013 12:20:06	✓
8	urn:nbn:de:101:1-201212052069	05.12.2012 03:42:08	15.03.2013 12:20:08	✓
9	urn:nbn:de:101:1-201212052082	05.12.2012 03:42:34	15.03.2013 12:20:09	✓

0 auf DIAS wartende Transaktion(en)

Nr.	urn	Eingangsdatum	letzte Änderung	Status
-----	-----	---------------	-----------------	--------

1 abgebrochene Transaktion(en)

Nr.	urn	Eingangsdatum	letzte Änderung	Status
1	urn:nbn:de:101:1-201212052049	05.12.2012 03:41:34	15.03.2013 12:20:08	✗

Abbildung 6: Detailsicht eines Harvesting-Vorgangs

Abgebrochene Transaktionen

Für eine Fehlersuche für die abgebrochene Transaktion kann durch anklicken des roten Kreuzsymbols die vom DIAS zurückgemeldete Fehlermeldung und die Ticket-Nummer des Vorgangs eingesehen werden (Abbildung 7). Hier sind sowohl die URN, als auch die

Ticket-Nummer des Ticket-Systems von DIAS angegeben. Zur Unterstützung der Fehlersuche, werden sowohl die Fehlermeldung als auch der Name des UOF-SIP's angezeigt. Das UOF-SIP kann im Error-Verzeichnis des DIAS-Produkktivsystems gefunden und analysiert werden.

urn:nbn:de:101:1-201212052049

Ausgabe

DIAS-Transfer fehlgeschlagen.

Fehlermeldung

DIAS reported: com.ibm.dias.architecturaltemplate.archive.EipTARArchive: extractFile



Transaktion erneut für den DIAS-Transfer vorbereiten.

Kriterium	Wert
Beginn der Transaktion	2013-03-15 12:19:41.0
Ende der Transaktion	2013-03-15 12:20:08.0
DIAS-interne ID	unbekannt
Ingest-Ticket	IBM_13633463633371
SIP-Name	1363346363245146.tar

Abbildung 7: Beispiel eine Fehlermeldung aus DIAS

3.5 Suche

Für die Suche nach durchgeführten Transaktionen kann innerhalb der Oberfläche die Suchmaske verwendet werden. Diese Option erreicht man über das „Fernglas-Symbol“. In Abbildung 8 ist die Suchmaske dargestellt.



Suche

Sie können hier entweder durch Angabe einer URN nach einem konkreten Paket suchen, oder sich abgelieferte Pakete (je nach Status der Transaktion) aus einem definierten Zeitraum anzeigen lassen. Bei Angabe einer konkreten URN hat der Status-Schalter keine Auswirkung.

URN:

Ingest von:

Ingest bis:

Status: archiviert
 nicht archiviert
 Archivierung aktuell in Bearbeitung
 Archivierung abgebrochen
 wartend auf DIAS-Bestätigung

Abbildung 8: Suchmaske des LTPLaunchers

Generell kann eine Suchanfrage entweder:

- Für eine spezifische URN oder
- Für einen spezifischen Ingest-Zeitraum

erfolgen. Das Ergebnis ist entweder eine einzelne Transaktion oder eine Liste aller für einen Zeitraum vorgefundene Transaktionen. Je nach gewählten Umfang des Suchzeitraumes kann die Ergebnisliste sehr umfangreich werden. Um die Ergebnisliste weiter einzuschränken kann in die Ergebnisliste anhand eines weiteren Kriteriums (Status) gefiltert werden. Es stehen hierfür fünf verschiedene nicht miteinander kombinierbare Filterkriterien zur Verfügung:

- archiviert,
- nicht archiviert,
- Archivierung aktuell in Bearbeitung
- Archivierung abgebrochen und
- wartend auf DIAS-Bestätigung.

Abbildung 9 zeigt das Ergebnis der Suche nach einer Transaktion für eine spezifische URN. Zu sehen ist, neben der URN, das Eingangsdatum, welches beschreibt wann das Dokument in das Repository der DNB aufgenommen worden ist, das Datum der letzten Änderung und der Status. In diesem Fall wird durch das rote Kreuz angezeigt, dass die Transaktion für die URN abgebrochen worden ist. Mit einem Klick auf das rote Kreuz kann die Fehlermeldung und sämtliche relevanten Daten zur Fehlersuche eingesehen werden.

Ein besonderes Augenmerk ist auf die in dieser Oberfläche eingebauten globalen Aktionen zu richten. Die erste globale Transaktion ermöglicht es den INGEST-Status für alle in der Liste als auf DIAS-wartend markierte Transaktionen den Status zu aktualisieren. Dies vereinfacht die Aktualisierung der Ergebnisliste der jeweiligen Suchabfrage.



Liste aller gesuchten Objekte

Sie sehen hier die Trefferliste der von Ihnen gesuchten Objekte und haben für noch nicht archivierte Objekte die Möglichkeit, direkt ein DIAS-Ingest anzustoßen oder sich über den bereits durchgeführten Ingest-Vorgang ins DIAS zu informieren.

Suchparameter:

Zeitraum: von 2012-01-01 bis heute
LZA-Status: true
in Bearbeitung: false
abgebrochen: false
wartend: false

Nr.	urn	Eingangsdatum	letzte Änderung	Status
1	urn:nbn:de:101:1-201212055984	05.12.2012 05:53:37	18.03.2013 09:20:41	✘

Seite 1

globale Aktionen

 [prüfe DIAS-Ingeststatus für alle wartenden Transaktionen.](#)

 [setze alle DNB-seitigen abgebrochenen Transaktionen auf TODO \(bereit für die Archivierung\).](#)

Abbildung 9: Ergebnisoberfläche für eine Suchanfrage nach einer spezifischen URN

Die zweite globale Aktion ermöglicht dem Benutzer alle in der Verwaltungsdatenbank als abgebrochene Transaktionen markierte Transaktionen wieder für die LZA-Anbindung bereitzustellen. Abgebrochene Transaktionen können entweder DNB-seitig abgebrochen oder von DIAS abgelehnte Objekte sein. Beim Zurücksetzen wird lediglich der Status der abgebrochenen Transaktionen von ERROR auf TODO gesetzt. Für die folgenden Harvesting-Aufträge stehen diese Transaktionen wieder zur Verfügung.

Beispiel:

Scenario: Suche nach allen abgebrochenen Transaktionen

Hierfür muss in der Suchmaske lediglich das Kriterium „Archivierung abgebrochen“ ausgewählt werden. Eine Beschränkung auf einen spezifischen Zeitraum muss nicht durchgeführt werden.

In diesem Fall werden alle abgebrochenen Transaktionen aufgelistet, die in der Datenbank als abgebrochen registriert worden sind. Abbildung 10 zeigt einen Auszug aus dieser Liste.

Liste aller gesuchten Objekte

Sie sehen hier die Trefferliste der von Ihnen gesuchten Objekte und haben für noch nicht archivierte Objekte die Möglichkeit, direkt ein DIAS-Ingest anzustoßen oder sich über den bereits durchgeführten Ingest-Vorgang ins DIAS zu informieren.

Suchparameter:

Zeitraum: von 2012-01-01 bis heute
 LZA-Status: false
 in Bearbeitung: false
 abgebrochen: true
 wartend: false

97 abgebrochene Transaktionen gefunden

Nr.	urn	Eingangsdatum	letzte Änderung	Status
1	urn:nbn:de:1111-201212205198	20.12.2012 21:48:33	25.05.2013 21:37:43	
2	urn:nbn:de:1111-20130109264	09.01.2013 03:58:59	01.06.2013 20:01:33	
3	urn:nbn:de:1111-20130109288	09.01.2013 04:05:50	01.06.2013 20:01:50	
4	urn:nbn:de:1111-201301102099	10.01.2013 06:08:28	02.06.2013 18:38:14	
5	urn:nbn:de:1111-20130111195	11.01.2013 03:13:45	03.06.2013 17:34:51	
6	urn:nbn:de:101:1-2013011211608	12.01.2013 07:58:19	04.06.2013 19:38:59	
7	urn:nbn:de:gbv:18-56244	10.12.2012 17:22:15	28.05.2013 19:12:32	
8	urn:nbn:de:gbv:18-44298	10.12.2012 17:29:12	05.06.2013 11:03:57	

Abbildung 10: Liste aller abgebrochenen Transaktionen seit Beginn.

In der Abbildung werden einleitend noch einmal die Suchparameter angegeben und daran anschließend die Gesamtzahl aller zutreffenden Transaktionen. In diesem Fall wurden 97 Transaktionen gefunden.

Fehlerkategorien:

In der Liste werden bisher zwei unterschiedliche Fehlerkategorien unterschieden

- 1.) Systemfehler und (Läufer mit rotem X)
- 2.) Fachliche Fehler (Rotes x)

Im ersten Fall werden Systemfehler, die vom DIAS-System gemeldet werden einsortiert: Abbildung 11 zeigt ein Beispiel:

Informationen zum DIAS-Transfer

Sie sehen hier relevanten Transaktionsdetails zur initialen Übertragung an DIAS.

urn:nbn:de:1111-201301102099

Ausgabe

DIAS-Transfer aufgrund einer DNB-seitigen Systemstörung fehlgeschlagen.

Fehlermeldung:

add file to uof failed: Technical metadata size exceeds DIAS limit of 262144 bytes



Transaktion erneut für den DIAS-Transfer vorbereiten.

Kriterium	Wert
Beginn der Transaktion	2013-01-10 06:03:54.0
Ende der Transaktion	2013-06-02 18:38:14.0
Transferpaketname	ddsid:ebok:978-3-642-16712-6

Abbildung 11: Beispiel eines DIAS-Systemfehlers

Die Abbildung 11 zeigt die vom DIAS-System zurück gemeldete Fehlermeldung inklusive der verfügbaren Transaktionsdaten. Mit Hilfe dieser Fehlermeldung kann nun die Fehlesuche beginnen. Für den Fall dass der Fehler beseitigt worden ist kann über den grünen Pfeil die Transaktion erneut für den Transfer an das DIAS-System vorbereitet werden.

Eine genaue Auflistung aller bekannten Fehlerfälle kann im JIRA-Ticket-System im Bereich „LZA-Ingest“ gefunden werden.

3.6 Statistik und Konfiguration

Abbildung 12 zeigt die für die LZA-Anbinung notwendigen Konfigurationselemente.



Konfiguration

Sie sehen hier die derzeit verwendete Konfiguration für das angebundene Langzeitarchivierungssystem.

Ausgabe

Name	Beschreibung	Wert
Access-Server	Adresse des Servers für den Access	
Ingest-Server	Der Ingest-Server des LZA-Systems	
Preload-Area	Upload-Bereich für SIPs	
LZA-Online-Status	Zeigt die Verfügbarkeit des LZA-Systems an	
DIAS-Retrieval-Interface	WebGUI für das Retrieval	
LZA-Monitoring	Monitoring-Informationen zum LZA-System	
SIP-Format	Ablieferungsformat eines SIPs	
DIP-Format	Auslieferungsformat eines DIPs	
DIAS-StorageSize	Größe des Gesamt-Speichervolumens in TB	
DIAS-Doku	DIAS-Dokumentation (GWDG gepflegt)	

Abbildung 12: Für die LZA-Anbindung mit dem DIAS-System notwendige Konfigurationselemente

Generell sind die Konfigurationsparameter nicht einseitig zu verändern. Eine Absprache mit dem externen Dienstleister von dem das DIAS-System gehostet und gewartet wird. Zur Zeit ist dies die Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen GmbH (GWDG).

Ein besonderes Augenmerk soll lediglich auf drei Aspekte dieser Übersichtsseite gerichtet werden:

- DIAS-Retrieval-Interface
- LZA-Monitoring und
- DIAS-Dokumentation

Das DIAS-System stellt eine Möglichkeit via (Web-Schnittstelle) zur Verfügung einzelne Objekte, durch die Angabe der externen oder internen URN, aus dem DIAS-System zu holen. Hierfür steht das DIAS-Retrieval-Interface zur Verfügung.

Monitoring

Gemeinsam mit der GWDG wird ein Monitoring-System für das DIAS-Produktivsystem aufgebaut worden. Dieses Monitoring-System, welches Verfügbarkeit, Auslastung und weitere Systemrelevante Daten des DIAS-Systems wiedergibt, ist über den angegebenen Link in der Zeile „LZA-Monitoring“ zu erreichen. Die Wartung des Monitoring-System erfolgt durch die GWDG.

DIAS-Dokumentation

Gemeinsam mit der GWDG wird eine WIKI-Seite aufgebaut und gepflegt, welche sämtliche Kommunikation und Dokumentationen bzgl. des DIAS-Systems und der Kommunikation zwischen der DNB und der GWDG enthält. Die Wiki-Seite ist über den angegebenen Link zu erreichen.

Statistik



Die LTPLauncher-Oberfläche bietet dem Nutzer eine Übersichtsseite, die einen zusammenfassenden Überblick über alle vom LTPLauncher durchgeführten Transaktionen wiedergibt. Darüber hinaus findet man Angaben über die Anzahl bereits langzeitarchivierter Objekte und die Anzahl von Objekten, die noch dem Langzeitarchiv übergeben werden müssen. Ein Überblick über die aktuelle Speicherauslastung des DIAS-System schließt die aktuelle Überblicksstatistik ab.

Ingeststatistik

Sie sehen hier eine Auswertung zum aktuellen Stand der LZA-Transaktionen. Ein Paket ist im Sinne dieser Anzeige archivierungsfähig, wenn es über den neuen Import-Service 2.0 (online seit **04.12.2012**) abgeliefert und erfolgreich verarbeitet wurde.

Ausgabe	
Pakete	
archivierungsfähige Pakete insg.	161613
Anteil archivierter Pakete	<div style="width: 23%;"></div> 23%
Gesamtzahl der noch nicht archivierten Pakete	131246
Gesamtzahl der bereits archivierten Pakete	30367
Transaktionen	
aktuell laufende Transaktionen	3
abgebrochene Transaktionen	97
wartende Transaktionen	0
Dateien	
Gesamtzahl der archivierten Dateien	41380
Speicher	
Speicherbelegung im Archiv	399169.01 MB
Speicherauslastung (bzgl. 180 TB verfügbarem Volumen)	<div style="width: 0%;"></div> 0%

Abbildung 11: Ingeststatistik über alle durchgeführten Transaktionen und Angaben zur Speicherauslastung im DIAS-System

4 Fehler-Handling

Die folgende Tabelle gibt einen Überblick über mögliche Fehlersituationen und stellt entsprechende Handlungsanweisungen zur Fehlerhebung bereit. Dem Leser wurden bereits einige Möglichkeiten zur Fehleranalyse mithilfe der Weboberfläche des LTPLaunchers vorgestellt. Darüber hinaus ist der gesamte Prozess der LZA-Anbindung auch an das DNB-interne JIRA-TicketSystem angeschlossen, d. h. die meisten Fehlerfälle führen automatisiert zur Erstellung eines entsprechenden Tickets. Damit steht dem Benutzer neben der Administration über die WebGUI ein 2tes Werkzeug für die Bearbeitung, Dokumentation und Statistik von Fehlerfällen zur Verfügung.

JIRA

Die folgende Tabelle gibt Auskunft über die derzeitige konfigurierte Anbindung an das JIRA-System:

JIRA-Space	lza-ingest
Link	-
Komponenten	<ul style="list-style-type: none">• DIAS-Ingest: Störungen bei der Verarbeitung von gesendeten SIPs auf Seiten von DIAS• DIAS-Metadata-Access : Störung bei Zugriff auf das DIAS-Datamanagement• DIAS-Retrieval: Störung beim Retrieval von in DIAS gespeicherten Objekten• DIAS-Ticket-System: Störungen beim Anfordern und der Statusabfrage an das DIAS-Ticket-System• LZA-Datenbank: Störung bei Zugriff auf die LZA-Verwaltungsdatenbank• SIP-Builder: Störung beim Bau des UOF-SIPs• SIP-Submitter: Störung beim Transfer eines SIPs an DIAS
Mail-Adresse	-
Nutzergruppen	lza-ingest-clearingstelle

HERANGEHENSWEISEN

Prinzipiell ist es die Entscheidung des Nutzers wie das Fehler-Handling im alltäglichen Betrieb gestaltet wird. Es gibt hierbei zwei mögliche Heransgehensweisen:

1. Reagieren auf Empfang von entsprechenden JIRA-Meldungen im Fehlerfall
2. Eigenständiges Prüfen mithilfe der WebOberfläche auf abgebrochene Transaktionen

Die aktuelle Erfahrung zeigt, dass ein bloßes Reagieren auf durch JIRA gemeldete Probleme (Variante 1.) für die Gewährleistung eines zuverlässigen Betriebes der LZA-Anbindung nicht ausreichend ist. Bestimmte Fehlersituationen wie z. B. das Stehenbleiben eines Harvestauftrages und ähnliche DNB-systeminterne Fehler werden im JIRA-System nicht immer sichtbar. Auch werden im JIRA-System Tickets gecacht, d. h. Fehler mit

derselben Ursache mitunter erst mit großer Verzögerung (derzeit 6h) ausgeliefert. Es ist also zu empfehlen, der Betreuung des LTP-Launchers zumindest ein regelmäßiges Zeitfenster im Tagesablauf einzuräumen. Es empfiehlt sich derzeit hierfür am Besten morgens eine halbe Stunde für die Betreuung einzuplanen.

FEHLERFÄLLE

Die nun folgende Tabelle stellt die derzeit bekannten Fehlersituationen und mögliche Maßnahmen zur Beseitigung zusammen.

Wenn nicht anders angegeben, beziehen sich alle Aktionen zur Fehlerbeseitigung auf Aktionen, die in der WebGUI des LTP-Launchers durchgeführt werden.

Zahlen in []-Klammern bezeichnen Fehler-IDs.

Allgemeine Hinweise

Fehler, die in der Ansicht „Ergebnisse abgeschlossener Tasks“ angezeigt werden, sind in der Regel keine permanenten Blocker, sondern beruhen zumeist auf sporadisch auftretene Systemstörungen. Eingegriffen werden sollte, wenn diese gehäuft über einen langen Zeitraum auftreten und zu einem Stillstand der LZA-Anbindung führen. Diese Kategorie betrifft derzeit die Fehler [1] und [2].

Id	Beschreibung	mögliche Ursachen	Bps.-JIRA-Ticket	Fehlerbehebung				
1	<p>Ein konfigurierte Harvesting-Auftrag wurde gestartet, aber es wurden keine Objekte verarbeitet.</p> <p>Ansicht: Ergebnisse abgeschlossener Tasks:</p> <table border="1" data-bbox="219 528 427 651"> <thead> <tr> <th data-bbox="219 528 327 587">Dauer</th> <th data-bbox="327 528 427 587">Status</th> </tr> </thead> <tbody> <tr> <td data-bbox="219 587 327 651">0 s</td> <td data-bbox="327 587 427 651"></td> </tr> </tbody> </table>	Dauer	Status	0 s		Tomcat-Server-Neustart, Netzwerkstörungen, Repository nicht erreichbar oder ausgelastet, LZA-Verwaltungsdatenbank nicht erreichbar	-	<p>Unter Aktionen des entsprechenden Tasks, rote Käfersymbol (Bug) anklicken, um Fehlermeldung einzusehen.</p> <p>Aktion</p>  <p>Fehler, die auf Systemstörungen (Repository, ItpSubmitter ist nicht erreichbar usw.) sollten bei wiederholtem Auftreten gemeldet werden. Bei sporadischen Systemstörungen läuft der nächste Task jedoch ohne Zutun wieder, d. h. es ist kein weiteres Eingreifen erforderlich.</p>
Dauer	Status							
0 s								
2	<p>Ansicht: Ergebnisse abgeschlossener Tasks:</p> <p>Zahl der erfolgreich geharvesteten Objekte liegt unterhalb der konfigurierten:</p> <p></p> <p><u>13 / 30</u></p>	Repository-Störung oder andere System-Störungen wie unter [1].	-	<p>Unter Aktionen, das Datenblatt-Symbol für weitergehende Analyse klicken.</p>  <p>Es kann sein, dass in der Zwischenzeit bereits die Pakete der mißglückten Transferversuche durch einen der nächsten Tasks wieder</p>				

				erfolgreich geharvetet worden sind.
3	<p>Ansicht: Aktuelle Warteschlange Die Anzeige „Harvesting läuft“ blinkt, obwohl der Beginn des Auftrags bereits lange zurückliegt und der Auftrag zeitlich bereits abgeschlossen sein müsste.</p> <div style="border: 1px solid gray; padding: 5px; background-color: #f0f0f0; text-align: center;"> <p>Beginn</p> <p>2013-06-05 19:37:00.0 Harvester läuft ...</p> </div>	System-Störungen wie unter [1], die zu langen langen Verarbeitungszeiten führen, so dass sich Aufträge ansammeln. Unkontrollierter Neustart des ItpLauncher und ItpSubmitters	-	<p>Alle geplanten Aufträge pausieren bzw. löschen und warten bis der noch laufende Task abgeschlossen ist.</p> <div style="border: 1px solid gray; padding: 2px; background-color: #e0e0e0; text-align: center; margin-bottom: 5px;"> <p>Aktion</p> </div> <div style="display: flex; justify-content: center; gap: 10px;">   </div> <p>Jedes Paket eines Tasks kann mitunter eine Wartezeit (TimeOut) von ca. 30min produzieren. Sollte der Auftrag nach einem Tag immer noch laufen, Produktverantwortlichen informieren.</p>

4	<p>Suche > Aktuell in Bearbeitung</p> <p>Eine Transaktion, die eigentlich bereits abgeschlossen sein müsste, steht immer noch auf RUNNING.</p> 	Interne System-Störungen wie unter [1]	-	<p>Sofern der Zeitraum mehr als 3h zurückliegt und sichergestellt ist, dass aktuell kein Auftrag läuft („Harvesting läuft“ in Ansicht: „Aktuelle Warteschlange“) kann durch Klick auf das „Läufer“-Symbol die Transaktion zurückgesetzt (auf TODO) werden.</p>
5	<p>Suche > Aktuell in Bearbeitung</p> <p>Eine Transaktion, die bereits abgeschlossen (grünes Häkchen) erscheint fälschlicherweise in dieser Ansicht.</p> 	<p>Bug. Es gibt offensichtlich 2 Ingest-Transaktionen (eine mit Status RUNNING und eine mit DONE), vermutlich durch einen sich zeitlich überlappenden wiederholten Import-Vorgang. Das sollte nur sehr selten auftreten.</p>	-	<p>Produktverantwortlichen informieren, damit doppelte Transaktions-Informationen aus LZA-Verwaltungsdatenbank gelöscht wird.</p>
6	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DNB-seitig abgebrochen:</p>  <p>Fehlermeldung: Error requesting a SIP ingest ticket from the DIAS. System message: Unable to parse DIAS response. Response corrupt</p>	Störung bei Zugriff auf das DIAS-Ticketsystem		<p>Durch klicken auf das Runner-Icon gelangt man zu Fehlerseite der betroffenen Transaktion.</p> <p>Da dieser Fehler (siehe Fehlermeldung) derzeit sporadisch auftaucht, sollte diese Transaktion zurückgesetzt werden.</p>  <p>JIRA-Ticket bitte entsprechend schließen.</p>

7	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DNB-seitig abgebrochen:</p>  <p>Fehlermeldung: add file to uof failed: Technical metadata size exceeds DIAS limit of 262144 bytes</p>	<p>Fehler bei Verarbeitung der Metadaten (mets.xml) – Metadaten-Volumen ist überschritten</p>		<p><i>Fehler wird derzeit noch untersucht.</i></p>
8	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DNB-seitig abgebrochen:</p>  <p>Fehlermeldung: SIP ingest ticket ID request failed. DIAS alert: Please wait, Maximum number of users are already accessing system</p>	<p>Störung bei Zugriff auf das DIAS-Ticketsystem</p>		<p><i>Warten bis das Dias-Ticket-system wieder neu gestartet ist! Denn aktuellen Harvestingvorgang solange pausieren lassen.</i></p>  <p>JIRA-Ticket bitte entsprechend schließen.</p>

9	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DIAS-seitig abgebrochen:</p>  <p>Fehlermeldung: Keine Fehlermeldung gefunden</p>	<p><i>Fehler wird derzeit noch untersucht.</i></p>		<p><i>Fehler wird derzeit noch untersucht.</i></p>
10	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DIAS-seitig abgebrochen:</p>  <p>Fehlermeldung: SIP ingest ticket ID request failed. DIAS alert: The SIP has been submitted successfully</p>			

11	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DIAS-seitig abgebrochen:</p>  <p>Fehlermeldung: z.B. no repository package for urn urn:nbn:de:101:1-201302226019 found</p>			
12	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DIAS-seitig abgebrochen:</p>  <p>Fehlermeldung: DIAS reported: java.lang.Exception: com.ibm.dias.loader.commands.Eip CommandConverter::execute - An error occurred while converting the metadata XML.</p>	Fehler bei Verarbeitung der Metadaten (mets.xml)		<p>Durck Klicken auf das Runner- Icon gelangt man zu Fehlerseite der betroffenen Transaktion.</p> <p>Da dieser Fehler (siehe Fehlermeldung) derzeit sporadisch auftaucht, sollte diese Transaktion zurückgesetzt werden.</p>  <p>JIRA-Ticket bitte entsprechend schließen.</p>

13	<p>Suche > Archivierung abgebrochen</p> <p>Eine Transaktion erscheint als DIAS-seitig abgebrochen:</p> <p></p> <p>Fehlermeldung: DIAS reported: Exception in substate: File check; com.ibm.dias.model.data.ModelException: File not exists: file:///./content/Meulen%202010%20Reading%20your%20daily%20newspaper%20without%20a%20PC.pdf</p>	<p>Problem bei Dateinamen mit Sonderzeichen (z. B. ein mit %20 encodiertes Leerzeichen). DIAS kann die Pfadangaben in der FileSec des Mets-Containers deshalb nicht richtig zuordnen.</p>		<p>Fehler muss DNB-seitig behoben werden. Bis dahin kein Eingreifen möglich.</p>
14	<p>Ansicht: Statistik</p> <p>Extrem hohe Werte bei „wartende Transaktionen“</p> <p>Der Werte sollte die aktuell verwendete Harvesting-Option (z. B. Anzahl=30) nicht überschreiten.</p> <p>Die PreLoad-Area (FTP-Location) enthält Pakete, die mit der Erweiterung „.IBM“ enden.</p>	<p>Fehler beim DIAS-Ingest. Das DIAS-Ingest-Modul scheint nicht zu funktionieren.</p>	-	<p>konfigurierte Aufträge (bis 18 Uhr) löschen/pausieren.</p> <p>FTP-Location prüfen. Mantis-Ticket erstellen, GWDG bitten, das Ingest-Modul zu starten.</p> <p>Wenn alle Pakete in der PreLoad-Area verarbeitet wurden, Harvestingauftrag neu einrichten.</p>

15	Ansicht: Statistik Hohe Werte (vgl. Harvesting-Option) bei laufende Transaktionen	Interne System-Störungen wie unter [1] oder anderweitige Störung.	-	konfigurierte Aufträge (bis 18 Uhr) löschen/pausieren. Produktverantwortlichen informieren.
----	---	---	---	--