

Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)

Eine gute Einführung in OAI-PMH findet sich hier: <https://libtechlaunchpad.com/2017/02/13/oai-pmh-basics-and-resources>.

Offizielle OAI-PMH-Dokumentation: <https://www.openarchives.org/OAI/openarchivesprotocol.html>.

In der Response steht jeweils ein responseDate und das Request-Verb. Beispiel:

```
<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2021-09-09T09:40:41Z</responseDate>
  <request verb="ListMetadataFormats">https://www.e-manuscripta.ch/oai/</request>
  ...
</OAI-PMH>
```

Verben

- GetRecord – Used to retrieve an individual metadata record.
- Identify – Used to retrieve repository information (ex. name, version).
- ListIdentifiers – Used to retrieve only headers.
- ListMetadataFormats – Used to retrieve the available metadata formats.
- ListRecords – Used to retrieve actual item metadata records.
- ListSets – Used to retrieve the set structure of a repository

Resumption Token

Die OAI-Schnittstellen liefern jeweils maximal eine bestimmte Anzahl Records zurück (bei uns in der Regel 10).

Der Tag `resumptionToken` enthält ein Attribut `completeListSize`, welches jeweils die gesamte Anzahl Treffer enthält. Diese Zahl ist nicht immer ganz exakt, sie kann sich während der Abfrage in Tranchen noch ändern.

Bsp:

```
<resumptionToken completeListSize="1279">0x726ca01fec751a64b6398913ea5ffe03-  
cursor_p_3D10_p_26set_p_3Dmaps_p_26metadataPrefix_p_3Doai_dc_p_26batch_size_p_3D11</resumptionToken>
```

Das Token selber enthält die Information, welche 10er-Tranche der Trefferliste als nächstes folgt.

Bsp. Für die Tranche ab Treffer Nr. 30:

```
<resumptionToken completeListSize="64">0x9bcb692a53b1f20609c8c4083fe1049a-  
cursor_p_3D30_p_26set_p_3Dzuz_p_26from__p_3D2021-09-  
15T00_p_253A00_p_253A00Z_p_26metadataPrefix_p_3Doai_dc_p_26batch_size_p_3D11_p_26until_p_3D2021-09-  
22T23_p_253A59_p_253A59Z</resumptionToken>
```

Für die Programmiersprache Python stehen die Libraries [Polymathea](#) und [Sickle](#) zur Verfügung, die das Harvesten von Daten via OAI-PMH vereinfachen.

Version #5

Erstellt: 8 Juli 2022 14:55:12 von Team Bibliotheks-informatik

Zuletzt aktualisiert: 20 September 2022 14:40:02 von Team Bibliotheks-informatik