OAI-PMH API

OpenArchives Initiative Protocol for Metadata Harvesting API

2024-07-15

|  |  |
| --- | --- |
|  | **Important**  These pages are under construction so the information given here might be incomplete or entirelly wrong. |

|  |  |
| --- | --- |
|  | API is available here: <https://api.aiscr.cz/oai?verb=Identify>. User registration is available here: <https://amcr.aiscr.cz/accounts/register/>. This guide is valid for version 2.0.0 of the service and schema. |

OAI-PMH AMČR API provides metadata from the AMČR database using the *OpenArchives Initiative Protocol for Metadata Harvesting* ([OAI-PMH](http://www.openarchives.org/pmh/)). The implementation supports two metadata standards:

* [Dublin Core](http://dublincore.org/) – part of the AMČR data set with metadata on documents and their descriptions,
* AMČR XML – native format with a complete representation of the whole AMČR database.

The API response is an *.xml* file.

## Access

The access to some metadata is protected by user roles:

* **A** (anonymous) – anyoune on the internet accessing AMČR services,
* **B** (researcher) – registered users,
* **C** (archaeologist) – archaeologist from a licensed organisation,
* **D** (archivist) – AMČR authorized archivist.

Registration to the system is available here: <https://amcr.aiscr.cz/accounts/register/>.

## Specification

### Schema

Schema (*.xsd*) for current version (**2.0**) of the service can be found here: <https://api.aiscr.cz/schema/amcr/2.0/>. Schema for previous version (1.0) is accessible here: <https://api.aiscr.cz/schema/amcr/1.0/>

???

### Filters

#### State and accessibility of the record

???

#### Sets

Selective querying can be achieved using predefined sets.

Available sets

| setSpec | setName | vocabulary |
| --- | --- | --- |
| projekt | Projekty / Projects | - |
| archeologicky\_zaznam | Archeologické záznamy / Archaeological Records | - |
| archeologicky\_zaznam:akce | Akce / Fieldwork Events | - |
| archeologicky\_zaznam:lokalita | Lokality / Sites | - |
| let | Lety / Flights | - |
| adb | Archeologické dokumentační body / Archaeological Documentation Points | - |
| dokument | Dokumenty / Documents | - |
| ext\_zdroj | Externí zdroje / External Sources | - |
| pian | PIAN / Spatial Units | - |
| samostatny\_nalez | Samostatné nálezy / Individual Finds | - |
| uzivatel | Uživatelé / Users | ✔ |
| heslo | Hesláře / Vocabularies | ✔ |
| ruian\_kraj | RÚIAN - Kraje / RÚIAN – Regions | ✔ |
| ruian\_okres | RÚIAN - Okresy / RÚIAN – Districts | ✔ |
| ruian\_katastr | RÚIAN - Katastrální území / RÚIAN - Cadastral Municipality | ✔ |
| organizace | Organizace / Organisations | ✔ |
| osoba | Osoby / Persons | ✔ |

#### Datestamps

In ListIdentifiers and ListRecords verbs it is possible to filter based on optional query parameter datestamp. (from – until).

???

## Login and authorization

To be able to access records with limited accessibility, it is neccessary to login using *Basic access authentication* while sending the requests.

### cURL

Use flag -u tu put in your username and password in each request.

curl -u <username>:<password> <GET request>

### Postman

In Postman, username and password can be set up on the *Authorization* tab. Select *Basic Auth* in the *TYPE* menu.

## Verbs

OAI-PMH protocol defines several verbs that allow metadata harvesting. OAI-PMH specification is available [here](http://www.openarchives.org/OAI/openarchivesprotocol.html).

### Identify

Verb Identify is used to get information about the repository.

*Request:* <https://api.aiscr.cz/oai?verb=Identify>

curl 'https://api.aiscr.cz/oai?verb=Identify'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-11T12:26:50Z</responseDate>  
 <request verb="Identify">http://digiarchiv:8080/oai</request>  
 <Identify>  
 <repositoryName>Archaeological Map of the Czech Republic (AMCR)</repositoryName>  
 <baseURL>https://api.aiscr.cz/oai</baseURL>  
 <protocolVersion>2.0</protocolVersion>  
 <adminEmail>info@amapa.cz</adminEmail>  
 <earliestDatestamp>1990-01-01</earliestDatestamp>  
 <deletedRecord>persistent</deletedRecord>  
 <granularity>YYYY-MM-DDThh:mm:ssZ</granularity>  
 <description>  
 <rightsManifest xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/rights/ http://www.openarchives.org/OAI/2.0/rightsManifest.xsd" appliesTo="http://www.openarchives.org/OAI/2.0/entity#metadata">  
 <rights>  
 <rightsReference ref="http://creativecommons.org/licenses/by-nc/4.0/rdf"/>  
 </rights>  
 </rightsManifest>  
 </description>  
 <description>  
 <oai\_dc:dc xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai\_dc/ http://www.openarchives.org/OAI/2.0/oai\_dc.xsd">  
 <dc:title xml:lang="cs">Archeologická mapa České Republiky (AMČR)</dc:title>  
 <dc:title xml:lang="en">Archaeological Map of the Czech Republic (AMCR)</dc:title>  
 [...]  
 <dc:identifier>https://api.aiscr.cz/</dc:identifier>  
 <dc:identifier>version 2.0.0</dc:identifier>  
 [...]  
 </oai\_dc:dc>  
 </description>  
 </Identify>  
</OAI-PMH>

### ListMetadataFormats

Verb ListMetadataFormats returns available metadata formats/standards.

*Request:* <https://api.aiscr.cz/oai?verb=ListMetadataFormats>

curl 'https://api.aiscr.cz/oai?verb=ListMetadataFormats'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-11T12:46:31Z</responseDate>  
 <request verb="ListMetadataFormats">http://digiarchiv:8080/oai</request>  
 <ListMetadataFormats>  
 <metadataFormat>  
 <metadataPrefix>oai\_dc</metadataPrefix>  
 <schema>http://www.openarchives.org/OAI/2.0/oai\_dc.xsd</schema>  
 <metadataNamespace>http://www.openarchives.org/OAI/2.0/oai\_dc/</metadataNamespace>  
 </metadataFormat>  
 <metadataFormat>  
 <metadataPrefix>oai\_amcr</metadataPrefix>  
 <schema>https://api.aiscr.cz/schema/amcr/2.0/amcr.xsd</schema>  
 <metadataNamespace>https://api.aiscr.cz/schema/amcr/2.0/</metadataNamespace>  
 </metadataFormat>  
 </ListMetadataFormats>  
</OAI-PMH>

Supported metadata formats are:

* Dublin Core – metadataPrefix=oai\_dc, schema available here <http://www.openarchives.org/OAI/2.0/oai_dc.xsd>,
* native AMČR *.xml* format – metadataPrefix=oai\_amcr, schema available here <https://api.aiscr.cz/schema/amcr/2.0/>.

### ListIdentifiers

Verb ListIdentifiers lists record headers. Required parameter is metadataPrefix, optional arguments allow filtering based on predefined sets (set) and/or datestamps (from, until). Only first page with limited amount of records is returned, to go to the next page, user has to submit a resumptionToken returned in the response.

*Request:* <https://api.aiscr.cz/oai?verb=ListIdentifiers&metadataPrefix=oai_amcr>

curl 'https://api.aiscr.cz/oai?verb=ListIdentifiers&metadataPrefix=oai\_amcr'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-11T12:51:03Z</responseDate>  
 <request verb="ListIdentifiers" metadataPrefix="oai\_amcr">http://digiarchiv:8080/oai</request>  
 <ListIdentifiers>  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/C-202013149</identifier>  
 <datestamp>2024-07-11T12:27:13.968Z</datestamp>  
 <setSpec>projekt</setSpec>  
 </header>  
 </record>  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/P-0334-000377</identifier>  
 <datestamp>2024-07-11T12:12:19.386Z</datestamp>  
 <setSpec>pian</setSpec>  
 </header>  
 </record>  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/C-DL-200400209</identifier>  
 <datestamp>2024-07-11T12:10:48.403Z</datestamp>  
 <setSpec>dokument</setSpec>  
 </header>  
 </record>  
 [...]  
 <resumptionToken completeListSize="859142" cursor="0">8F37FA46A1BE68D815CAE66EDFD7EE64</resumptionToken>  
 </ListIdentifiers>  
</OAI-PMH>

### ListRecords

Verb ListRecords lists records. Required parameter is metadataPrefix, optional arguments allow filtering based on predefined sets (set) and/or datestamps (from, until). Only first page with limited amount of records is returned, to go to the next page, user has to submit a resumptionToken returned in the response.

*Request:* <https://api.aiscr.cz/oai?verb=ListRecords&metadataPrefix=oai_dc>

curl 'https://api.aiscr.cz/oai?verb=ListRecords&metadataPrefix=oai\_dc'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-11T13:25:02Z</responseDate>  
 <request verb="ListRecords" metadataPrefix="oai\_dc">http://digiarchiv:8080/oai</request>  
 <ListRecords>  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/C-202013149</identifier>  
 <datestamp>2024-07-11T12:27:13.968Z</datestamp>  
 <setSpec>projekt</setSpec>  
 </header>  
 <metadata>  
 <oai\_dc:dc xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai\_dc/ http://www.openarchives.org/OAI/2.0/oai\_dc.xsd">  
 <oai\_dc:dc xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai\_dc/ http://www.openarchives.org/OAI/2.0/oai\_dc.xsd">  
 <dc:title xml:lang="cs">AMČR - projekt C-202013149</dc:title>  
 <dc:identifier>C-202013149</dc:identifier>  
 <dc:subject xml:lang="cs">projekt</dc:subject>  
 <dc:description xml:lang="cs">Stav: 6</dc:description>  
 <dc:type>http://base\_url/id/HES-000865</dc:type>  
 [...]  
 </oai\_dc:dc>  
 </oai\_dc:dc>  
 </metadata>  
 </record>  
 <resumptionToken completeListSize="859142" cursor="0">2A997D362431AC3FC670E2680E61D782</resumptionToken>  
 </ListRecords>  
</OAI-PMH>

### GetRecord

Verb GetRecord returns individual record. Required parameters are metadataPrefix and identifier (unique identifier of the record, URI). Unique identifier of individual record is in the header of each record in the element with tag <identifier>. The contents of <identifier> element is derived from a unique identifier of the given record stored in the <amcr:ident\_cely> prefixed with string https://api.aiscr.cz/id/.

*Request:* <https://api.aiscr.cz/oai?verb=GetRecord&identifier=https://api-test.aiscr.cz/id/M-FT-110598700&metadataPrefix=oai_amcr>

curl 'https://api.aiscr.cz/oai?verb=GetRecord&identifier=https://api-test.aiscr.cz/id/M-FT-110598700&metadataPrefix=oai\_amcr'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-15T09:04:59Z</responseDate>  
 <request verb="GetRecord" identifier="https://api-test.aiscr.cz/id/M-FT-110598700" metadataPrefix="oai\_amcr">http://digiarchiv:8080/oai</request>  
 <GetRecord>  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/M-FT-110598700</identifier>  
 <datestamp>2024-05-09T12:39:30.474Z</datestamp>  
 <setSpec>dokument</setSpec>  
 </header>  
 <metadata>  
 <amcr:amcr xsi:schemaLocation="https://api.aiscr.cz/schema/amcr/2.0/ https://api.aiscr.cz/schema/amcr/2.0/amcr.xsd http://www.opengis.net/gml/3.2 http://schemas.opengis.net/gml/3.2.1/gml.xsd"> HTTP/1.1 403 Forbidden </amcr:amcr>  
 </metadata>  
 </record>  
 </GetRecord>  
</OAI-PMH>

### ListSets

Verb ListSets lists available sets.

*Request:* <https://api.aiscr.cz/oai?verb=ListSets>

curl 'https://api.aiscr.cz/oai?verb=ListSets'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-11T13:28:01Z</responseDate>  
 <request verb="ListSets">http://digiarchiv:8080/oai</request>  
 <ListSets>  
 <set>  
 <setSpec>projekt</setSpec>  
 <setName>Projekty / Projects</setName>  
 <setDescription>  
 <oai\_dc:dc xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/oai\_dc/ http://www.openarchives.org/OAI/2.0/oai\_dc.xsd">  
 <dc:description xml:lang="cs">  
 Evidenční jednotky terénní činnosti badatelského nebo záchranného rázu evidované již ve fázi přípravy. Pro vymezení projektu je rozhodující podnět k výzkumu a provádějící subjekt (oprávněná organizace), lokalizace a příp. projektová dokumentace. Na projekt zpravidla navazuje jedna či (méně často) více terénních akcí či evidence samostatných nálezů.  
 </dc:description>  
 <dc:description xml:lang="en">  
 Records of field activities of a research or development-led nature recorded in the preparation phase. For the definition of the project, the motivation and the implementing body (authorised organisation), the location and, if applicable, the project documentation are decisive. The project is usually followed by one or (less frequently) more fieldwork, or the recording of individual finds.  
 </dc:description>  
 </oai\_dc:dc>  
 </setDescription>  
 </set>  
 [...]  
 </ListSets>  
</OAI-PMH>

## Parameters

Some verbs have required parameters (typically specifying the metadata format with metadataPrefix) and some have optional parameters allowing filtration, pagination etc.

Required and optional query parameters

| verb | required | optional |
| --- | --- | --- |
| Identify | - | - |
| ListMetadataFormats | - | - |
| ListIdentifiers | metadataPrefix | set, from, until, resumptionToken |
| ListRecords | metadataPrefix | set, from, until, resumptionToken |
| GetRecord | identifier, metadataPrefix | - |
| ListSets | - | - |

### Pagination

In case of ListIdentifiers and ListRecords verbs, only the first page of results is returned. To get another page, a resumptionToken returned at the end of the response must be used as a parameter in the following request. Parameter resumptionToken should be always the last parameter in the request.

*Request:* <https://api.aiscr.cz/oai?verb=ListRecords&metadataPrefix=oai_amcr>

curl 'https://api.aiscr.cz/oai?verb=ListRecords&metadataPrefix=oai\_amcr'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-15T11:57:23Z</responseDate>  
 <request verb="ListRecords" metadataPrefix="oai\_amcr">http://digiarchiv:8080/oai</request>  
 <ListRecords>  
 <record>  
 <header>  
 [...]  
 </header>  
 <metadata>  
 [...]  
 </metadata>  
 </record>  
 <resumptionToken completeListSize="859162" cursor="0">55CC4AC4A4BDAA3245F6B3E0E8B1CE0A</resumptionToken>  
 </ListRecords>  
</OAI-PMH>

*Follow up request with resumptionToken:* <https://api.aiscr.cz/oai?verb=ListRecords&resumptionToken=55CC4AC4A4BDAA3245F6B3E0E8B1CE0A>

curl 'https://api.aiscr.cz/oai?verb=ListRecords&resumptionToken=55CC4AC4A4BDAA3245F6B3E0E8B1CE0A'

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-15T12:01:28Z</responseDate>  
 <request verb="ListRecords" resumptionToken="55CC4AC4A4BDAA3245F6B3E0E8B1CE0A">http://digiarchiv:8080/oai</request>  
 <ListRecords>  
 <record>  
 <header>  
 [...]  
 </header>  
 <metadata>  
 [...]  
 </metadata>  
 </record>  
 <resumptionToken completeListSize="859162" cursor="100">AC59E153A7C3CDBAC4200F62CAC67880</resumptionToken>  
 </ListRecords>  
</OAI-PMH>

Notice the attributes of the resumptionToken element completeListSize indicating total number of results and cursor indicating that the first record in the given response is n-th returned record.

### Filters

Selective harvesting is allowed thanks to datestamps of records and sets.

#### Record accessibility

???

#### Sets

Available sets of records are returned using the ListSets verb. The contents of the <setSpec> element can be used in a set parameter in ListIdentifiers and ListRecords verbs.

*Request:* <https://api.aiscr.cz/oai?verb=ListIdentifiers&set=pian&metadataPrefix=oai_amcr>

curl 'https://api.aiscr.cz/oai?verb=ListIdentifiers&set=pian&metadataPrefix=oai\_amcr'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-15T12:17:21Z</responseDate>  
 <request verb="ListIdentifiers" set="pian" metadataPrefix="oai\_amcr">http://digiarchiv:8080/oai</request>  
 <ListIdentifiers>  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/P-1223-101288</identifier>  
 <datestamp>2024-07-15T11:13:45.237Z</datestamp>  
 <setSpec>pian</setSpec>  
 </header>  
 </record>  
 [...]  
 <resumptionToken completeListSize="219014" cursor="0">A828ECF79EF4EABB40CDD5EAB26F62B1</resumptionToken>  
 </ListIdentifiers>  
</OAI-PMH>

#### Datestamps

Parameters from and until in verbs ListIdentifiers and ListRecords allow filtering based on a datestamp of the given record. Use dates in ISO-8601 standard format, i.e. YYYY-MM-DD.

*Request:* <https://api.aiscr.cz/oai?verb=ListIdentifiers&set=projekt&from=2023-04-21&until=2024-04-27&metadataPrefix=oai_amcr>

curl 'https://api.aiscr.cz/oai?verb=ListIdentifiers&set=projekt&from=2023-04-21&until=2024-04-27&metadataPrefix=oai\_amcr'

*Sample response:*

<OAI-PMH xsi:schemaLocation="http://www.openarchives.org/OAI/2.0/ http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">  
 <responseDate>2024-07-15T12:20:21Z</responseDate>  
 <request verb="ListIdentifiers" set="projekt" from="2023-04-21" until="2024-04-27" metadataPrefix="oai\_amcr">http://digiarchiv:8080/oai</request>  
 <ListIdentifiers>  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/M-202204394</identifier>  
 <datestamp>2024-04-26T13:08:37.256Z</datestamp>  
 <setSpec>projekt</setSpec>  
 </header>  
 </record>  
 [...]  
 <record>  
 <header>  
 <identifier>https://api-test.aiscr.cz/id/C-202401944</identifier>  
 <datestamp>2024-03-17T06:32:08.143Z</datestamp>  
 <setSpec>projekt</setSpec>  
 </header>  
 </record>  
 <resumptionToken completeListSize="180105" cursor="0">6ABB1C7ACC374C52B5106D1ABCA113C6</resumptionToken>  
 </ListIdentifiers>  
</OAI-PMH>