

Inventory of persistent identifiers for cultural heritage objects and web resources provided to Europeana

Revision	<i>Version 1.1</i>
Last updated	<i>11/08/2023</i>
Author(s)	<i>Nuno Freire (Europeana Foundation), Hugo Manguinhas (Europeana Foundation), Antoine Isaac (Europeana Foundation), Valentine Charles (Europeana Foundation)</i>
Dissemination Level	<i>Public</i>

Version history

Revision no.	Date	Author	Organisation	Description
1.0	14/03/2023	<i>Nuno Freire, Hugo Manguinhas, Antoine Isaac, Valentine Charles</i>	Europeana Foundation	First version
1.1	05/10/2023	<i>Nuno Freire, Antoine Isaac</i>	Europeana Foundation	General revisions. Added listing of data providers using PIDs. New sections: 'related word', 'methodology', 'data quality' and references.

Table of Contents

1. Introduction	3
2. Related work	3
3. Methodology	4
4. Results	6
4.1 General PID usage statistics	6
4.2. Statistics on usage of ARK	10
4.3. Statistics on usage of HANDLE	11
4.4. Statistics on usage of URN	14
4.5. Statistics on usage of DOI	15
4.6. Statistics on usage of PURL	16
4.6. Use of multiple PIDs in the same Europeana record	16
5. Uniqueness of the persistent identifiers in Europeana across records and datasets	19
6. Data quality issues	21
7. Conclusion	21
References	22
Annex A - Europeana data providers using PIDs	24

1. Introduction

The importance of persistent identification has long since been recognized in cultural heritage (CH) and over the past years, several PID schemes have been adopted in CH and academia. For example, national libraries in several countries use the National Bibliography Number (NBN), an identification system for publications, which predates modern Persistent Identifier (PID) schemes.

PIDs are relevant for embracing the FAIR principles (PIDs without persistence policy or with non-registered prefixes are less commendable and may be used to refine data governance policies and data quality reporting. To inform future decisions regarding the use of PIDs in the common European data space for cultural heritage¹, we have analysed the usage of PIDs in the metadata that CH institutions (CHIs) deliver to Europeana, focusing on identification of CH objects and their digital representations, reliability of the PIDs, and data quality issues.

2. Related work

Several projects and researchers have addressed the topics of persistent identification in cultural heritage in recent years. In [1], McKenna & Fokke present PIDs requirements and recommendations for CH institutions, and recently, Koster has conducted a comprehensive analysis of PIDs for applications in CH [2], examining the current infrastructural practices, and existing PID systems, while highlighting their advantages and disadvantages. There has also been some work carried out in the context of Europeana in 2013, when an PID resolution service for Europeana was prototyped [3].

The actual persistence of DOI identifiers, a highly relevant PID scheme in the scholarly domain, was for example investigated by Klein & Balakireva, 2021. From their experiments, the authors concluded that persistence is not guaranteed, and responses to DOIs may return inconsistent results over a period of time [4]. In our study, we have not investigated the persistence of PIDs, but we identified reliability issues in the PIDs present in the metadata.

Worldwide, or even Europe-wide, quantitative data about the current PID usage in CH institutions is not available. It is probably possible to extract some figures from the previous mentioned works or general web studies². Also, one could try to obtain the number of, say, DOIs that are minted by doi.org, from online reports from these initiatives³. But these are not focused on CH, and thus probably do not give hints on the adoption of PIDs in CH (i.e. how much of the collections are provided with PIDs). They also will not tell us much about what is identified by PIDs, i.e. which resource has a PID, and where do PIDs appear in the metadata.

1

<https://digital-strategy.ec.europa.eu/en/news/commission-proposes-common-european-data-space-cultural-heritage>

2

[https://www.researchgate.net/publication/324633938 Can Common Crawl Reliably Track Persistent Identifier PID Use Over Time](https://www.researchgate.net/publication/324633938_Can_Common_Crawl_Reliably_Track_Persistent_Identifier_PID_Use_Over_Time)

3

<https://www.doi.org/the-identifier/resources/factsheets/key-facts-on-digital-object-identifier-system>.

Closer to CH, one can find relevant surveys, including one at Europeana⁴ and one conducted by the Heritage PIDs project in the United Kingdom in 2020, which indicated that more than 75% of the respondents were aware of PIDs, and more than 50% of the CH institutions were using PIDs [5, 6]. Our work aims to contribute additional information about how PIDs are currently being actually used by CH institutions.

3. Methodology

In order to know more accurately which, and to what extent, the different PID schemes are nowadays in use by CH institutions in Europe, we processed the metadata that is delivered to Europeana by these institutions.

To meet our goal, we processed the metadata that CHIs deliver to Europeana, using a dump of the Europeana dataset⁵ from February 2023. Although CHIs may use PIDs for identifying several type of entities⁶, for this study our interest is on the identification of "original" CH objects ("CHOs" in the Europeana Data Model - EDM⁷) and their digital representations ("web resources" in EDM). This means that we exclude PIDs used for, e.g., concepts or persons that appear as the subject or creator of a CHO.

All references to cultural heritage objects (CHO) and of their web resources were checked for candidate PID schemes, by inspecting the following data elements of the EDM:

- Property dc:identifier in ore:Proxy (the ore:Proxy from Europeana was excluded)
- Property edm:isShownBy in ore:Aggregation
- Property edm:isShownAt in ore:Aggregation
- Property edm:hasView in ore:Aggregation

Note that the metadata made available by Europeana includes more data than that of the providers (i.e., it also includes data from intermediate aggregators and Europeana itself), but in our investigation we analysed only the metadata originally provided by the CH institutions. This was possible because EDM allows the provenance of the metadata to be traced across all aggregation steps, via a data modelling pattern based on OAI-ORE (Open Archives Initiative Object Exchange and Reuse⁸) proxies and aggregations.

It is important to also note that though the rdf:about of the edm:ProvidedCHO can also contain values that may correspond to a PID, it is not possible to check if a persistent URI was sent by the provider, since the original value of the rdf:about is not kept in the metadata after the ingestion process of Metis (nb. new datasets processed via Metis keep the original value found in rdf:about as dc:identifier).

Based on our knowledge of the dataset and on an ad-hoc search across the values of the EDM elements that may contain PIDs, we decided to investigate the usage of five PID schemes:

⁴ The call for participation in the survey is available at: <https://pro.europeana.eu/post/help-us-to-make-cultural-heritage-data-more-persistent>

⁵ Obtained via <https://pro.europeana.eu/page/harvesting-and-downloads#downloads>

⁶ ARK identifiers, for example, have been assigned to bibliographic records, persons, organisations, vocabulary terms, archeological artefacts, etc. Cf. Kunze, J.: ARK identifiers FAQ. (2021). Retrieved from

<https://wiki.lyrasis.org/display/ARKs/ARK+Identifiers+FAQ#ARKIdentifiersFAQ-ARKedThingsWhatkindsofthingsareARKsassignedto?>

⁷ <http://pro.europeana.eu/edm-documentation>

⁸ <https://www.openarchives.org/ore/1.0/>

- Archival Resource Key (ARK)
- Digital Object Identifier (DOI)
- HANDLE⁹
- Persistent URL (PURL)
- Uniform Resource Name (URN)

Table 1 describes how the usage of these identifiers was checked. When applicable, we tried to detect PIDs expressed in both forms of an HTTP(S) URI or in compact URI form (see Table 1 for examples)¹⁰. The detection methods also take into consideration that some PIDs may be expressed as HTTP(S) URIs using different (resolver) hostnames.

Table 1. The detection methods applied in this investigation for the five candidate PID schemes

PID Scheme	Detection methods	Examples in HTTP(S) URI and compact URI forms
ARK	<ul style="list-style-type: none"> • When an HTTP(S) URI has a path that starts with 'ark:' • Or, when an URI start with 'ark:' 	https://n2t.net/ark:/12148/btv1b8449691v/f29 ark:/12148/btv1b8449691v/f29 ark:12148/btv1b8449691v/f29
DOI	<ul style="list-style-type: none"> • When an HTTP(S) URI has the host 'doi.org' or 'dx.doi.org' • Or, when an URI start with 'doi:' 	http://doi.org/10.2298/BG20130206SPEHAR doi:10.2298/BG20130206SPEHAR
HANDLE	<ul style="list-style-type: none"> • When an HTTP(S) URI has a host that starts with 'handle.' or contains '.handle.' • Or, when an URI start with 'hdl:' 	http://hdl.handle.net/10062/5516 hdl:10062/5516
PURL	<ul style="list-style-type: none"> • When an HTTP(S) URI has a host that starts with purl.' or contains '.purl.' 	http://www.purl.org/yoolib/inha/10136
URN	<ul style="list-style-type: none"> • When an HTTP(S) URI has a path that starts with '/urn:'¹¹ • Or, when an URI start with 'urn:' 	http://urn.fi/URN:NBN:fi-fd2010-00003198 URN:NBN:fi-fd2010-00003198

For the inventory we have also analysed the prefixes of some of the candidate PID schemes. The prefixes allow the identification of relevant registration authorities and namespaces of a PID scheme, which may give a good indication of the reliability of the persistence of the PIDs. Another relevant aspect is the uniqueness of the PIDs in Europeana. With this in mind, we checked for the existence of cases where more than one CHO have an equal persistent identifier, and also for cases of CHOs that contain more than one persistent identifier.

⁹ DOIs are also based on the HANDLE system (prefix '10') but given their prominence, we decided to analyse them separately.

¹⁰ All of these compact schemes (prefixes) can be resolved by identifiers.org and by n2t.net. We are not certain however if the URN resolvers they both redirect to can cope with all the NIDs present in Europeana.

¹¹ Although the URN scheme does not mandate the existence of resolution services, some cultural heritage institutions have deployed resolution services for the namespaces under their responsibility. The most common of these cases are the national libraries that are responsible for the National Bibliography Numbers and use the NBN namespace of the URN scheme. For example, <https://urn.nsk.hr/urn:nbn:hr:211:395205> and <http://urn.fi/URN:NBN:fi-fd2010-00003211>.

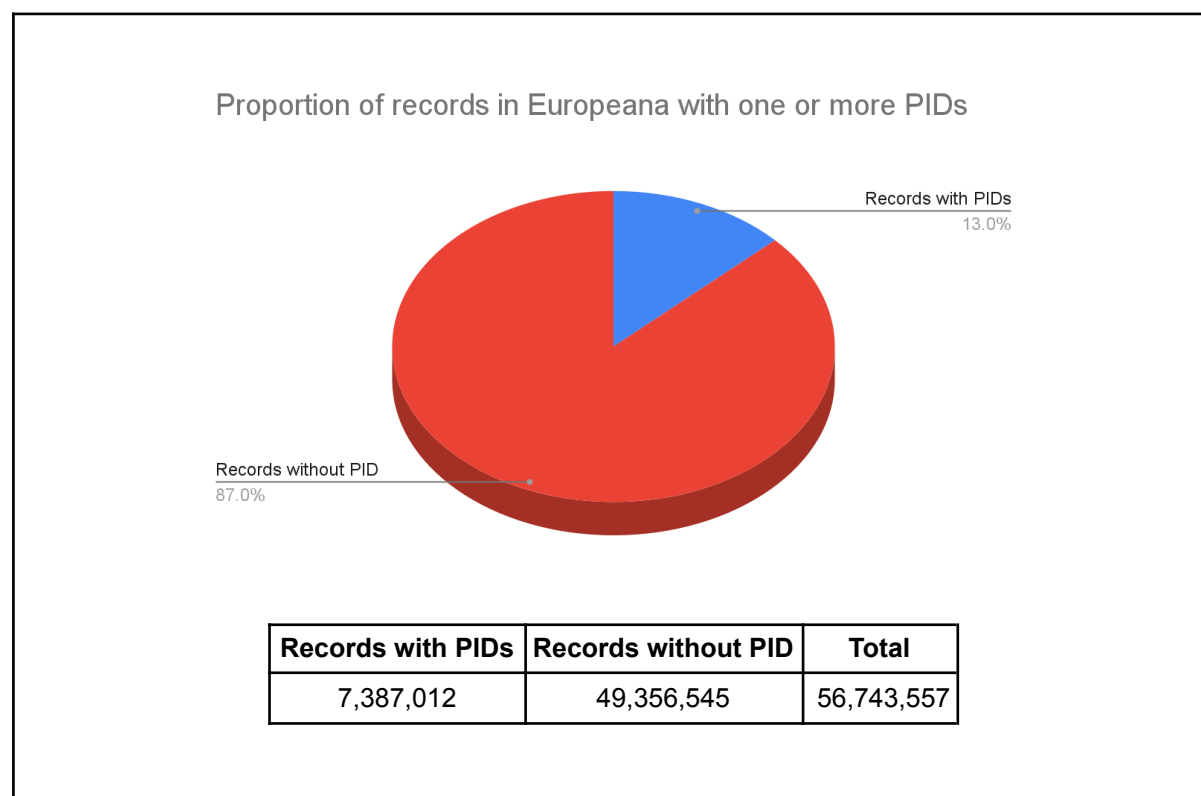
In EDM, a PID in the dc:identifier of an ore:Proxy identifies a CHO, while a PID in edm:isShownBy, edm:isShownAt and edm:hasView properties identifies a web resource. From the semantics of edm:isShownAt and edm:isShownAt properties, we could also draw some conclusions as to whether the PIDs are resolving to a webpage that gives access to the digital resource, or directly to a media file¹².

4. Results

This section presents statistics on PID usage in general and for individual schemes, and reports on the (non-)uniqueness of PIDs and other data quality issues.

4.1 General PID usage statistics

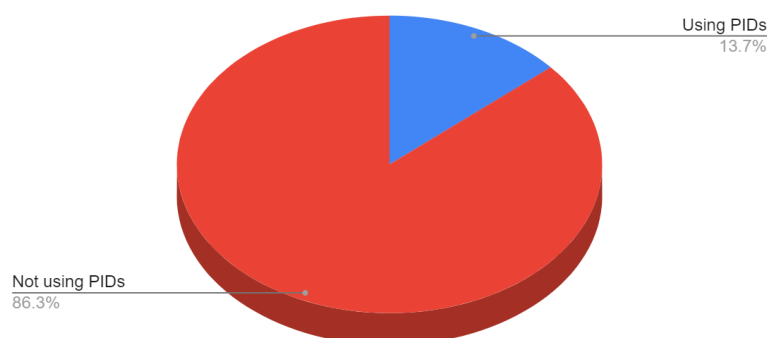
The following charts present an overview of the proportion of each PID scheme found in Europeana data, and in which properties they are being used. We have found 13% of records (including tier 0 records) in Europeana to contain at least one PID.



The existing PIDs originate from 512 data providers, corresponding to approximately 13.7% of the Europeana data providers. The list of data providers using PIDs may be consulted in Annex A.

¹² This assumes that these properties are used in the way recommended by the EDM, which has not always proven to be the case and should in future work be further verified

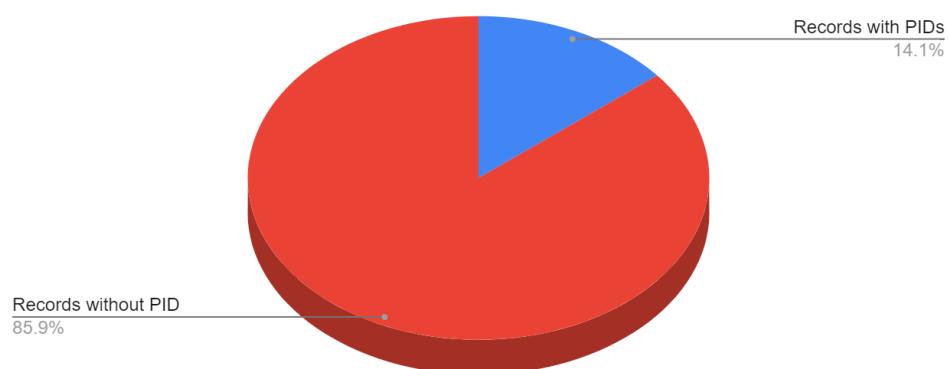
Proportion of Europeana data providers using PIDs



Using PIDs	Not using PIDs	Total providers
512	3,227	3,739

We have also analysed the Europeana dataset excluding records that are in tier 0 from the Europeana Publishing Framework. In this subset of Europeana, 14.1% of records contain at least one PID.

Proportion of non-Tier0 records in Europeana with one or more PIDs



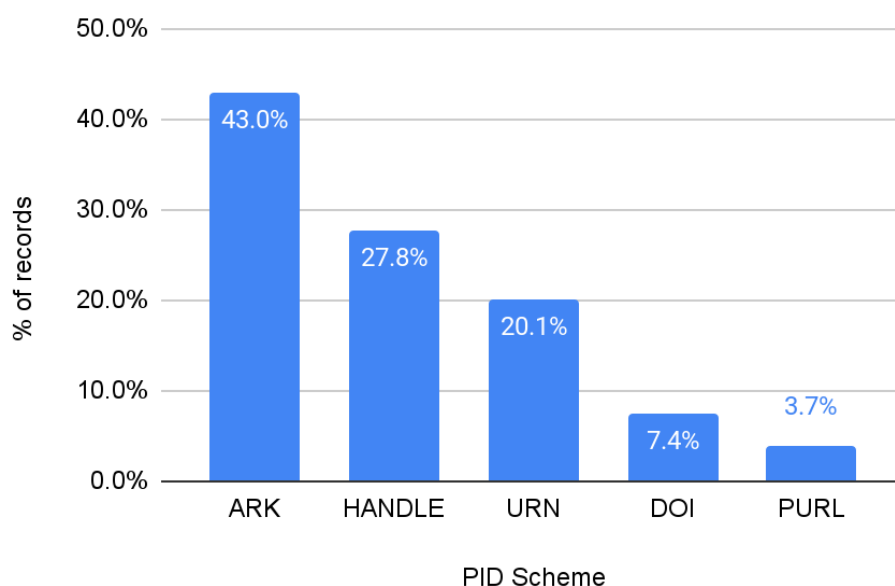
Records with PIDs	Records without PID	Total
7,121,542	43,467,850	50,589,392

NOTE: in the remainder of this document the tables and charts are based on all the Europeana records (including tier 0).

The following two charts indicate the proportion of Europeana records that an individual PID scheme appears in and their representation counting all PIDs that appear in Europeana. Note that it is common for records to contain more than one PID - sometimes from different PID Schemes - since CHOs and web resources may - and should! - be referenced using different PIDs, and dc:identifier is a repeatable property in EDM.

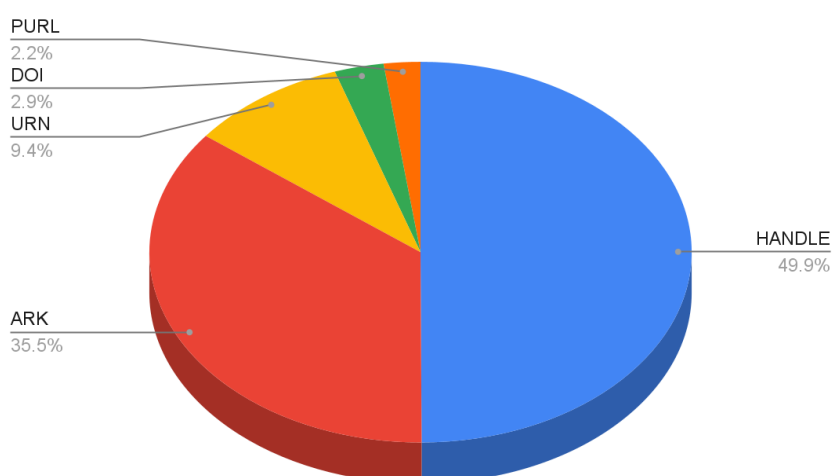
ARK and HANDLE are the most frequently used candidate schemes, with ARK occurring in more records, but HANDLE has the highest number of PIDs (as shown in Section 4.3, the high number of HANDLES originates from uses in edm:hasView).

Proportion of PIDs schemes in use in records



PID Scheme	Record count	% of records
ARK	3,176,131	43.0%
HANDLE	2,051,327	27.8%
URN	1,482,709	20.1%
DOI	549,298	7.4%
PURL	275,196	3.7%

Proportion of PID schemes in statements with PIDs

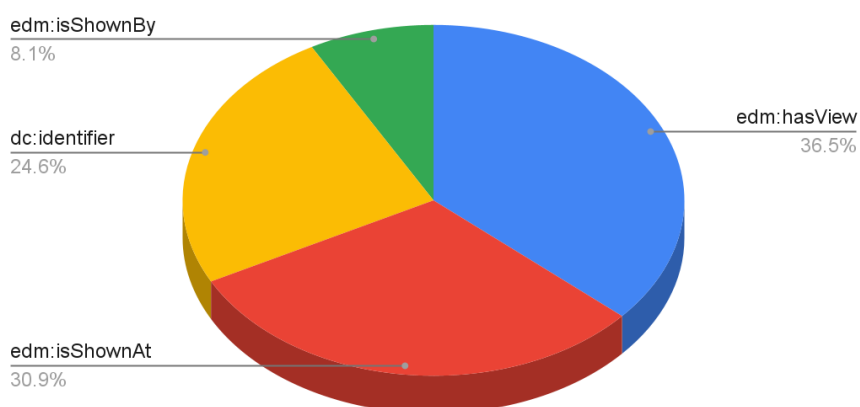


PID Scheme	Total PIDs
HANDLE	9,369,773
ARK	6,664,671
URN	1,768,135
DOI	551,634
PURL	409,677
Total	18,763,890

The table below shows the metadata properties where PIDs are used. `edm:hasView` contains the most PIDs, followed by `edm:isShownAt`, `dc:identifier` (of the CHO) and `edm:isShownBy`. The EDM semantics allow the values of `edm:hasView` - which is repeatable, thus contributing to its ranking first - to refer to either (1) web pages giving access to a digital object or (2) media files directly. PIDs in `edm:isShownBy` shall resolve directly to media files and those in `edm:isShownAt` and CHOs' `dc:identifier` to web pages that give access to the digital object and `dc:identifier` could identify the object itself¹³. Our results thus hint that a majority of the PIDs resolve to web pages rather than to media files.

¹³ I.e., this would be the identifier of a "real-world object" in the Semantic Web paradigm (see <https://www.w3.org/TR/cooluris/#semweb>). One would expect such PID to resolve "for humans" to web pages, but for machines the same identifier could resolve to some data (see <https://www.w3.org/TR/cooluris/#solutions>). Checking the extent to which this happens could be included in future work.

Proportion of properties with PIDs



Property with PIDs	Total PIDs
edm:hasView	6,841,900
edm:isShownAt	5,795,176
dc:identifier	4,613,432
edm:isShownBy	1,513,382
Total	18,763,890

4.2. Statistics on usage of ARK

ARK is the PID candidate scheme used in most records in Europeana. A very small percentage of ARKs are used in `edm:isShownBy` (1.7%). This is an indication that ARKs typically resolve to a webpage that gives access to the digital object. But we cannot draw general conclusions here since the ARKs in Europeana originate from only four Name Assigning Authorities¹⁴ (see Table 3), and one of them (the Bibliothèque nationale de France) is by far the greatest source. In fact, diving deeper we can see an association between the assigning authorities and the usage of ARKs in some properties: all cases of use in `edm:isShownBy` and `edm:hasView` were assigned by Data Archiving and Networked Services from the Netherlands.

¹⁴ <https://arks.org/about/ark-naans-and-systems/>

Table 2. Distribution of ARK usage across properties

Property	Count	%
edm:isShownAt	3,024,583	45.4%
dc:identifier	2,150,108	32.3%
edm:hasView	1,374,557	20.6%
edm:isShownBy	115,423	1.7%
total	6,664,671	

Table 3. The NAAN (Name Assigning Authority Number) present in the ARKs in use in Europeana

NAAN	Authority	Count
12148	Bibliothèque nationale de France (National Library of France)	6,395,780
73189	Data Archiving and Networked Services	200,103
81055	British Library	64,831
86084	Blavatnik Archive Foundation	3,954

4.3. Statistics on usage of HANDLE

The HANDLE PID scheme comes second in number of records using it, but it has the highest number of total PIDs because it is used very often in the repeatable edm:hasView. Similarly to the usage of ARKs, the results indicate that HANDLES most frequently resolve to a webpage that gives access to the digital object, but the usage of HANDLES in edm:isShownBy reaches a significant 8.7%. These cases of HANDLES in edm:isShownBy originate from a single provider that is using HANDLES to identify web resources (images) and the HANDLES resolve to IIIF Image API requests¹⁵.

Table 4. Distribution of HANDLE usage across properties

Property	Count	%
edm:hasView	5,363,603	57.2%
edm:isShownAt	1,871,274	20.0%
dc:identifier	1,334,324	14.2%
edm:isShownBy	800,572	8.5%
total	9,369,773	

A large number of distinct HANDLE prefixes (74, see Table 5) are present in Europeana: more Europeana providers are using HANDLE than ARK. '10648', the prefix with highest count (10648) is Friedrich-Ebert-Stiftung, followed by '11088', which is the Rijksmuseum. It is relevant to highlight that two prefixes are Multi-Primary Administrators (MPAs) of the DONA Foundation:

- Prefix '21' - Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG), which is the institution responsible for the assignment of HANDLES in ePIC, a European consortium that provides PID services for the European Research Community,

¹⁵ For example <http://hdl.handle.net/10622/30051000549359>

- Prefix '20' - Corporation for National Research Initiatives (CNRI), an organisation formed for supporting research and development for the USA's national information infrastructure.

On the other hand, some HANDLE prefixes are dubious¹⁶. We also identified a data provider whose HANDLE resolver was offline for some months (one of their now working items is <http://hdl.handle.net/10797/14218>).

Table 5. The HANDLE prefixes in use in Europeana

Prefix	Count
10648	5,302,114
11088	1,003,623
10934	679,048
10891	348,611
10622	321,454
10796	237,488
10687	196,585
10366	189,365
10107	174,401
10970	170,529
20	134,066
2027	107,375
10999	83,457
11441	43,313
10251	33,198
10062	31,898
11525	26,471
21	26,268
10442	20,911
1839	18,559
11259	16,469
11649	15,728
1887	15,110

¹⁶ For example <http://hdl.handle.net/123456789/54611> returns the following message: "This handle is not valid. The numbers 123456789 are just a place holder for a valid handle prefix. It appears the DSpace site you encountered has not yet implemented handles. You will need to send an email to a contact from the DSpace site. There is nothing we can do from this end. Good luck."

<http://hdl.handle.net/987654321/76135> (which appears at https://data.europeana.eu/item/2022712/lod_oai_gredos_usal_es_987654321_76135_ent0) returns a "Prefix [987654321] Not Found" error.

Prefix	Count
123456789	13,397
11533	13,043
10797	12,713
10459	11,398
11631	10,151
11653	8,633
11728	6,236
10357	6,099
11650	5,911
11725	5,801
11739	5,588
11013	5,188
10347	5,144
10405	5,114
11741	4,246
11051	4,194
11638	3,989
11609	3,899
11340	3,762
10651	3,268
2454	2,872
11736	2,772
10272	2,652
11737	2,424
11734	2,258
10017	2,176
10316	2,106
10400	1,914
10419	1,721
11688	1,460
10498	1,104
11219	1,102
11412	609
2437	508
11735	402
10411	390

Prefix	Count
11596	321
10192	224
11535	149
987654321	56
11702	34
2077	6
10810	2
2043	2
10045	2
10486	1
11705	1
10071	1
10138	1

4.4. Statistics on usage of URN

URN comes third and is used differently from ARKs and HANDLES; URNs are indeed mostly used for identifying CHOs (dc:identifier) (58.3%). The number of uses in edm:isShownAt and edm:isShownBy (to refer to web resources) is similar.

Table 6. Distribution of URN usage across properties

Property	Count	%
dc:identifier	1,031,407	58.3%
edm:isShownAt	379,304	21.5%
edm:isShownBy	357,326	20.2%
edm:hasView	98	0.0%
total	1,768,135	

Six URN Namespace Identifiers (NID) are present in Europeana. The NID for the National Bibliography Number ('NBN') is clearly the main one being used (70% of the existing URNs). Table 7 lists all the NIDs in use, and also indicates the country code in the case of NBN. Within NBN, we found 14 country codes, 5 having over a thousand PIDs: Germany, Norway, Finland, Croatia and Slovakia.

A closer look at URN namespaces raises concerns of the general reliability of URNs. The URN namespace is often used liberally. Some of the NIDs appear to have been used ad-hoc, and may not indicate any commitment for persistent identification. Especially, three of the NIDs are not registered in IANA¹⁷, and some of the URNs under the NBN NID have invalid country codes (for example, 'nbn:imp272'). There are also 254,643 incomplete URNs that omit the NID (for example, 'URN:db_read/ft/52747/1928'). Further investigation is required to determine which URNs ensure persistent resolution services.

¹⁷ The three cases are 'imss', 'rs' and 'repox.ist.utl.pt'. Cf. the IANA URN NID registry: <https://www.iana.org/assignments/urn-namespaces/urn-namespaces.xhtml>

Table 7. The URN NIDs (Namespace Identifiers) in use in Europeana

NID	Count	Registered in IANA?
nbn:de	542,838	Yes
nbn:no	358,032	Yes
nbn:si	333,284	Yes
repox.ist.utl.pt	197,832	No
rs	46,691	No
nbn:fi	13,271	Yes
imss	9,219	No
isbn	4,866	Yes
nbn:hr	2,408	Yes
nbn:sk	2,390	Yes
nbn:at	966	Yes
nbn:cz	856	Yes
nbn:pt	542	Yes
issn	64	Yes
nbn:lv	20	Yes
nbn:imp272	6	Yes
nbn:fr	2	Yes
nbn:es	1	Yes
nbn:se	1	Yes
nbn:imp269	1	Yes
nbn:imp268	1	Yes
nbne	1	No
nbn:ch	1	Yes

4.5. Statistics on usage of DOI

Although very well-known in academia, DOIs are not frequent in Europeana. They almost always resolve to web pages, but a few cases exist where they resolve to media (typically, to a PDF file). Note that upon inspection, the 35 edm:isShownBy statements are used with DOIs that resolve to a webpage; these identifiers should thus be used with edm:isShownAt or edm:hasView.

Table 8. Distribution of DOI usage across properties

Property	Count	%
edm:isShownAt	474,461	86.0%
dc:identifier	77,136	14.0%
edm:isShownBy	35	0.0%
edm:hasView	2	0.0%
total	551,634	

4.6. Statistics on usage of PURL

The PURL is the least used PID Scheme in Europeana. It is mainly used with edm:isShownBy, but still has a 11.1% usage with edm:isShownAt. Some providers also use PURLs in the dc:identifier of the cultural heritage object.

Table 9. Distribution of PURL usage across properties

Property	Count	%
edm:isShownBy	240,026	58.6%
edm:hasView	103,640	25.3%
edm:isShownAt	45,554	11.1%
dc:identifier	20,457	5.0%
total	409,677	

There are five different PURL services in use, with one of them being the widely known purl.org.

Table 10. The PURL domain names in use in Europeana

Domain	Count
purl.ox.ac.uk	333,269
www.purl.org and purl.org	41,018
purl.pt	35,061
purl.sgmf.pt and purl.sgmf.gov.pt	328
purl.access.gpo.gov	1

4.6. Use of multiple PIDs in the same Europeana record

We also investigated cases of records containing more than one distinct PID. These cases are frequent: around 34% of the records that contain PIDs contain more than one. In some cases, a PID is used for the CHO and one or more PIDs are used for web resources, or in other cases, a PID is used for each web resource but none for the CHO. And there are also cases where a CHO simply has more than one PID assigned. It is also relevant to note that it is not uncommon for one resource within a record to be assigned a PID from multiple persistent identifier schemes ,

which is not a bad practice unless these come from the same PID scheme¹⁸. The following table lists the number of cases found, grouped by (pair of) PID scheme(s), and gives some examples of each configuration.

Table 11. Cases with more than one distinct PID on a Europeana record

PID Schemes	Count	Examples
ARK	1,252,133	CHO: 9200365/BibliographicResource_2000081597945 PIDs: - http://gallica.bnf.fr/ark:/12148/bpt6k6228241b - http://gallica.bnf.fr/ark:/12148/bpt6k6228241b/f1.zoom
		CHO: 9200365/BibliographicResource_2000081597803 PIDs: - http://gallica.bnf.fr/ark:/12148/bpt6k884246x - http://gallica.bnf.fr/ark:/12148/bpt6k884246x/f1.zoom
HANDLE	829,878	CHO: 660/ providedCHO https hdl handle net 10622 COLL00467 https hdl handle net 10622 COLL00467 10 PIDs: - https://hdl.handle.net/10622/COLL00467.10?locatt=view:level3 - https://hdl.handle.net/10622/COLL00467-10 (NB: this is in fact an invalid PID; it should be https://hdl.handle.net/10622/COLL00467.10)
		CHO: 659/ providedCHO https hdl handle net 10622 COLL00462 https hdl handle net 10622 COLL00462 11 PIDS: - https://hdl.handle.net/10622/COLL00462.11?locatt=view:level3 - https://hdl.handle.net/10622/COLL00462-11
URN	266,172	CHO: 9200234/download_type_document_docid_356377 PIDS: - urn:isbn:8884532019 - urn:isbn:8884532027
		CHO: 9200234/download_type_document_docid_432175 PIDS: - urn:isbn:9789949321520 - urn:isbn:9789949321988
ARK, DOI	88,691	CHO: 467/10 17026 dans z38 3zt9 PIDs: - https://doi.org/10.17026/dans-z38-3zt9 - https://datavault.dans.knaw.nl/ark:/73189/96e746b6-6654-4ca9-869e-5d24f112eba7/data/images/PAN%2D00025505%2D003%2Ejpg - https://datavault.dans.knaw.nl/ark:/73189/96e746b6-6654-4ca9-869e-5d24f112eba7/data/images/PAN%2D00025505%2D001%2Ejpg

¹⁸ It may also be legitimate in situations like the ARK one, where the use of "passthrough suffixes" qualify an identifier to refer to a different (media) representation of a resource, or a part of it, without really amounting to a duplicate identification.

		<p>CHO: 467/10_17026_dans_z5g_w2uv</p> <p>PIDs:</p> <ul style="list-style-type: none"> - https://doi.org/10.17026/dans-z5g-w2uv - https://datavault.dans.knaw.nl/ark:/73189/e6bac95b-9ac7-4fec-ba68-c332fed9088c/data/images/PAN%2D00080358%2D001%2Ejpg - https://datavault.dans.knaw.nl/ark:/73189/e6bac95b-9ac7-4fec-ba68-c332fed9088c/data/images/PAN%2D00080358%2D003%2Ejpg
DOI, URN	57,426	<p>CHO: 92040/URN_RS_NAE_4ed7e7a5_cb5b_4ca3_857f_03c8515c1f2dcho</p> <p>PIDs:</p> <ul style="list-style-type: none"> - URN:RS:NAE:4ed7e7a5-cb5b-4ca3-857f-03c8515c1f2dcho - http://dx.doi.org/10.2298/GSGD0902003P
		<p>CHO: 92040/URN_RS_NAE_4fdf0fe8_8d69_401b_9fb6_b5b40b6f2f7ccho</p> <p>PIDs:</p> <ul style="list-style-type: none"> - http://dx.doi.org/10.2298/TEM0202019C - URN:RS:NAE:4fdf0fe8-8d69-401b-9fb6-b5b40b6f2f7ccho
PURL	42,546	<p>CHO: 9200143/BibliographicResource_2000069391478</p> <p>PIDs:</p> <ul style="list-style-type: none"> - http://purl.ox.ac.uk/uuid/bc46f132699f4db784e5a905a66741a0 - http://purl.ox.ac.uk/uuid/599514d8509841f5be0a8ca23c589176
		<p>CHO: 9200143/BibliographicResource_2000069361143</p> <p>PIDs:</p> <ul style="list-style-type: none"> - http://purl.ox.ac.uk/uuid/5fdb432ea11a4a339d6fef981a2ad80c - http://purl.ox.ac.uk/uuid/00c76cde251c4db3ba7387c534239c74 - http://purl.ox.ac.uk/uuid/2cbaf993e97f4fd78ce69c3b7a56a42d - http://purl.ox.ac.uk/uuid/0dcc658a6f574ec980ddf80829c8b310 - http://purl.ox.ac.uk/uuid/637d8fa621264757b6a9ce1bbd9ec7b5 - http://purl.ox.ac.uk/uuid/7369b20a4a084eebb7bf911dce1bca59
DOI, HANDLE	1,400	<p>CHO: 2022712/lod_oai_gredos_usal_es_10366_121882_ent0</p> <p>PIDs:</p> <ul style="list-style-type: none"> - http://dx.doi.org/10.1016/j.physletb.2008.02.049 - http://hdl.handle.net/10366/121882
		<p>CHO: 2048425/item_C5COZKU3O4I2RINLSBDNNQOGBS5OPTV</p> <p>PIDs:</p> <ul style="list-style-type: none"> - http://hdl.handle.net/10419/167710 - https://doi.org/10.4054/DemRes.2017.36.55
DOI	432	<p>CHO: 2064501/https_data_ucd_ie_data_ivrla_35250</p> <p>PIDs:</p> <ul style="list-style-type: none"> - http://dx.doi.org/info:fedora/ivrla:34550 - doi:10.7925/drs1.ivrla_35250
		<p>CHO: 2064501/https_data_ucd_ie_data_ivrla_35543</p> <p>PIDs:</p> <ul style="list-style-type: none"> - http://dx.doi.org/info:fedora/ivrla:34550 - doi:10.7925/drs1.ivrla_35543 ;

PURL, URN	99	CHO: 10501/bib_rnod_51333 PIDs: - http://purl.sgmf.pt/COL-MF-0022/1/COL-MF-0022_master/COL-MF-0022_PDF/versaointegral.pdf - http://purl.sgmf.pt/COL-MF-0022/1/ - URN:ISBN:972-9244-23-5
		CHO: 10501/bib_rnod_24362 PIDs: - https://purl.pt/6359 - URN:NBN:pt-25525/89
HANDLE, URN	31	CHO: 10501/bib_rnod_238630 PIDs: - http://hdl.handle.net/10400.2/6727 - URN:ISBN:978-972-674-551-8
		CHO: 2048425/item_FKAKWZI2JKJKUBWUWU5NCLDHNOS7P6S PIDs: - http://hdl.handle.net/10419/191936 - urn:nbn:de:0168-ssaoar-65698-9
ARK, PURL	2	CHO: 9200143/BibliographicResource_2000069296189 PIDs: - http://gallica.bnf.fr/ark:/12148/bpt6k279983.notice - http://purl.ox.ac.uk/uuid/05fbf66d83e24f5d96922b5bab8ca435 - http://purl.ox.ac.uk/uuid/db139ac33b9441deb8d9881788b9e818 - http://purl.ox.ac.uk/uuid/75d5d7e1272845ef953100aba2541e3e - http://purl.ox.ac.uk/uuid/56333efcf996465c9df61b3cf95a70fb - http://purl.ox.ac.uk/uuid/3788bcd7f88444c28dad658997a3c676 - ... (more 20 PURLs not shown)
		CHO: 9200143/BibliographicResource_2000069296544 PIDs: - http://purl.ox.ac.uk/uuid/13e943875f83425092dbf4d2bc0b11d3 - http://gallica.bnf.fr/ark:/12148/bpt6k1249466
Total	2,538,396	

5. Uniqueness of the persistent identifiers in Europeana across records and datasets

Uniqueness of identifiers is strongly desirable for Europeana objects and their representation, which are expected to be distinct. With this in mind, we checked for the existence of cases where more than one record shares the same persistent identifier, after normalising the PIDs that can be expressed in different forms (HTTP(S) URI or compact URI) and the PIDs that can be attached to different resolvers.

The Europeana dataset contains 385,796 cases where more than one record share the same PID. This occurs across all PID Schemes, but is more frequent in ARKs (see Table 12).

The large majority of shared PIDs happens in only two records. Many of these cases are actually valid, i.e., when the same PID appears in records about the same CHO aggregated via two

different data aggregation routes. In fact these cases may help us to de-duplicate objects across datasets. There are cases where a PID is shared by many more records, in some cases more than 10 (see Table 13). The most extreme cases happen in periodic publications and collections, when the PID of the periodic publication/collection is included in all the records of the individual issues, resulting in hundreds of records with the same PID. Other cases appear to be data quality problems. Some examples are shown in Table 14.

Table 12. Number of cases, per PID scheme, with more than one record with the same PID

PID Scheme	Count
ARK	222,493
HANDLE	143,646
URN	11,905
DOI	7,711
PURL	41
Total	385,796

Table 13. Number of occurrences of the same PID in different Europeana records

Number of records with the same PID	Number of cases
2	320,301
3	31,909
4	20,458
5	8,565
6	1,505
7	1,360
8	584
9	553
10	70
more than 10	491
Total	385,796

Table 14. Examples of a same PID appearing in different Europeana records

PID	Example records	Description
https://hdl.handle.net/11653/per60201	2021624/eazines_IJV_602014_I_AV_00602014_001_05_pdf 2021624/eazines_IJV_602014_I_AV_00602014_001_03_pdf	The PID appears in edm:isShownAt on 956 records. It identifies a periodic publication.
https://doi.org/10.5962/bhl.title.21280	08711/https_www_biodiversitylibrary_org_item_65936 08711/https_www_biodiversitylibrary_org_item_66080	The PID appears in edm:isShownAt on 195 records. It identifies a periodic publication.
http://hdl.handle.net/10934/RM0001.HARCHIEF.29529	90402/HA_0029529 90402/HA_0030458	The PID appears in dc:identifier on 44 records. It identifies a collection.
http://gallica.bnf.fr/ark:/12148/bpt6k4132888	9200408/BibliographicResource_3000115190290 9200365/BibliographicResource_1000056001438	The PID appears in dc:identifier on 2 records. It is not clear what it identifies, since it resolves to a page that shows only a title and author.
http://www.arkivverket.no/URN:NBN:no-a1450-tl20071115630467.jpg	2058626/loccloud_preg_d94599a8_bc7a_40f4_896f_997b9b56b432 2058626/loccloud_preg_74ce7944_375f_49d0_ad13_23ef69b5a94f	The PID appears in edm:isShownBy on 2 records. Either these records are duplicates or there is an error in the edm:isShownBy of one of the records.
http://purl.pt/471	9200364/BibliographicResource_3000098464712 10501/bib_rnod_39846	The PID appears in edm:isShownAt on 2 records. It appears to be a CHO that was aggregated twice, via different aggregators.

6. Data quality issues

Although it was not an initial objective of our study to analyse data quality issues related with PIDs, during our work we came across some cases which are relevant to be reported.

Many of the cases presented earlier in Section 5 are data quality problems where a PID is attributed to a CHO when it actually identifies a collection or a serial publication.

We have also detected that some providers unintentionally use PID schemes in their identifiers. Some HANDLE prefixes are not valid, e.g., '123456789'. It appears that some digital library systems have built-in support for automatically outputting HANDLE identifiers in the metadata. The data providers may have neglected to deactivate or configure this functionality properly. A similar case was found for URNs. Some URNs have a namespace identifier that refers to a system that structures its identifiers as URNs (for example, the namespace 'repor.ist.utl.pt'). When

deploying this system, the data provider did not configure the namespace identifier to its own case.

In the same vein, we found a few trivial typos, such as repeated prefixes - 'urn:urn:nbn:...' (e.g., 'URN:URN:NBN:SI:doc-DU27Q92N' instead of 'URN:NBN:SI:doc-DU27Q92N' , a problem that occurs for 333,284 PIDs!) or using an '-' (hyphen) instead of a '.' (dot) (e.g., 'hdl:10622/COLL00467-10' instead of 'hdl:10622/COLL00467.10').

Finally, we have detected uses of PIDs that do not fit EDM semantics. Especially, some PIDs in `edm:isShownBy` resolve to a web page, while this property should point directly to media files such as images or videos. This issue is not specific to PID values and has been detected earlier by the Europeana Data Quality Committee¹⁹.

7. Conclusion

This inventory shows that PIDs are already being delivered to Europeana to a small but significant extent. 13% of the records in Europeana contain a possible PID. This percentage goes up to 14.1% if records in tier 0 are excluded from the analysis.

Three of the selected PID Schemes in use are supposed to be supported by reliable persistent identifier policies and resolution systems - ARK, HANDLE and DOI - while it is unclear which URNs and PURLs have an underlying persistence policy. ARK and HANDLE are the most frequently used. ARK is the most used, but ARK appears to be in use only by a few data providers (only 4 name assigning authorities are present), while HANDLE is used by more data providers (72 different prefixes are in use). DOI is not frequently used by Europeana data providers.

URNs are frequently encountered in Europeana, particularly URNs using the National Bibliography Number namespace. But it is not certain if all countries using them have reliable resolution services. URNs may have other problems of data quality, because not all URN namespaces in use are registered in IANA, and others appear to be invalid.

The PURL scheme is the least used in Europeana. Additionally, it is not guaranteed that the institutions assigning PURLs have a reliable persistence policy in place, therefore the use of the PURL scheme may be unreliable.

It is important to highlight these concerns regarding the PIDs in use in Europeana. Not all of the PID schemes in use actually ensure persistence of identification, or they may not have an associated resolution service. We spotted issues with HANDLES, for example. In addition, we have also identified data quality problems on some PIDs, such as invalid namespaces, more than one record having the same PID, or wrong use of (EDM) metadata elements.

Overall, this study indicates that the usage of PIDs still needs to increase and mature, so that they can fully benefit the common European data space for cultural heritage that is being built²⁰. It also brings concrete elements that can be used to refine data governance policies and data quality reporting that are crucial to such data spaces, especially regarding the adoption of the FAIR principles (PIDs without persistence policy or with non-registered prefixes are less commendable). Some results of this inventory have fed into the ongoing work on assessing and enhancing persistence of identifiers in Europeana. For example, the main Key Performance Indicator for the data space has been changed from "Records with a unique permanent identifier

¹⁹ <https://pro.europeana.eu/project/data-quality-committee#problem-patterns>

²⁰ <https://pro.europeana.eu/post/help-us-to-make-cultural-heritage-data-more-persistent>

in the data repository" to "The number of items published on the Europeana website that contain identifiers using a persistent identifier scheme with an underlying persistence policy.". Also, some of our efforts have already contributed to the work of the Europeana Data Quality Committee.

In future work we would like to investigate methods for determining which PIDs are reliably persistent over time, e.g. exploiting IANA registration status and perhaps reusing past studies on lack of persistence (link rot). We also want to better detect data quality problems related to PIDs, especially (1) invalid non-unique PIDs and (2) usage of EDM elements with PIDs that refer to types of digital resources not expected in these elements. Finally, recommendations for data providers are needed so that these issues are better addressed across the common European data space for cultural heritage.

References

- [1] Gordon McKenna, G., Fokke, C.: D 2.2 State of the art report on persistent identifier standards and management tools. Project Linked Heritage. (2013). Retrieved from <https://www.linkedheritage.eu/getFile.php?id=556>
- [2] Koster, L.: Persistent identifiers for heritage objects. Code4Lib Journal, issue 47. (2020).
- [3] Svensson, L. G.: Universal Access to Cultural Heritage Material: The Europeana Resolution Discovery Service for Persistent Identifiers. International Conference on Dublin Core and Metadata Applications, 184–185. (2010). Retrieved from <https://dcpapers.dublincore.org/pubs/article/view/1008>
- [4] Klein, M., Balakireva, L.: An extended analysis of the persistence of persistent identifiers of the scholarly web. *Int J Digit Libr* 23, 5–17 (2022). <https://doi.org/10.1007/s00799-021-00315-w>
- [5] Kotarski, R., Kirby, J., Madden, F., Mitchell, L., Padfield, J., Page, R., Palmer, R., Woodburn, M.: PIDs as IRO Infrastructure-Early Findings. (2020). <https://doi.org/10.23636/1214>.
- [6] Kotarski, R., Madden, F.: Persistent Identifiers as IRO Infrastructure: Survey Data. British Library. (2020). [Dataset]. <https://doi.org/10.23636/1210>

Annex A - Europeana data providers using PIDs

This annex lists the 512 Europeana data providers that have delivered at least one record containing a PID.

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Archdiocese of Dublin	60	0
Μεγάλη Μουσική Βιβλιοθήκη της Ελλάδας «Λίλιαν Βουδούρη» - Σύλλογος Οι Φίλοι της Μουσικής	13395	11027
Leibniz Universität Hannover	37	0
Ljutomer Public Library	2	0
National and University Library of Strasbourg	21320	0
University of Toronto	41	13
Josip Juraj Strossmayer University of Osijek, Faculty of Medicine	59	0
Rota do Românico	33	11
ROM Technik	18	0
Statistical Office of the Republic of Slovenia	11	0
UCD College of Human Sciences	9	0
National and University Library of Slovenia	336435	428
Ίδρυμα Μουσείου Νίκου Καζαντζάκη	2551	0
Transport Infrastructure Ireland	2578	4
National Museum of World Cultures Foundation	17812	0
Sterling and Francine Clark Art Institute Library	6	0
Slovenian music information centre	1598	0
Singapore Botanic Gardens, National Parks Board Singapore	242	52
Secretaria Geral do Ministério das Finanças	160	0
School of Theology, Boston University	2	0
State Library of North Carolina	25	27
State Library of Massachusetts	1	2
State Botanical Collection, Royal Botanic Gardens Victoria	59	7
Scott - York University Libraries	9	0
Joseph Palmer Knapp Library, UNC School of Government	1	0
Joint Committee on Taxation, US Congress	1	0
Εθνικό Κέντρο Τεκμηρίωσης & Ηλεκτρονικού Περιεχομένου (ΕΚΤ)	551	0
National Library of Estonia	269	99986
John Carter Brown Library	579	10
John Adams Library at the Boston Public Library	730	268

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Universidade do Porto	5114	0
National Library of Wales	174409	671
Institute for Blind and Partially Sighted Children Ljubljana	10	0
Trinity College Dublin	12541	8
Library of Neapolis University Paphos	3118	0
Academy of Natural Sciences of Drexel University, Library and Archives	71	146
Abraham Lincoln Presidential Library	1	1
Agriculture and Agri-Food Canada	331	17
Seminário Livre de História das Ideias	6	15
Bayerische Staatsbibliothek	1241359	127
UCD School of Irish, Celtic Studies, Irish Folklore and Linguistics	191	0
Alberta Legislature Library	23	32
UCLA Library Preservation Department	8	0
Erasmus Smith Schools Archive	64	0
Universität Bremen, Forschungsstelle für Bienenkunde	1	0
Culture and Convention Centre Cankarjev Dom	1	1
Allen County Public Library Genealogy Center	39	152
Irish Jesuit Archives	20	0
Waterford City and County Council	7	0
Amsterdam Museum	3972	0
Amsab-Institute of Social History	67452	0
American Museum of Natural History Library	1505	3230
American Entomological Society	1	1
Institut für Film und Bild in Wissenschaft und Unterricht, Abteilung Hochschule und Forschung	1	0
Petro Jacyk Central & East European Resource Centre, University of Toronto	9	0
University of Galway	50	0
Ταμείο Αρχαιολογικών Πόρων και Απαλλοτριώσεων	10055	0
Archaeology Data Service	57733	36765
International Institute of Social History	165110	240
University of Zagreb, Faculty of Metallurgy	2	0
University of Split, Faculty of Civil Engineering, Architecture and Geodesy	5	0
Velenje Library	74	73

Data provider name	Nr. of records with a PID	Nr- of records without a PID
National Irish Visual Arts Library	20	0
Chesapeake Bay Foundation	5	14
Dr Franc Sušnik's Central Library of Carinthia	280	0
Chicago Botanic Garden, Lenhardt Library	133	48
University of Rijeka, Faculty of Tourism and Hospitality Management	20	0
Ίδρυμα Κωνσταντίνος Κ. Μητσοτάκης	1129	0
Museu Nacional de Arte Antiga	1	264
Museum and Institute of Zoology, Polish Academy of Sciences	1	0
Museum of Design of Barcelona	834	0
Postal Cabinet Museum of Barcelona (Marull's Collection)	49300	0
Museum of Cultures of the World	550	0
Sweet Briar College	9	0
University College Dublin	558	3843
National Library of Serbia	48477	5751
The Blavatnik Archive	1070	0
University of Rijeka, Faculty of Law	5	0
Josip Juraj Strossmayer University of Osijek. FACULTY OF FOOD TECHNOLOGY. Department of Process Engineering. Sub-department of Unit Operations.	2	0
University of Split, School of Medicine	1	0
Χριστιανική Αδελφότητα Νέων Θεσσαλονίκης (ΧΑΝΘ)	5588	0
Josip Vošnjak Library, Slovenska Bistrica	381	105
Ivan Potrč Library Ptuj	8695	1891
Catholic Faculty of Theology, University of Zagreb	13	0
Miran Jarc Library Novo Mesto	87	1903
Národní lékařská knihovna	755	28177
University of Split, Faculty of Catholic Theology	2	0
Juraj Dobrila University of Pula	2	1228
Archiv der Sozialen Demokratie	133032	0
University of Rijeka, Faculty of Informatics and Digital Technologies	2	0
Architekturmuseum der Technischen Universität Berlin	156055	83
Archives of Ontario	1	0
University of Zagreb. Catholic Faculty of Theology. Department of Religious Pedagogy and Catechetics.	1	0
University College London Library Services	51	0
Royal Botanic Gardens Kew	37	597099

Data provider name	Nr. of records with a PID	Nr- of records without a PID
ASC - York University Libraries	59	41
Assembleia da República - Arquivo Histórico Parlamentar	1	1436
Atria, Institute on Gender Equality and Women's History	8601	0
Dublin City Library and Archive	22270	7
Australian Museum	2	12
Australian Institute of Marine Science	10	0
University of Navarra	7744	0
UNAM Institute of Biology	170	10
Badische Landesbibliothek Karlsruhe	9121	13
Internet Archive	449	176901
Bauhaus-Universität Weimar. Universitätsbibliothek	20656	638
School of Dental Medicine, University of Zagreb	54	0
Stiftung Universität Hildesheim	12	0
Helmholtz-Zentrum Potsdam Deutsches GeoForschungsZentrum	10	0
Επιτροπή Ποντιακών Μελετών	385	0
Institut für den Wissenschaftlichen Film	160	0
Eidgenössische Technische Hochschule Zürich	1	0
Zentrum für interdisziplinäre Forschung	7	0
Lindau Nobel Laureate Meetings	201	0
Arquivo Regional e Biblioteca Pública da Madeira	2	434
Laserinstitut Hochschule Mittweida	16	0
Hochschule Ravensburg-Weingarten	15	0
Carl von Ossietzky Universität Oldenburg	17	0
Weierstraß-Institut für Angewandte Analysis und Stochastik	17	0
Institut für sozial-ökologische Forschung	13	0
Ίδρυμα Βαρώνου Μιχαήλ Τσούτσα	1036	0
Noranda Earth Sciences Library	180	93
Tribunal de contas	2	93
Biblioteca das Artes da Madeira	118	352
Gesellschaft für Informatik	60	0
International Union of Crystallography	22	0
Forschungszentrum Küste	18	0
IMAGINARY gGmbH	26	0
Math.space - Verein für Mathematik als kulturelle Errungenschaft	25	0
Institut des Hautes Études Scientifiques	227	0

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Ruhr-Universität Bochum	194	0
Field Museum of Natural History	551	16238
Library and Archives Canada - Bibliothèque et Archives Canada	2	0
Library of Congress	6519	142
Αρχαιολογικό Μουσείο Χανίων	609	0
Faculty of Science, University of Zagreb	246	0
University College Dublin's School of History and Archives	612	0
University of Zagreb. Catholic Faculty of Theology. Department of Fundamental Theology.	1	0
Biblioteca do Exército Português	743	15191
University College Cork	462	62
National Great Rivers Research and Education Center	19	1
Paul Jerabek, Julian Hegemann, Andreas Authmann	99	0
Rijksmuseum	341407	300
Robert Dale Kloppenburg	4	37
Knihovna Akademie věd České republiky	3903	60729
National Archives of the Netherlands	9671	729
Bodleian Libraries, University of Oxford	53598	29945
University of Cyprus	4507	0
University of Southampton	229	2
Faculty of Economics and Business, University of Zagreb	14	0
Complutense University of Madrid	107342	18255
Národné osvetové centrum (nocka / V-klub)	105	0
NCSU Libraries	1889	174
New College of California	2	0
Natural History Museum in Paris	1	504423
New York Botanical Garden	3921	6290
Nature Kenya, East Africa Natural History Society	1	10
NC History of Health Digital Collection	1	1
Naval Postgraduate School	473	3
Natural History Museum Library, London	968	6090
PHOTOCONSORTIUM	31755	121521
Joint Library of Earth Sciences, Institute of Marine Sciences and Limnology, UNAM	103	2
Museum and Galleries of Ljubljana	14	622

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Meise Botanic Garden	204	1384055
Museums Victoria	134	121
Nipissing University	1	0
Josip Juraj Strossmayer University of Osijek, Faculty of Civil Engineering	15	0
NIH Library	56	18
Commonwealth of Australia (Geoscience Australia)	4	0
Comissão para a Cidadania e a Igualdade de Género	27	299
Cork City Libraries	11	0
Tulsk History Society	11	0
Cornell University Library	4721	1720
Polytechnic University of Valencia	33198	0
Claire T. Carney Library, UMass Dartmouth	7	0
Occitan Media Library	355	1607
National Commission for the Conservation of Biodiversity	27	176
Columbia University Libraries	470	53
Irish Qualitative Data Archive	725	1
Deutsche Nationalbibliothek	299212	4842
Lancaster County Historical Society	10	0
KNAW-DANS	48118	0
Emory University, Woodruff Health Sciences Center Library	1	0
Emory University, Robert W. Woodruff Library	8	0
Emory University, Pitts Theology Library	10	0
Environmental Services & Library Systems	2	0
Harold B. Lee Library	252	25
Harvard University Botany Libraries	985	1333
Frost - York University Libraries	1	0
The Hunt Museum	152	326
Emory University, Manuscript, Archives and Rare Book Library	7	0
Harvard University, Museum of Comparative Zoology, Ernst Mayr Library	3395	7395
Smithsonian Institution Archives	2997	342
GESIS - Leibniz-Institut für Sozialwissenschaften. Bibliothek Köln	54445	500
Faculty of Kinesiology, University of Zagreb	31	0
Tourism Society Vurberk	2	0
Hemeroteca Municipal de Lisboa	89	372

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Faculty of Croatian Studies, University of Zagreb	45	0
The Government of Catalonia	34885	0
Εν Αθήναις Αρχαιολογική Εταιρεία	35502	0
Fordham University Libraries	3	0
The British Library	64647	262194
Universitätsbibliothek Johann Christian Senckenberg Frankfurt am Main	3849	32669
Forschungsbibliothek Gotha	9556	165
Smithsonian Libraries	11332	16637
Francis A. Countway Library of Medicine	217	12
Leiden University Libraries	9087	452
Ινστιτούτο Ιστορικών Ερευνών	19083	4873
Ελληνική Ολυμπιακή Επιτροπή	1460	0
Central Library Srečko Vilhar Koper	57	314
Celje Central Library	1851	4630
Κέντρον Ερεύνης της Ελληνικής Λαογραφίας της Ακαδημίας Αθηνών	1217	0
Severočeská vědecká knihovna	146	236
National Library of Medicine	77	4
Nasjonalbiblioteket	255	157
Cork LGBT Archive	167	1
Kočevje Regional Museum	93	0
Open Society Archives at Central European University	116116	0
National Library Board, Singapore	96	2
Repositori Fons Especials UdL	11398	1
National and University Library in Zagreb	1	21301
University of Rijeka, Faculty of Medicine	92	0
Pennsylvania Horticultural Society	8	15
Penn State University	251	429
Universidade Aberta	1914	35
E-codices - Virtual Manuscript Library of Switzerland	1554	0
Duke University Libraries	156	2
Faculty of Agriculture, University of Zagreb	22	0
Lower Silesian Digital Library	4497	54367
Faculty of Forestry and Wood Technology, University of Zagreb	28	0
European Fashion Heritage Association	18487	445130

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Faculty of Mining, Geology and Petroleum Engineering, University of Zagreb	24	0
Dorothy H. Hoover Library, Ontario College of Art & Design	8	0
Faculty of Pharmacy and Biochemistry, University of Zagreb	54	0
Ontario Ministry of Natural Resources Library	2	27
Open Access Publishing in European Networks Foundation	3308	4912
North Carolina State Museum of Natural Sciences	36	13
Northeastern University, Snell Library	45	12
University of Rijeka, Faculty of Maritime Studies	40	0
University of Toronto Mississauga	12	1
Josip Juraj Strossmayer University of Osijek, Faculty of Economics	38	0
Institute of Botany, Chinese Academy of Sciences	9	942
Institut für deutsche Sprache - Bibliothek	5532	4361
Europeana Local Österreich	1	0
History of Modern Biomedicine Research Group	1	0
EuropeanaTravel	254	0
Κέντρο Διάδοσης Επιστημών και Μουσείο Τεχνολογίας	2729	0
SciFox	46	0
Museum für Naturkunde Berlin	2	44178
Autonomous University of Barcelona	14	21341
Documentary Repository of the University of Salamanca	94758	0
The European Film Gateway	182	299730
Hochschul- und Landesbibliothek Fulda	66794	754
France Bevk Public Library Nova Gorica	3539	1275
Göteborgs stadsmuseum	2	40839
Goldfarb Library, Brandeis University	112	0
The Lepidopterists' Society	4	0
The Language Archive	18392	0
The Johns Hopkins University Sheridan Libraries	37	0
Εθνική Βιβλιοθήκη της Ελλάδος	101	0
The European Library	36	664
Technische Informationsbibliothek	32	0
University of Rijeka, Faculty of Humanities and Social Sciences	39	0
Digital Repository of Ireland	97	5
Aggtelek National Park Directorate	1	2034

Data provider name	Nr. of records with a PID	Nr- of records without a PID
The Oireachtas Library & Research Service	1567	1568
The Newberry Library	119	3
Josip Juraj Strossmayer University of Osijek	4	0
Joe Lee Community Based Films	8	0
Goethe University Frankfurt	10295	3
University of Split, Faculty of Humanities and Social Sciences in Split	8	0
University of Split, Faculty of Economics, Business and Tourism	4	0
University of Rijeka, Department of Biotechnology	6	0
Dublin City University Library	12	0
Raptor Research Foundation	1	203
Ίδρυμα Μιχάλης Κακογιάννης	2424	0
House for the Social Sciences and Humanities of Dijon	2116	0
CARARE	40920	0
Faculty of Transport and Traffic Sciences, University of Zagreb	9	0
Faculty of Education and Rehabilitation Sciences, University of Zagreb	8	0
University of Dubrovnik	2	0
Sustainable History Monograph Pilot (SHMP)	1	0
Sancho el Sabio Fundazioa	6099	0
Naturalis Biodiversity Center	802	4533370
Ολυμπιακό Μουσείο	1842	0
Euskal Herriko Unibertsitatea	566	0
University of Split, Faculty of Kinesiology	42	0
Εταιρεία Τοπικής Ανάπτυξης Λέσβου Α.Ε	1738	0
National Library of France	2999306	4
Digital Library of the Formal Linguistics Department at the University of Warsaw	1	101
Digital Library of the Silesian University of Technology	41	32109
Library of the Wroclaw University	80	111238
Biblioteca Pública Municipal do Porto	34	687
John Gilmore (archive.org)	2	0
Ontario Ministry of the Environment	78	14
London School of Hygiene & Tropical Medicine Library & Archives Service	82	0
Lucian Blaga Central University Library, Cluj-Napoca, Romania	1	58373

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Lincoln Financial Collection	30	1
Bibliothek der Friedrich-Ebert-Stiftung	201518	0
Library of the Marine Corps	1	0
Universitäts- und Landesbibliothek Darmstadt	6521	1
Unknown	7	61
Universitätsbibliothek Kiel	6425	164
Universitätsbibliothek Heidelberg	34149	107640
Universitätsbibliothek der Humboldt-Universität zu Berlin	675	2
Universitätsbibliothek 'Georgius Agricola' der TU Bergakademie Freiberg	101	2052
Universitäts- und Landesbibliothek Tirol	751	5
Irish Architectural Archive	30	0
Universidade de Coimbra	2106	79
University of Oviedo	3268	0
Zentral- und Landesbibliothek Berlin	6870	32522
Milutin Bojić Library	72	0
UNEP-WCMC, Cambridge	128	234
Autonomous University of Nuevo León	22	0
Universal Digital Library	1	0
Univ. of Mass Medical School, Lamar Soutter Library	26	0
United States Geological Survey Libraries Program	75	40
UMass Amherst Libraries	422	1450
Galileo Museum	9219	2
Digital Library Memory of Madrid	197832	0
Biblioteca Municipal de Montalegre	1	2
National Library of Scotland	14	17189
University of Toronto - Royal Ontario Museum Library & Archives	138	3
The University of Western Ontario, Western Archives	20	0
University of Pretoria, Library Services	15	0
University of Tartu	17532	0
University of Toronto - OISE Library	37	5
University of Toronto - Regis College Library	2	0
University of Toronto - Robarts Library	1348	442
Wellcome Library	769	20
University of Toronto - St. Michael's College (John M. Kelly Library)	127	20

Data provider name	Nr. of records with a PID	Nr- of records without a PID
University of Toronto - Thomas Fisher Rare Book Library	1681	176
University of Pennsylvania Libraries	113	21
Josip Juraj Strossmayer University of Osijek, Faculty of Humanities and Social Sciences	41	0
University of Pittsburgh Library System	196	34
University of Rijeka, Faculty of Engineering	36	0
Työväen Arkisto	27821	0
U.S. Government Printing Office	4	1
U.S. Dept. of the Treasury, Treasury Library	1	0
U.S. Department of Agriculture	52992	21576
Tufts University	2	0
Trade Union Congress Library	3653	0
Österreichische Akademie der Wissenschaften	5	434
Trinity College in the University of Toronto	8	0
Tilburg University	195	0
University of Toronto - Victoria University (Emmanuel College Library)	2	0
University Library Maribor	2125	465
Uppsala Universitet	34	109976
Tisch Library	27	0
Prelinger Library	181	97
New York Public Library	99	1
Repositorio Institucional de la Universidad de Huelva	1143	0
Regional Archive Nijmegen	35	47889
Public Library Bor	57	0
Cherokee Garden Library	8	36
Auckland War Memorial Museum Tāmaki Paenga Hira	10	33
Harvey Cushing/John Hay Whitney Medical Library	36	0
Delaware Museum of Nature and Science	32	0
Regional Archives Tilburg	6569	0
San Diego Zoo Wildlife Alliance Library	30	0
Universität Potsdam	27	0
Φεστιβάλ Κινηματογράφου Θεσσαλονίκης	702	0
Queen's University, W.D. Jordan Special Collections & Music Library	70	0
Λαογραφικό και Εθνολογικό Μουσείο Μακεδονίας-Θράκης	11722	0

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Princeton Theological Seminary Library	88	4
Υπηρεσία Συντήρησης Μνημείων Ακρόπολης	3989	0
Κρατική Ορχήστρα Θεσσαλονίκης	234	0
Hopkins Marine Station of Stanford University	6	0
Entomological Society of Latvia	3	0
Natural History Museum of Los Angeles County	1	2
Kranj City Library	1603	569
Ljubljana City Library	1301	2933
Frederic Marès Museum	6159	38
Sonoma County Library	4	0
Texas Ornithological Society	7	15
Dr. José María Luis Mora Research Institute	11	5
Jewish Heritage Network	952	69477
The New York Academy of Medicine	11	0
California Academy of Sciences	226	1674
Oliver Wendell Holmes Library	10	0
The Modern Records Centre, University of Warwick	21279	0
MBLWHOI Library	5528	7067
North York Central Library	11	0
Canadian Museum of Nature Library	12	260
Canadiana	2691	25
Cambridge University Library	242	22
Maryland Ornithological Society	1	250
Josip Juraj Strossmayer University of Osijek, Faculty of Food Technology	44	0
Κοινοφελής Επιχείρηση Δήμου Μάνδρας	35	0
Morris Library, Southern Illinois University Carbondale	1	0
Montana State Library	1338	786
Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology	15	0
Booth Library, Eastern Illinois University	1	0
Minerva, Institutional Repository of the University of Santiago de Compostela	5144	0
Museum of Vertebrate Zoology, University of California, Berkeley	2	0
Oak Spring Garden Foundation	2	0
MIT Libraries	2	0

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Missouri Botanical Garden	1705	5305
State Library of New South Wales	1	0
University of California	1	0
Bitoteko, Hispanic Esperanto Federation Library	5188	0
Slovenské technické múzeum	2285	0
Municipal Library of Chambéry	1399	0
Pontifical Institute of Mediaeval Studies	9	0
Εθνικό Κέντρο Τεκμηρίωσης (ΕΚΤ)	119	0
National Institute for Art History Library	18704	0
Brock University	3	48
Brown University Library	20	0
Greek Aggregator SearchCulture.gr	1693	2426
Bureau of Land Management	834	398
Bundesanstalt für Wasserbau	65	17827
Arkivverket	2995029	158900
Biblioteca Nacional de Portugal	24832	608
Boston College Libraries	94	5
Boston University Libraries	4	0
Brigham Young University-Idaho	7	2
Brigham Young University Hawaii, Joseph F. Smith Library	3	0
Brandeis University Libraries	19	0
University of Ottawa	54	0
University of North Carolina at Chapel Hill	16	3
University of New Hampshire Library	6	557
University of Michigan	173	131
University of Maryland	23	2
Williams College Libraries	7	0
Wisconsin Entomological Society	3	80
University of Glasgow Library	240	5
University of Florida, George A. Smathers Libraries	571	212
University of Connecticut Libraries	381	74
University of California, Davis Libraries	270	0
University of California Libraries	11156	115
University of British Columbia Library	379	5
University of Bristol	18	0

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Wellesley College Library	360	6
West Virginia University Libraries	57	3
University of Illinois	3	0
University of Leeds Library	20	0
Webster Family Library of Veterinary Medicine	1196	1
University Library, University of North Carolina at Chapel Hill	64	31
University Library, University of Illinois Urbana Champaign	4001	2248
University library Svetozar Markovic, Belgrade	911	1357
Herzog August Bibliothek	108	65670
Universitätsbibliothek Würzburg	3611	0
University of Alberta Libraries	1613	896
e-archivo Universidad Carlos III de Madrid	175	0
Εφορεία Αρχαιοτήτων Ευβοίας	2554	0
Εφορεία Αρχαιοτήτων Μεσσηνίας	321	0
Εφορεία Αρχαιοτήτων Φωκίδος	149	0
Ίδρυμα Γεωργίου Ζογγολόπουλου	997	0
Δήμος Προσοτσάνης	34	0
Δήμος Χαϊδαρίου	335	0
Ελληνικό Κέντρο Βιοτόπων – Υγροτόπων	1881	0
Εφορεία Αρχαιοτήτων Αργολίδας	44	0
Royal Ontario Museum	86	8
Regional and Study Library Murska Sobota	885	836
South African Institute for Aquatic Biodiversity	6	16
Svenska Litteratursällskapet i Finland, Finland	3258	27888
South African National Biodiversity Institute	26	2
Royal Irish Academy	614	0
Royal College of Physicians, London	194	1
Royal College of Surgeons of England	8	0
Royal College of Physicians in Edinburgh	465	0
Royal Botanic Garden Edinburgh	1	484165
Roman Catholic Archdiocese of Toronto	1	0
Philadelphia Museum of Art, Library	1	0
School of Government Library, University of North Carolina at Chapel Hill	1	0
Ευρωπαϊκό Πολιτιστικό Κέντρο Δελφών	530	0

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Arts and Culture in Education Research Repository	62	1
San Francisco Public Library	35	76
Repository of Teaching and Research Objects of the University of Cádiz	3127	0
University of Rijeka, Faculty of Economics and Business	62	0
Saint Mary's College of California	11	0
Saint Louis Public Library	1	0
Kraljeva knjižnica, Danska nacionalna knjižnica in univerzitetna knjižnica v Kopenhagenu	5	0
Ryerson University Library Special Collections	1	0
Rutgers University Libraries	5	0
Faculty of Organization and Informatics, University of Zagreb	41	0
Αμερικανική Σχολή Κλασικών Σπουδών	4246	0
Κέντρο Ανάπτυξης και Επιχειρηματικότητας	4173	0
BHL SciELO	147	362
Silesian Digital Library	85	267471
Museum of Texas Tech University	207	96
University of Zadar	21	0
King's College London, Foyle Special Collections Library (archive.org)	197	9
Digital Repository Research Institutes	63	36459
Yale Peabody Museum of Natural History	200	0
Research Library, The Getty Research Institute	783	65
University of Zagreb. Catholic Faculty of Theology. Department of the History of Christian Literature and Christian Teaching.	1	0
Monaghan County Council	20	1
Kansalliskirjasto	4859	540017
Andrew Pollock	20	0
Σύλλογος Φίλων Ιδρυμάτων Στέφανου Κότσιανου	5418	0
E_Buah Biblioteca Digital de la Universidad de Alcalá	2176	0
Deutsches Exilarchiv	5343	32
Centro de documentação 25 de Abril da UC	206	160
Brooklyn Botanic Garden	4	78
San Diego Natural History Museum Research Library	31	49
University of Texas Libraries	5	45
Cheng Yu Tung East Asian Library - University of Toronto	3	0

Data provider name	Nr. of records with a PID	Nr- of records without a PID
Phillips Library at The Peabody Essex Museum	2	59
Lloyd Library and Museum	51	1
Historical Image Archive for Migrants	108	4
Cantonal Museum of Zoology, Lausanne	127	6
Leibniz-Informationszentrum Technik und Naturwissenschaften	550	0
University of Zagreb. Catholic Faculty of Theology. Department of the Holy Scripture of the New Testament.	1	0
Ιστορική & Εθνολογική Εταιρεία της Ελλάδος	4535	0
Επιμελητήριο Εικαστικών Τεχνών Ελλάδος	6570	0
Gerstein Science Information Centre	3314	1883
Ίδρυμα Ωνάση	7980	0
Josip Juraj Strossmayer University of Osijek. FACULTY OF FOOD TECHNOLOGY. Department of Process Engineering. Sub-department of Modelling, Optimisation and Automatisatation.	2	0
Staatsbibliothek Bamberg	2296	0
Ίδρυμα Κωνσταντίνου Σημίτη	15728	0
Zuiderzee Museum	19639	910