

ECDL 2009 Tutorial

Aggregation and reuse of digital objects' metadata from distributed digital libraries

Prepared by:

PSNC Digital Libraries Team

(<http://dl.psnc.pl/>)

Speaker:

Marcin Werla

(mwerla@man.poznan.pl)

Part 2: The most important aspects of... the metadata aggregation and reuse

ECDL 2009 Tutorial: Aggregation and reuse of digital objects' metadata
from distributed digital libraries

Metadata aggregation and reuse

- Foundation for building new, advanced network services
- Gathering metadata records from diverse, distributed digital libraries can simplify metadata reuse scenarios
 - Metadata unification and cleaning
 - Forwarding metadata to other services
 - e.g. Europeana, Driver

Metadata aggregation and reuse

- Metadata exchange can be realized using different means one of them is OAI-PMH

Metadata exchange using OAI-PMH

- OAI-PMH – Open Archives Protocol for Metadata Harvesting
 - Defines two actors
 - data providers (repositories)
 - service providers (harvesters)
 - Service providers provide services based on the metadata harvested from data providers

Metadata exchange using OAI-PMH

- OAI-PMH – Open Archives Protocol for Metadata Harvesting
 - XML based
 - Built on top of HTTP protocol
 - Widely used
 - Over 1 000 repositories worldwide containing several millions of records
 - More info: <http://www.openarchives.org/>

Metadata exchange using OAI-PMH

- OAI-PMH Repository data structure
 - Repository contains items
 - Item may be described in multiple metadata schemas
 - Item must be described with Dublin Core schema (15 general attributes – title, author etc.)
 - Metadata is exchanged via records (entire description of an item in the particular schema) or headers (basic description - id, modification date etc.)

Metadata exchange using OAI-PMH

- OAI-PMH Repository data structure
 - Repository may define sets (groups of items) – possible usage:
 - Subject sets, MIME type sets, ...
 - Set may contain subsets
 - Each item may belong to zero or more sets
 - Sets are used for selective harvesting – they allow to harvest only a part of a repository

Metadata reuse

- Overview of how Dublin Core is used in Polish digital libraries
- List of the most important requirements and recommendations coming from:
 - Europeana
 - Driver
 - „e-Guide for digital librarians”
 - recommendations gathered by Polish digital librarians

Analysis of DC usage

- Data gathered in the beginning of September 2009
- Approximately 280 000 objects
- Selected Dublin Core attributes will be described using:
 - General metrics
 - Number of unique **values** introduced in **all language versions** of objects' description
 - Metrics applied to **Polish version** of objects' description
 - Number of unique values
 - Number of **objects** with **at least one** occurrence of given attribute in description
 - Average number of values per object
 - Most frequent values for given attribute

References

- e-Guide for digital librarians
 - Version from 3rd July of 2008
 - Document exist only in Polish version
 - <http://www.bibliotekacyfrowa.pl/dlibra/doccontent?id=25069>

References

- Europeana Semantic Elements specifications
 - Version 3.2, 07/08/2009
 - https://group.europeana.eu/c/document_library/get_file?uuid=c56f82a4-8191-42fa-9379-4d5ff8c4ff75&groupId=10602
- Metadata Mapping & Normalisation Guidelines for the Europeana Prototype
 - Version 1.2, 07/08/2009
 - https://group.europeana.eu/c/document_library/get_file?uuid=58e2b828-b5f3-4feo-aa46-3dcbcoa2a1fo&groupId=10602

References

- DRIVER Guidelines 2.0: Guidelines for content providers - Exposing textual resources with OAI-PMH
 - Version 2.0, 13/11/2008
 - http://www.driver-support.eu/documents/DRIVER_Guidelines_v2_Final_2008-11-13.pdf

Characterization of references

- Documents from Europeana and Driver **are not** saying how objects should be classified
- It's a set of recommendations which are aimed to :
 - Simplify process of searching/browsing through metadata which comes from diverse sources
 - Automate workflows associated with metadata processing.

Selected Dublin Core Elements

- Subject
- Publisher
- Contributor
- Date
- Type
- Source
- Language
- Relation
- Coverage
- Rights

Subject - Metrics

- General metrics
 - Number of unique values: 93 795
- Metrics applied to Polish version of objects' description
 - Number of unique values: 92 149
 - Number of objects with at least one occurrence of given attribute in description: 169 794
 - Average number of values per object : 3,29



Subject - Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
gazety regionalne	12214	2,56%
czasopisma	7716	1,62%
prasa polska	5424	1,14%
czasopisma niemieckie	5009	1,05%
gazety sublokalne	4968	1,04%
Grodków	4962	1,04%
Grottkau	4961	1,04%
Wielkopolska	4422	0,93%
19 w.	4249	0,89%
Prusy	4164	0,87%
Czasopisma regionalne i lokalne polskie -19 w.	4140	0,87%
wiadomości polityczne	4094	0,86%
Gazety polskie - 1918-1939 r.	4077	0,85%
kultura	4071	0,85%
czasopisma sublokalne	3813	0,80%
Górny Śląsk	3731	0,78%
architektura	3566	0,75%
Wrocław	3515	0,74%
Śląsk	3448	0,72%
budownictwo	3388	0,71%



Subject - e-Guide for digital librarians

- Value should come from controlled vocabulary
- When creating new values it should be considered that users would use those words while searching
- If possible use the plural form of word/phrase



Subject – Europeana SE

- Object's subject



Subject – DRIVER Guidelines

- Two general possibilities:
 - Keywords
 - It is recommended to choose the most significant and unique keyword describing the object's content. Avoid to general keywords.
 - Classification
 - Dewey Decimal Classification (DDC) is the European recommendation



Subject – Conclusions/remarks/...

- Subject is often confused with the coverage attribute



Publisher - Metrics

- General metrics
 - Number of unique values: 22 563
- Metrics applied to Polish version of objects' description:
 - Number of unique values: 21 163
 - Number of objects with at least one occurrence of given attribute in description: 257 220
 - Average number of values per object: 1,70



Publisher – Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
Poznań	54943	12,62%
Telecomp Service na zlecenie PBI	22310	5,12%
Kraków	13662	3,14%
Warszawa	11245	2,58%
Toruń	11221	2,58%
Katowice	8187	1,88%
Drukarnia Polska	7998	1,84%
Drukarnia Dziennika Poznańskiego T.A.	6828	1,57%
Warszawa : Telecomp Service na zlecenie PBI	6824	1,57%
Drukarnia Dziennika Poznańskiego S.A.	5785	1,33%
Nakładem F[ranciszka] T[adeusza] Rakowicza	5406	1,24%
Kielce	5292	1,22%
Krakowskie Wydawnictwo Prasowe RSW "Prasa"	5137	1,18%
Breslau	5130	1,18%
E. Neugebauer	4959	1,14%
Wangefield	4959	1,14%
Grottkau	4959	1,14%
Bydgoszcz	4752	1,09%
Drukarnia Dziennika Poznańskiego	3923	0,90%
Drukarnia J. I. Kraszewskiego	3869	0,89%



Publisher - e-Guide for digital librarians

- Information about how object was published, printed and distributed coming from title page of object
- This attribute should be extended with Place where object was published
 - Base value (from controlled vocabulary) should be used e.g. „Wrocław”, other names under which given localization is known should be added as synonyms
 - e.g.. „Breslae, Breslau, Wratislavia, Wratislaviae”.



Publisher – Europeana SE

- Name of publisher/publishing house



Publisher – DRIVER guidelines

- Name of publisher/publishing house



Publisher – Conclusions/Remarks/...

- This field should contain name of publisher
- Definitely worth to expose the information about where object was published
 - Lack of appropriate metadata profile (none of dc:terms fits here)



Contributor - Metrics

- General metrics
 - Number of unique values: 15 379
- Metrics applied to Polish version of objects' description:
 - Number of unique values: 15 379
 - Number of objects with at least one occurrence of given attribute in description: 71 378
 - Average number of values per object: 1,36



Contributor – Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
Majerowicz Jan (red.)	3360	3,46%
Sienicki, Stanisław. Red.Wyd.	3010	3,10%
Grydzewski, Mieczysław (1894-1970)	1669	1,72%
Kossowska, Stefania (1909-2003)	1668	1,72%
Chmielowiec, Michał (1918-1974)	1668	1,72%
Biedermann, Marcin. Wydaw.	1494	1,54%
Wróbel Ludwik (red.)	1477	1,52%
Koszutski Kazimierz (red.)	1151	1,19%
Czajewski, Wiktor (1857-1922). Red.	1123	1,16%
Kleczyński, Jan (1837-1895). Red.	1084	1,12%
Rajchman, Aleksander (1855-1915). Red.	1008	1,04%
Chołodziński Jerzy (red.)	989	1,02%
Machejek Władysław (1920-1991). Red.	972	1,00%
Grodek, Anna. Red.	971	1,00%
Majerski, Michał (1873-1934)	948	0,98%
Roman, Jan Nepomucen	948	0,98%
Keller, Szczepan (1827-1872)	948	0,98%
Reich, Juliusz	948	0,98%
Michałowski, Edward Bolesław (1856-1905)	948	0,98%
Machejek, Władysław (1920-1991). Red.	946	0,97%
Findel Josef Gabriel (1828-1905) (hrsg.)	944	0,97%



Contributor - e-Guide for digital librarians

- Individual who cooperated during the original object creation/compilation
- It can be person or institution
- This attribute should contain the most widely known name of individual and indication which describes type of contribution e.g. Borkowski, Jan. Trans.



Contributor - e-Guide for digital librarians

- If one person contributed in a few different aspects of object creation all those activities should be mentioned
- Contributors names and contribution type indicators should come from controlled vocabulary



Contributor – Europeana SE

- An entity responsible for making contributions to the resource.
- Attribute's value describes source object creators
- This could be a person, an organization or a service.



Contributor- Driver Guidelines

- An entity responsible for making contributions to the content of the resource
- Names should be written in the same manner as in original edition
 - If this is not possible Driver recommends to use American Psychological Association (APA) style.



Contributor – Conclusions/remarks/...

- Diverse and bitty conventions for contribution type indicators
 - „Wyd.“, „Wydaw.“, „(red. odp.)“, „-red.“
- In analyzed data there is no coherent convention to tell which word is the surname, first name etc.
 - Sometimes there is no division at all
- Short and full names present
 - e.g. A. Einstein and Albert Einstein



Date - Metrics

- General metrics
 - Number of unique values : 52 395
- Metrics applied to Polish version of objects' description :
 - Number of unique values: 51 531
 - Number of objects with at least one occurrence of given attribute in description: 253 669
 - Average number of values per object : 1,0086



Date – Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
2004	4201	1,98%
2005.	3809	1,80%
2007.	3163	1,49%
2004.	2698	1,27%
2008	2494	1,18%
2007	2387	1,13%
1927	2167	1,02%
1926	2141	1,01%
1928	2098	0,99%
1867-1921	2098	0,99%
1929	2046	0,96%
1925	1732	0,82%
1924	1671	0,79%
2006	1645	0,78%
1930	1530	0,72%
1923	1503	0,71%
2005	1458	0,69%
1931	1394	0,66%
1922	1271	0,60%
1933	1268	0,60%



Date – Most frequent patterns

(Polish version of objects' description)

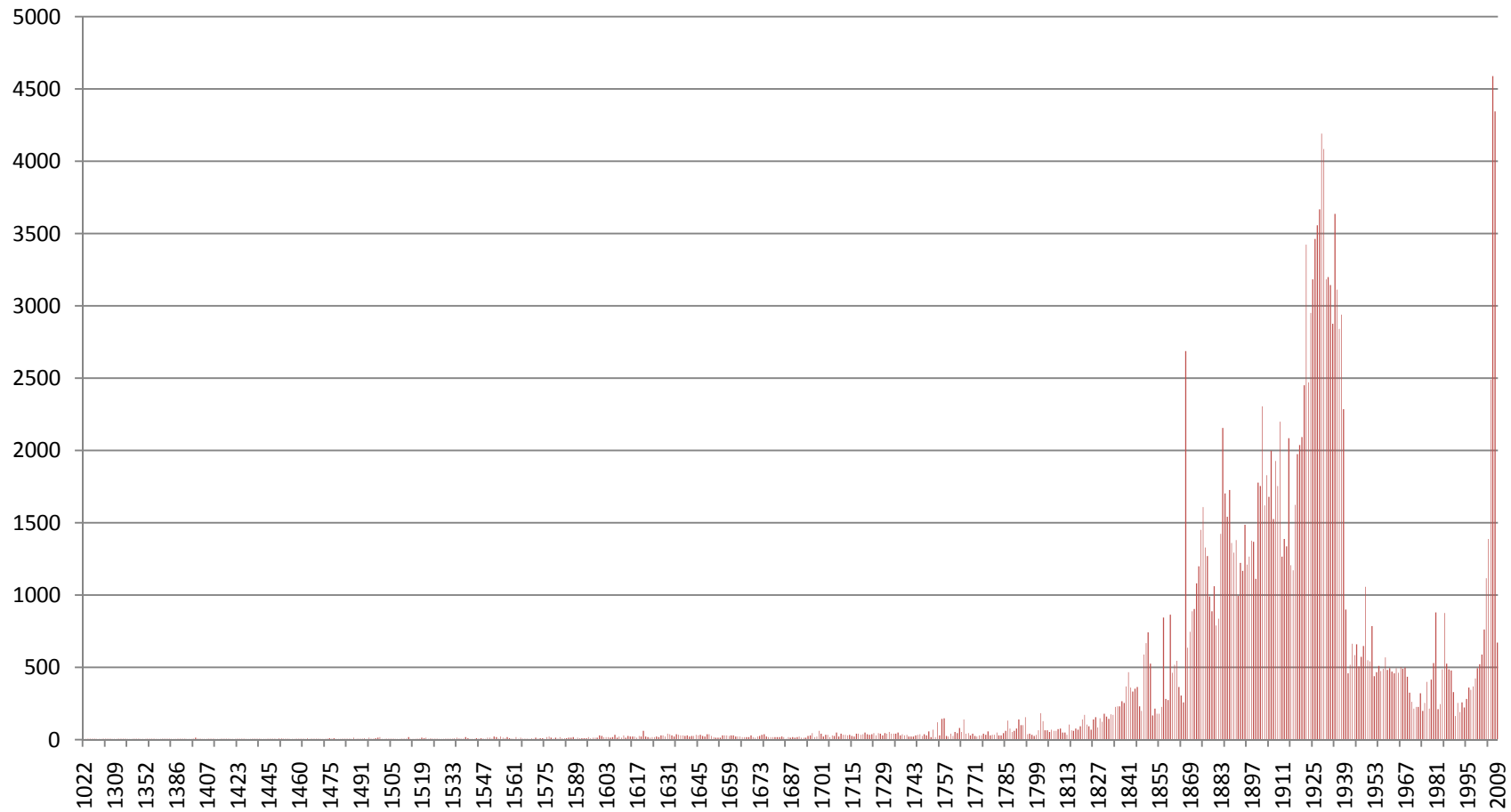
Pattern	No. of associations	% of all associations
DDDD	92 402	43,93%
DDDD.DD.DD	62 710	29,82%
DDDD-DD-DD	18 287	8,69%
DDDD-DDDD	8 935	4,25%
[DDDD]	4 327	2,06%
[ca DDDD]	3 208	1,53%
[DDDD-DDDD]	2 202	1,05%
[ante DDDD]	924	0,44%
DDDD.	906	0,43%
DDDD.DD	840	0,40%

(analysis from march 2009)



Date – Most frequent dates

(Polish version of objects' description)



2009-09-27

ECDL 2009, Corfu, Greece

(analysis from march 2009)



Date - e-Guide for digital librarians

- Date when original object was published
- Date should be encoded as defined in ISO 8601 (YYYY-MM-DD)
- If there is no information about publication date square brackets should be used to denote approximate date
- Date should be written using Arabic digits :
 - e.g. 17 w. = 1600-1700



Date – Europeana SE

- A point or period of time associated with an event in the lifecycle of the resource.
- Refined by dc:created (creation date) and dcterms:issued (publication date)
- At the moment there is no support for BC, BCE i BP
- Textual time periods should be converted into digits
 - 17 w. = 1601-1700



Date – DRIVER Guidelines

- ISO 8601
- When there is more than one important date in lifecycle of the source (e.g. creation, publication date etc.) this attribute should be used for publication date.
- Textual time periods should be converted into one the most representative date
 - 17 w. = 1650



Date - Conclusions/remarks/...

- ISO 8601
 - But more than 30% of date has a form of „DDDD.DD.DD”
- How express information about time periods?
- Controlled dictionary for date refinements?
 - dr., cop., dyst., ...
- How to distinguish approximate date?
 - [1234]



Type - Metrics

- General metrics
 - Number of unique values: 962
- Metrics applied to Polish version of objects' description
 - Number of unique values: 852
 - Number of objects with at least one occurrence of given attribute in description: 274 402
 - Average number of values per object: 1,01



Type – Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
czasopismo	66919	24,17%
gazety	50703	18,31%
Czasopismo	18174	6,56%
książka	15764	5,69%
gazeta	10323	3,73%
czasopisma	8673	3,13%
Gazeta	6653	2,40%
pocztówka	6363	2,30%
text	4894	1,77%
grafika	4451	1,61%
fotografia	4140	1,50%
artykuł z czasopisma	3343	1,21%
artykuł	2793	1,01%
dokument piśmienniczy	2159	0,78%
dzienniki urzędowe	2018	0,73%
starodruki	1743	0,63%
stary druk	1591	0,57%
rysunek	1540	0,56%
Czasopisma i gazety	1477	0,53%
rękopis	1239	0,45%
sprawozdanie szkolne	1117	0,40%
mapa	1091	0,39%
gazeta (gazety)	956	0,35%
rozprawa doktorska	954	0,34%
Czasopisma	919	0,33%



Type - e-Guide for digital librarians

- Type describes type or genre of original object
- Controlled vocabulary simplifies statistical analysis, but controlled voc. Is optional
- „e-Guide...” contains a proposition of such a vocabulary



Type – Europeana SE

- The nature or genre of the resource.
- Type includes terms describing general categories, functions, genres, or aggregation levels for content.
- The recommended best practice is to select a value from a controlled vocabulary
 - np. <http://dublincore.org/documents/dcmi-type-vocabulary>
- Europeana: TEXT, IMAGE, SOUND and VIDEO



Type – Driver Guidelines

- Type can appear more than once, but it's required (see below)
- Type can be used to describe:
 - Determine the type of original material
 - **required**, Driver's controlled vocabulary
 - Determine the type using vocabulary different than that defined by Driver
 - optional



Type – Driver Guidelines

- Type can be used to describe:
 - Indicate publishing lifecycle phase proper for described object
 - **recommended**, controlled vocabulary
 - e.g draft, submittedVersion



Type – Conclusions/remarks/...

- Different spelling
 - Small and big letters, diverse inflection
- English words used in Polish language version of object description
 - e.g. text (4894 occur.) or article (171 occur.)



Source - Metrics

- General metrics
 - Number of unique values : 25 307
- Metrics applied to Polish version of objects' description:
 - Number of unique values: 23 820
 - Number of objects with at least one occurrence of given attribute in description: 67 705
 - Average number of values per object : 1,09



Source – Most frequent values

(Polish version of objects' description)

Value	No. of associatiions	% of all associations
Biblioteka Narodowa	9173	12,41%
GSL P 31207 IV	4959	6,71%
GSL P 31731 IV	3666	4,96%
GSL P 28882 III	2902	3,92%
Miejska Biblioteka Publiczna w Radomiu	1587	2,15%
GSL P 31519 IV	1494	2,02%
08220 (BUW)	1073	1,45%
1010702 III	909	1,23%
GSL P 31741 IV	896	1,21%
GSL P 31521 IV	818	1,11%
GSL P 31883 IV	604	0,82%
Niniejszy rocznik "Kuriera Poznańskiego" został zdigitalizowany w Bibliotece Narodowej, która...	599	0,81%
Niniejszy rocznik "Kuriera Poznańskiego" został zdigitalizowany w Bibliotece Narodowej, która...	586	0,79%
011398 (BUW)	463	0,63%
2g.28.7.8. (BUW)	463	0,63%
4274 III	409	0,55%
4105 II	409	0,55%
Niniejszy rocznik "Kuriera Poznańskiego" został zdigitalizowany w Bibliotece Narodowej, która...	383	0,52%
djvu	348	0,47%
Niniejszy rocznik "Kuriera Poznańskiego" został zdigitalizowany w Bibliotece Narodowej, która...	346	0,47%
4453 IV	331	0,45%



Source - e-Guide for digital librarians

- Reference to the source object, which originates described object
- „e-Guide“ recommends to use some kind of formal identifier system e.g. archival signature of the source



Source – Europeana SE

- A related resource from which the described resource is derived in whole or in part.



Source – Driver Guidelines

- This attribute is reserved for results of digitisation
- For born-digital objects one should use dc:relation.
- It is recommended to use conventions described in:
 - <http://dublincore.org/documents/dc-citation-guidelines/>



Source – Conclusions/remarks/...

- This attribute is present in the description of appr. 25% resources.
- In 14% of all occurrences it is used in some different way (than described in DCMI)
 - e.g. Biblioteka Narodowa, MBP Radom



Language - Metrics

- General metrics
 - Number of unique values : 270
- Metrics applied to Polish version of objects' description:
 - Number of unique values: 205
 - Number of objects with at least one occurrence of given attribute in description: 273 549
 - Average number of values per object : 1,01



Language – Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
pol	225721	81,17%
ger	27708	9,96%
pol.	5362	1,93%
lat	4803	1,73%
niemiecki	3125	1,12%
eng	2889	1,04%
fre	2459	0,88%
und	2206	0,79%
rus	714	0,26%
polski	554	0,20%



Language - e-Guide for digital librarians

- Normalised 3 letter code (ISO 639-2)
- If object consists of a lot of languages **mul** should be used
- If information about object language is unknown, please use word **und**



Language - Wytyczne Europeany

- RFC 4646
 - <http://www.rfc-archive.org/getrfc.php?rfc=4646>
 - *language-script-region-variant-extension-privateuse*
 - Examples
 - Language information
 - en – english (ISO 639-1 and ISO 639-2, we choose ISO 639-1)
 - mas – masaian (ISO 639-2)
 - Language + region
 - fr-CA – French used in Canada
 - es-419 – Spanish used in Latin America
 - Language + script
 - zh-Hans – simplified Chinese



Language - Wytyczne DRIVERa

- ISO 639-3



Language - Conclusions/remarks/...

- Most often ISO 639-3
- When there is a need to describe language in a more detailed way use RFC 4646?



Relation - Metrics

- General metrics
 - Number of unique values : 12 171
- Metrics applied to Polish version of objects' description :
 - Number of unique values: 11 813
 - Number of objects with at least one occurrence of given attribute in description: 59 921
 - Average number of values per object: 1,41



Relation – Most frequent values

(Polish version of objects' description)

Value	No of associations	% of all associations
Dostępny również w formie drukowanej.	9159	11%
Bibliografia Gazety Toruńskiej	5406	6%
http://www.bu.uni.torun.pl/gazeta_torunska/	5406	6%
Mf. 41381	4959	6%
DVD GSL00035	4952	6%
DVD GSL00032	3608	4%
DVD GSL00006	2940	3%
Mf. 19882	2940	3%
Mf. 3913	1494	2%
DVD GSL00029	1494	2%
Echo Muzyczne 1877-1882	1073	1%
Echo Muzyczne i Teatralne 1906-1907	1065	1%
Echo Muzyczne i Teatralne 1883-1884	1008	1%
DVD GSL00027	896	1%
Czasopismo Techniczne. Architektura	640	1%
http://www.wbc.poznan.pl/publication/41558	637	1%
Mf. 19737	604	1%
Polska Gospodarcza 1930-1939	463	1%
Demobil 1921-1924	463	1%
Gazeta Lwowska	423	0%
Acta Universitatis Nicolai Copernici. Nauki Humanistyczno-Społeczne	332	0%



Relation - e-Guide for digital librarians

- Reference to other object with which source object is in some kind of relation
- Use in two cases:
 - For hierarchical bibliographic relations (e.g. series, parts of albums etc.)
 - To show relation between source and reference



Relation – Europeana SE

- Information about resources that are related to the original analog or born digital object
- The recommended best practice is to identify the resource using a formal identification scheme.
- ESE introduces the following refinements:
 - (dcterms:) isVersionOf, hasVersion
 - (dcterms:) isReplacedBy, replaces
 - (dcterms:) isRequiredBy, requires



Relation – Europeana SE

- ESE introduces the following refinements:
 - (dcterms:) isPartOf, hasPart
 - (dcterms:) isReferencedBy, references
 - (dcterms:) isFormatOf, hasFormat
 - (dcterms:) conformsTo



Relation – Driver Guidelines

- Link to related resources
- It is recommended to identify the resource using a formal identification scheme.
- Different digital manifestations of the same object should be related using **dc:relation**



Relation – Conclusions/remarks/...

- HTML embedded in attribute values
 - „źródło”



Coverage - Metrics

- General metrics
 - Number of unique values : 101
- Metrics applied to Polish version of objects' description:
 - Number of unique values: 72
 - Number of objects with at least one occurrence of given attribute in description : 2 531
 - Average number of values per object : 1,00



Coverage – Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
Dostępne tylko w sieci uczelnianej PŁ	1036	40,80%
Czasopismo przekazane w darze przez Ośrodek "KARTA"	882	34,74%
Kolekcja zawiera kilkaset pozycji. Prezentujemy jej fragment	491	19,34%
Polska	11	0,43%
Nr 1-13	8	0,32%
Nr 14-26	7	0,28%
Nr 27-39	6	0,24%
Nr 40-52	6	0,24%
Małopolskie, Województwo (Polska ; 1999-)	6	0,24%
Biblioteka Publiczna im. Stefana Rowińskiego w Ostrowie Wielkopolskim	5	0,20%
1954	4	0,16%
Podkarpackie, Województwo (Polska ; 1999-)	3	0,12%
20 w.	3	0,12%
Europa	3	0,12%
Skala [ca 1:300.000], podziałka w wiorstach.	3	0,12%
Galicja (region)	3	0,12%
Kraków (Polska)	2	0,08%
Skala [ca 1:26.000].	2	0,08%
Śródziemne, Morze (region)	2	0,08%
Nr 27-40	2	0,08%
Nr 1-12	2	0,08%



Coverage - e-Guide for digital librarians

- The spatial or temporal topic of the resource, the spatial applicability of the resource, or the jurisdiction under which the resource is relevant.
 - e.g. named place, a location, a spatial coordinate, a period, date, date range, name of administrative unit
- It's recommended to use controlled vocabulary



Coverage – Europeana ESE

- Defined as above (slide about „e-Guide for digital librarians“)
- ESE encourages to use two Dublin Core refinements:
 - **dcterms:spatial** – spatial characteristics of the original analog or born digital object
 - e.g. named place, spatial coordinate
 - **dcterms:temporal** – temporal characteristics of the original analog or born digital object.



Coverage – Driver Guidelines

- Defined as above (slide about „e-Guide for digital librarians“)
- It is recommended to use controlled vocabulary
 - e.g. Getty Thesaurus of Geographic Names (TGN)



Coverage – Conclusions/remarks/...

- Scarce usage – only 2539 objects
- In most cases (95%) it contains very weird things e.g.
 - „Available only from University Network“
 - „This journal was donated by...“



Rights - Metrics

- General metrics
 - Number of unique values : 1 453
- Metrics applied to Polish version of objects' description:
 - Number of unique values: 1 283
 - Number of objects with at least one occurrence of given attribute in description : 200 981
 - Average number of values per object : 1,40



Rights – Most frequent values

(Polish version of objects' description)

Value	No. of associations	% of all associations
Biblioteka Kórnicka PAN	54840	19,47%
Biblioteka Uniwersytecka w Poznaniu	22915	8,13%
http://www.bu.uni.wroc.pl/bc/regulamin.html	20076	7,13%
Biblioteka Narodowa	20056	7,12%
Biblioteka Uniwersytecka we Wrocławiu	19999	7,10%
Regulamin korzystania z Biblioteki Cyfrowej Uniwersytetu Wrocławskiego	19595	6,96%
Wojewódzka Biblioteka Publiczna w Krakowie	16206	5,75%
Wydawnictwo Jagiellonia SA	7605	2,70%
http://www.dziennik.krakow.pl/	7270	2,58%
http://www.wbp.krakow.pl/	5319	1,89%
Wydawca	4832	1,72%
Biblioteka Uniwersytecka im. Jerzego Giedroycia w Białymstoku	4535	1,61%
Pracownia Działu Digitalizacji i Zbiorów Cyfrowych WBP Kielce.	4430	1,57%
WiMBP im. Marszałka J. Piłsudskiego w Łodzi	3390	1,20%
Biblioteka Uniwersytetu Zielonogórskiego	3356	1,19%
Biblioteka Główna AGH	2929	1,04%
http://www.wbp.krakow.pl	2812	1,00%
Archiwum Państwowe w Poznaniu (zasób zdigitalizowany ze środków Ministra Kultury i Dziedzictwa Narodowego)	2790	0,99%
Publikacja chroniona prawem autorskim - reprodukcja cyfrowa dostępna w czytelnich BN	2787	0,99%
Uniwersytet Mikołaja Kopernika w Toruniu	2391	0,85%



Rights - e-Guide for digital librarians

- „ Information about user's rights to use the digital object“



Rights – Europeana SE

- Information about:
 - Intellectual Property Rights
 - Access rights or license arrangements for the digital object (digitized or born digital).



Rights – DRIVER Guidelines

- See the definitions above (see slide about Europeana SE)
- It is preferred to refer to a rights service where the reuse rights are made clear to the end-user by using a URL
- Examples:
 - (c) University of Bath, 2003
 - <http://creativecommons.org/licenses/by-sa/2.0/uk/>
 - cc-by-sa, Andrew Smith



Rights - Conclusions/remarks/...

- Usually it is information about owner of IPR
- In most of cases :
 - there is no clear explanation what user can do with the digital content
 - Only information that this resource is available from university's intranet

