Formats, metadata, standards and vocabularies for national bibliographic databases

Ivanović Dragan

University of Novi Sad

ENRESSH Training school
Outline

1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
The first faculty in Novi Sad was founded in 1954

The University of Novi Sad was founded on 28th of June 1960

Today, UNS represents an autonomous institution for education, science and arts
The first faculty in Novi Sad was founded in 1954.
The University of Novi Sad was founded on 28th of June 1960.
Today, UNS represents an autonomous institution for education, science and arts.
The first faculty in Novi Sad was founded in 1954
The University of Novi Sad was founded on 28th of June 1960
Today, UNS represents an autonomous institution for education, science and arts
Rectorate building

UNIVERSITY OF NOVI SAD-14 FACULTIES, 50.000 STUDENTS, 5.000 EMPLOYEES

Harmony of nature and modern architecture on the banks of the Danube-Rectorate Building
Introduction
Good practices
Metadata mapping
Conclusion

University of Novi Sad Cities

Formats, metadata, standards and vocabularies
Outline

1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
Metadata vs data

- Metadata commonly are understood as ‘data about data’
- The content of bibliographic databases are bibliographic metadata referring to research output
- Research outputs (pdf, xls, etc) represent data, while bibliographic databases store metadata - data about research outputs
- That is especially case if you are looking at bibliographic database as source for publications discovery (information retrieval)
- However, if you are looking at bibliographic database as source for bibliometrics analysis or research evaluation, then content of database could be called data
Metadata vs data

- Metadata commonly are understood as ‘data about data’
- The content of bibliographic databases are bibliographic metadata referring to research output
- Research outputs (pdf, xls, etc) represent data, while bibliographic databases store metadata - data about research outputs
- That is especially case if you are looking at bibliographic database as source for publications discovery (information retrieval)
- However, if you are looking at bibliographic database as source for bibliometrics analysis or research evaluation, then content of database could be called data
Metadata vs data

- Metadata commonly are understood as ‘data about data’
- The content of bibliographic databases are bibliographic metadata referring to research output
- Research outputs (pdf, xls, etc) represent data, while bibliographic databases store metadata - data about research outputs
- That is especially case if you are looking at bibliographic database as source for publications discovery (information retrieval)
- However, if you are looking at bibliographic database as source for bibliometrics analysis or research evaluation, then content of database could be called data
Metadata vs data

- Metadata commonly are understood as ‘data about data’
- The content of bibliographic databases are bibliographic metadata referring to research output
- Research outputs (pdf, xls, etc) represent data, while bibliographic databases store metadata - data about research outputs
- That is especially case if you are looking at bibliographic database as source for publications discovery (information retrieval)
- However, if you are looking at bibliographic database as source for bibliometrics analysis or research evaluation, then content of database could be called data
Metadata vs data

- Metadata commonly are understood as ‘data about data’
- The content of bibliographic databases are bibliographic metadata referring to research output
- Research outputs (pdf, xls, etc) represent data, while bibliographic databases store metadata - data about research outputs
- That is especially case if you are looking at bibliographic database as source for publications discovery (information retrieval)
- However, if you are looking at bibliographic database as source for bibliometrics analysis or research evaluation, then content of database could be called data
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
  - national evaluation rule-books
  - mandatory vs optional
  - rich vs light

- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
  - structured database vs csv vs xml vs json, etc
  - metadata schema
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
    - national evaluation rule-books
    - mandatory vs optional
    - rich vs light

- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
  - structured database vs csv vs xml vs json, etc
  - metadata schema
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
  - national evaluation rule-books
  - mandatory vs optional
  - rich vs light

- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
  - structured database vs csv vs xml vs json, etc
  - metadata schema
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
  - national evaluation rule-books
  - mandatory vs optional
  - rich vs light

- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
  - structured database vs csv vs xml vs json, etc
  - metadata schema
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
  - national evaluation rule-books
  - mandatory vs optional
  - rich vs light

- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
  - structured database vs csv vs xml vs json, etc
  - metadata schema
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
  - national evaluation rule-books
  - mandatory vs optional
  - rich vs light
- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
    - structured database vs csv vs xml vs json, etc
    - metadata schema
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
  - national evaluation rule-books
  - mandatory vs optional
  - rich vs light
- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
  - structured database vs csv vs xml vs json, etc
  - metadata schema
Which metadata vs in which format

- Which metadata should be preserved in bibliographic database is one question
  - purpose
  - needs
  - national evaluation rule-books
  - mandatory vs optional
  - rich vs light

- In which format metadata should be preserved is the another question
  - how to select best format for preservation?
  - structured database vs csv vs xml vs json, etc
  - metadata schema
Standards

Which standard formats to be supported for export?

Which protocol should be implemented for harvesting metadata from/to the system?

OAI-PMH, OpenAIRE guidelines, SRU/W, etc.
Standards

Which standard formats to be supported for export?
Which protocol should be implemented for harvesting metadata from/to the system?
OAI-PMH, OpenAIRE guidelines, SRU/W, etc.
Standards

- Which standard formats to be supported for export?
- Which protocol should be implemented for harvesting metadata from/to the system?
- OAI-PMH, OpenAIRE guidelines, SRU/W, etc.
Vocabularies

- Not related to the structure of metadata or format
- Related to the content - allowed values/terms for metadata
- Publication types?
- Question very important for interoperability of systems
- If we speak languages which have similar rules and structures (nouns, verbs, etc), but we use different terms - can we communicate?
Vocabularies

- Not related to the structure of metadata or format
- Related to the content - allowed values/terms for metadata
- Publication types?
- Question very important for interoperability of systems
- If we speak languages which have similar rules and structures (nouns, verbs, etc), but we use different terms - can we communicate?
Vocabularies

- Not related to the structure of metadata or format
- Related to the content - allowed values/terms for metadata
- Publication types?
  - Question very important for interoperability of systems
  - If we speak languages which have similar rules and structures (nouns, verbs, etc), but we use different terms - can we communicate?
Vocabularies

- Not related to the structure of metadata or format
- Related to the content - allowed values/terms for metadata
- Publication types?
- Question very important for interoperability of systems
- If we speak languages which have similar rules and structures (nouns, verbs, etc), but we use different terms - can we communicate?
Vocabularies

- Not related to the structure of metadata or format
- Related to the content - allowed values/terms for metadata
- Publication types?
- Question very important for interoperability of systems
- If we speak languages which have similar rules and structures (nouns, verbs, etc), but we use different terms - can we communicate?
Outline

1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
Recommendation 3

- Define the data model and/or metadata schema, taking into account the database’s purpose and recognized standards
  - Ensures that the system can fulfill its purpose, while following recognized standards simplifies the work and can benefit interoperability
  - Majority of bibliographic standards do not take evaluation purposes into account
Recommendation 3

- Define the data model and/or metadata schema, taking into account the database’s purpose and recognized standards
- Ensures that the system can fulfil its purpose, while following recognized standards simplifies the work and can benefit interoperability
- Majority of bibliographic standards do not take evaluation purposes into account
Recommendation 3

- Define the data model and/or metadata schema, taking into account the database’s purpose and recognized standards

- Ensures that the system can fulfil its purpose, while following recognized standards simplifies the work and can benefit interoperability

- Majority of bibliographic standards do not take evaluation purposes into account
Recommendation 4

- Select a suitable technical solution and design the technical structure of the database
  - Contributes to the functionality, performance, and maintainability of the database
  - Purpose, budget, the estimated number of records/requests, contemporary technologies/databases, experience of staff - technicians and librarians should be taken into account
Recommendation 4

- Select a suitable technical solution and design the technical structure of the database
- Contributes to the functionality, performance, and maintainability of the database
- Purpose, budget, the estimated number of records/requests, contemporary technologies/databases, experience of staff - technicians and librarians should be taken into account
Recommendation 4

- Select a suitable technical solution and design the technical structure of the database
- Contributes to the functionality, performance, and maintainability of the database
- Purpose, budget, the estimated number of records/requests, contemporary technologies/databases, experience of staff - technicians and librarians should be taken into account
Outline

1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
Recommendation 12

- Maintain authority lists for publication channels
  - Contributes to the accuracy of data on publication channels and the functionality of the database
  - Journals, conferences, publishers, etc
  - local/external identifiers, title, etc.
Recommendation 12

- Maintain authority lists for publication channels
- Contributes to the accuracy of data on publication channels and the functionality of the database
  - Journals, conferences, publishers, etc
  - local/external identifiers, title, etc.
Recommendation 12

- Maintain authority lists for publication channels
- Contributes to the accuracy of data on publication channels and the functionality of the database
- Journals, conferences, publishers, etc
- local/external identifiers, title, etc.
Recommendation 12

- **Maintain authority lists for publication channels**
- Contributes to the accuracy of data on publication channels and the functionality of the database
- Journals, conferences, publishers, etc
- local/external identifiers, title, etc.
Recommendation 13

- **Maintain authority lists for authors and organisations**
  - Contributes to the accuracy of data on authors and organisations and the functionality of the database
  - Local/external (ORCID) identifiers, inside/outside database scope (national/international person/organization), name, history/variations of names, etc.
Recommendation 13

- **Maintain authority lists for authors and organisations**
- Contributes to the accuracy of data on authors and organisations and the functionality of the database
- local/external (ORCID) identifiers, inside/outside database scope (national/international person/organization), name, history/variations of names, etc.
**Recommendation 13**

- **Maintain authority lists for authors and organisations**
- Contributes to the accuracy of data on authors and organisations and the functionality of the database
- Local/external (ORCID) identifiers, inside/outside database scope (national/international person/organization), name, history/notifications of names, etc.
Recommendation 14

- Use international persistent identifiers where possible
  - Increases interoperability with other national and international databases and systems
  - ORCID, DOI, ISSN, ISBN, etc
Recommendation 14

- Use international persistent identifiers where possible
- Increases interoperability with other national and international databases and systems
- ORCID, DOI, ISSN, ISBN, etc
Recommendation 14

- Use international persistent identifiers where possible
- Increases interoperability with other national and international databases and systems
- ORCID, DOI, ISSN, ISBN, etc
Recommendation 15

- Use as much as possible terms from well-known and standardized vocabularies
  - Enhances the interoperability and functionality of the database
  - Languages’ and countries’ codes, publication types and scientific fields (problematic!)
Recommendation 15

- Use as much as possible terms from well-known and standardized vocabularies
- Enhances the interoperability and functionality of the database
- Languages’ and countries’ codes, publication types and scientific fields (problematic!)
Recommendation 15

- Use as much as possible terms from well-known and standardized vocabularies
- Enhances the interoperability and functionality of the database
- Languages’ and countries’ codes, publication types and scientific fields (problematic!)
Recommendation 16

- **When developing own vocabulary, consult stakeholders and relevant experts**
- Ensures that the vocabulary is usable and captures all use cases
- "From scratch" or extended standard vocabulary, human/machine readable vocabulary, SKOS semantic relations vocabulary
Recommendation 16

- When developing own vocabulary, consult stakeholders and relevant experts
- Ensures that the vocabulary is usable and captures all use cases
- "From scratch" or extended standard vocabulary, human/machine readable vocabulary, SKOS semantic relations vocabulary
Recommendation 16

- When developing own vocabulary, consult stakeholders and relevant experts
- Ensures that the vocabulary is usable and captures all use cases
- "From scratch" or extended standard vocabulary, human/machine readable vocabulary, SKOS semantic relations vocabulary
Outline

1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
Recommendation 22

- Specify procedures for data access
  - Enhances the usability of the database
  - user interface, API, protocol(s) for harvesting/federated search, etc.
  - Take into account licences (GDPR), needs of different users, different ways to transfer data
Recommendation 22

- **Specify procedures for data access**
- Enhances the usability of the database
  - user interface, API, protocol(s) for harvesting/federated search, etc.
- Take into account licences (GDPR), needs of different users, different ways to transfer data
Recommendation 22

- **Specify procedures for data access**
- Enhances the usability of the database
- user interface, API, protocol(s) for harvesting/federated search, etc.
- Take into account licences (GDPR), needs of different users, different ways to transfer data
Recommendation 22

- Specify procedures for data access
- Enhances the usability of the database
- user interface, API, protocol(s) for harvesting/federated search, etc.
- Take into account licences (GDPR), needs of different users, different ways to transfer data
Recommendation 23

- Offer research output metadata in multiple representations
  - Ensures that users with different needs and preferences can efficiently use the data
  - User profiles and preferences are different, option to customize the display and format, export to standardized formats, XML, BibTex, JSON, RDF - semantic web (FAIR principles)
Recommendation 23

- Offer research output metadata in multiple representations
- Ensures that users with different needs and preferences can efficiently use the data
  - user profiles and preferences are different, option to customize the display and format, export to standardized formats, XML, Bibtex, JSON, RDF - semantic web (FAIR principles)

dragan.ivanovic@uns.ac.rs

Formats, metadata, standards and vocabularies
Recommendation 23

- **Offer research output metadata in multiple representations**
- Ensures that users with different needs and preferences can efficiently use the data
- User profiles and preferences are different, option to customize the display and format, export to standardized formats, XML, Bibtex, JSON, RDF - semantic web (FAIR principles)
Recommendation 24

- **Provide access to the data through a functional user interface**
- Enables consulting the database in various ways and increases transparency
- Searching (basic and advance), browsing, downloading
Recommendation 24

- Provide access to the data through a functional user interface
- Enables consulting the database in various ways and increases transparency
- Searching (basic and advance), browsing, downloading
Recommendation 24

- Provide access to the data through a functional user interface
- Enables consulting the database in various ways and increases transparency
- Searching (basic and advance), browsing, downloading
Recommendation 25

- Facilitate automated access to the data through an API or a metadata harvesting protocol
- Enables automated and efficient use of the database
- REST, JSON vs XML, authentication and authorization (A1.2 FAIR principle), OAI-PMH, OAI-ORE, etc.
Recommendation 25

- Facilitate automated access to the data through an API or a metadata harvesting protocol
- Enables automated and efficient use of the database
- REST, JSON vs XML, authentication and authorization (A1.2 FAIR principle), OAI-PMH, OAI-ORE, etc.
Recommendation 25

- Facilitate automated access to the data through an API or a metadata harvesting protocol
- Enables automated and efficient use of the database
- REST, JSON vs XML, authentication and authorization (A1.2 FAIR principle), OAI-PMH, OAI-ORE, etc.
Recommendation 26

- Enable crawling of bibliographic records by web search engines
- Ensures that database content can be found through academic search engines
- Crawlers, Robots Exclusion Protocol (robots.txt), specific crawling guidelines (Google Scholar)
Recommendation 26

- Enable crawling of bibliographic records by web search engines
- Ensures that database content can be found through academic search engines
- Crawlers, Robots Exclusion Protocol (robots.txt), specific crawling guidelines (Google Scholar)
Recommendation 26

- Enable crawling of bibliographic records by web search engines
- Ensures that database content can be found through academic search engines
- Crawlers, Robots Exclusion Protocol (robots.txt), specific crawling guidelines (Google Scholar)
Outline

1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
Why

- EU services - evaluation for EU funded projects, reporting, etc
- Publications/outputs discovery
Why

- EU services - evaluation for EU funded projects, reporting, etc
- Publications/outputs discovery
Approaches

- **Distributed vs Centralized**
  - The distributed approach makes it easier to have complete information in real-time, since it does not require propagation of updates to the central catalogue - federated search SRU/W
  - However, for data-intensive operations, the centralized approach doesn’t have the problem of querying multiple sites, and has more complete overview of the data available when executing operations - harvesting data OAI-PMH
Approaches

- Distributed vs Centralized

The distributed approach makes it easier to have complete information in real-time, since it does not require propagation of updates to the central catalogue - federated search SRU/W

However, for data-intensive operations, the centralized approach doesn’t have the problem of querying multiple sites, and has more complete overview of the data available when executing operations - harvesting data OAI-PMH
**Approaches**

- **Distributed vs Centralized**
  - The distributed approach makes it easier to have complete information in real-time, since it does not require propagation of updates to the central catalogue - federated search SRU/W
  - However, for data-intensive operations, the centralized approach doesn’t have the problem of querying multiple sites, and has more complete overview of the data available when executing operations - harvesting data OAI-PMH
Distributed SRU/W based approach
Centralized OAI-PMH based approach

![Diagram of Centralized OAI-PMH based approach](image)
Centralized approach

- Data provider (nodes) and Service provider (Integrated European Publication Information Service)
  - Protocols for harvesting metadata should be implemented on both side (OAI-PMH, ResourceSync, etc.)
  - Target metadata format(s) should be selected
  - All nodes (partner systems) have to export metadata to (at least one) target metadata format
  - All nodes (data providers) have to map its metadata to target metadata format
Centralized approach

- Data provider (nodes) and Service provider (Integrated European Publication Information Service)
- Protocols for harvesting metadata should be implemented on both side (OAI-PMH, ResourceSync, etc.)
- Target metadata format(s) should be selected
- All nodes (partner systems) have to export metadata to (at least one) target metadata format
- All nodes (data providers) have to map its metadata to target metadata format
Centralized approach

- Data provider (nodes) and Service provider (Integrated European Publication Information Service)
- Protocols for harvesting metadata should be implemented on both side (OAI-PMH, ResourceSync, etc.)
- Target metadata format(s) should be selected
  - All nodes (partner systems) have to export metadata to (at least one) target metadata format
  - All nodes (data providers) have to map its metadata to target metadata format
Centralized approach

- Data provider (nodes) and Service provider (Integrated European Publication Information Service)
- Protocols for harvesting metadata should be implemented on both side (OAI-PMH, ResourceSync, etc.)
- Target metadata format(s) should be selected
- All nodes (partner systems) have to export metadata to (at least one) target metadata format
- All nodes (data providers) have to map its metadata to target metadata format
Centralized approach

- Data provider (nodes) and Service provider (Integrated European Publication Information Service)
- Protocols for harvesting metadata should be implemented on both side (OAI-PMH, ResourceSync, etc.)
- Target metadata format(s) should be selected
- All nodes (partner systems) have to export metadata to (at least one) target metadata format
- All nodes (data providers) have to map its metadata to target metadata format
Process

- Matching source schema entities to target schema entities
- Matching source attributes to target attributes
- Expressing the mapping in some format/language
- Implementation of mappings rules in source system
Process

- Matching source schema entities to target schema entities
- Matching source attributes to target attributes
- Expressing the mapping in some format/language
- Implementation of mappings rules in source system
Process

- Matching source schema entities to target schema entities
- Matching source attributes to target attributes
- Expressing the mapping in some format/language
- Implementation of mappings rules in source system
Process

- Matching source schema entities to target schema entities
- Matching source attributes to target attributes
- Expressing the mapping in some format/language
- Implementation of mappings rules in source system
A key aspect to consider in the mapping process is the involvement of various actors who play crucial roles in ensuring the success of the project. These actors include:

1. **Expert(s) for source schema**
2. **Expert(s) for target schema**
3. **Expert(s) for source/target vocabularies**
4. **Software developers**

Each role is essential for different aspects of the mapping process, from understanding the source and target schemas to integrating the vocabularies and implementing the mapping tools.
A. Actors

1. Expert(s) for source schema
2. Expert(s) for target schema
3. Expert(s) for source/target vocabularies
4. Software developers
Actors

1. Expert(s) for source schema
2. Expert(s) for target schema
3. Expert(s) for source/target vocabularies
4. Software developers
Actors

1. Expert(s) for source schema
2. Expert(s) for target schema
3. Expert(s) for source/target vocabularies
4. Software developers
Collaboration

- Collaboration between schema/vocabularies experts is usually not a problem
- However, collaboration between those experts and software developers could be a problem
  - Don’t "speak" the same language
  - The process of implementation of mappings rules in source system is error-prone and time-consuming
  - Can we automate the process? Can complete process be performed by schema/vocabularies experts?
Collaboration

- Collaboration between schema/vocabularies experts is usually not a problem
- However, collaboration between those experts and software developers could be a problem
  - Don’t "speak" the same language
  - The process of implementation of mappings rules in source system is error-prone and time-consuming
  - Can we automate the process? Can complete process be performed by schema/vocabularies experts?
Collaboration

- Collaboration between schema/vocabularies experts is usually not a problem
- However, collaboration between those experts and software developers could be a problem
  - Don’t "speak" the same language
  - The process of implementation of mappings rules in source system is error-prone and time-consum ing
  - Can we automate the process? Can complete process be performed by schema/vocabularies experts?
Collaboration between schema/vocabularies experts is usually not a problem.

However, collaboration between those experts and software developers could be a problem:

- Don’t "speak" the same language.
- The process of implementation of mappings rules in source system is error-prone and time-consuming.
- Can we automate the process? Can complete process be performed by schema/vocabularies experts?
Collaboration

- Collaboration between schema/vocabularies experts is usually not a problem.
- However, collaboration between those experts and software developers could be a problem:
  - Don’t "speak" the same language.
  - The process of implementation of mappings rules in source system is error-prone and time-consuming.
  - Can we automate the process? Can complete process be performed by schema/vocabularies experts?
Outline

1. Introduction
   - My University
   - Questions/challenges

2. Good practices
   - Design
   - Vocabularies, authority control and identifiers
   - Data use

3. Metadata mapping
   - Integrated European Publication Information Service
   - Mapping process
   - Mapping tools

4. Conclusion
The process of matching and mapping implies a lot of time and effort from experts on the source and target schemata.

To simplify and accelerate the process, a tool needs to be adopted for automation.

Besides enhancement of mapping development, such a tool should make the implementation of mappings more effective and shareable.
Why

- The process of matching and mapping implies a lot of time and effort from experts on the source and target schemata.
- To simplify and accelerate the process, a tool needs to be adopted for automation.
- Besides enhancement of mapping development, such a tool should make the implementation of mappings more effective and shareable.
The process of matching and mapping implies a lot of time and effort from experts on the source and target schemata.

To simplify and accelerate the process, a tool needs to be adopted for automation.

Besides enhancement of mapping development, such a tool should make the implementation of mappings more effective and shareable.
The X3ML toolkit with the 3M editor could be used to automate the mappings.

This toolkit allows several steps and tasks of the process of harvesting, matching, mapping and integrating the data from the sources to the target catalogue.

3M (one component of X3ML toolkit) guides the user to specify the schemata matchings and the instances generators.

X3ML engine (the another X3ML toolkit component) automatically transforms the source data into target format.
X3ML toolkit

- The X3ML toolkit with the 3M editor could be used to automate the mappings.
- This toolkit allows several steps and tasks of the process of harvesting, matching, mapping and integrating the data from the sources to the target catalogue.
- 3M (one component of X3ML toolkit) guides the user to specify the schemata matchings and the instances generators.
- X3ML engine (the another X3ML toolkit component) automatically transforms the source data into target format.
The X3ML toolkit with the 3M editor could be used to automate the mappings.

This toolkit allows several steps and tasks of the process of harvesting, matching, mapping and integrating the data from the sources to the target catalogue.

3M (one component of X3ML toolkit) guides the user to specify the schemata matchings and the instances generators.

X3ML engine (the another X3ML toolkit component) automatically transforms the source data into target format.
The X3ML toolkit with the 3M editor could be used to automate the mappings.

This toolkit allows several steps and tasks of the process of harvesting, matching, mapping and integrating the data from the sources to the target catalogue.

3M (one component of X3ML toolkit) guides the user to specify the schemata matchings and the instances generators.

X3ML engine (the another X3ML toolkit component) automatically transforms the source data into target format.
X3ML toolkit

Integrated European Publication Information Service
Mapping process
Mapping tools

Formats, metadata, standards and vocabularies
3M eases the process of matching by parsing and analyzing the source and target schemata, thus allowing auto-completion when selecting the entities and properties to be matched.

This mechanism speeds the matching process and allows non-expert users (users that do not have an extended knowledge of the whole schema) to define a matching.
3M eases the process of matching by parsing and analyzing the source and target schemata, thus allowing auto-completion when selecting the entities and properties to be matched.

This mechanism speeds the matching process and allows non-expert users (users that do not have an extended knowledge of the whole schema) to define a matching.
The description of the matching is homogenized, which reduces the misunderstandings between experts and software developers.

3M also includes a versioning mechanism that allows storage of different versions of the matchings.

The X3ML engine can be used exhaustively to test any version of the matching at any time just by providing a sample of data and applying the transformation.

The result is immediately available and can be analysed to check for defaults or implemented corrections.
The description of the matching is homogenized, which reduces the misunderstandings between experts and software developers.

3M also includes a versioning mechanism that allows storage of different versions of the matchings.

The X3ML engine can be used exhaustively to test any version of the matching at any time just by providing a sample of data and applying the transformation.

The result is immediately available and can be analysed to check for defaults or implemented corrections.
The description of the matching is homogenized, which reduces the misunderstandings between experts and software developers.

3M also includes a versioning mechanism that allows storage of different versions of the matchings.

The X3ML engine can be used exhaustively to test any version of the matching at any time just by providing a sample of data and applying the transformation.

The result is immediately available and can be analysed to check for defaults or implemented corrections.
The description of the matching is homogenized, which reduces the misunderstandings between experts and software developers.

3M also includes a versioning mechanism that allows storage of different versions of the matchings.

The X3ML engine can be used exhaustively to test any version of the matching at any time just by providing a sample of data and applying the transformation.

The result is immediately available and can be analysed to check for defaults or implemented corrections.
3M demo

- Dublin Core is the source - link
- CERIF RDF should be the result
- https://isl.ics.forth.gr/3M
- Mapping Project - ENRESSH Dublin Core to CERIF 1.6
3M demo

- Dublin Core is the source - link
- CERIF RDF should be the result
- https://isl.ics.forth.gr/3M
- Mapping Project - ENRESSH Dublin Core to CERIF 1.6
3M demo

- Dublin Core is the source - link
- CERIF RDF should be the result
- https://isl.ics.forth.gr/3M
- Mapping Project - ENRESSH Dublin Core to CERIF 1.6
3M demo

- Dublin Core is the source - link
- CERIF RDF should be the result
- https://isl.ics.forth.gr/3M
- Mapping Project - ENRESSH Dublin Core to CERIF 1.6
Summary

- Part of Manuel of good practices: CREATING AND MAINTAINING A NATIONAL BIBLIOGRAPHIC DATABASE FOR RESEARCH OUTPUT has been presented

- In order to improve reusability of metadata, the system could be a data provider and could export metadata to some Service Provider(s)

- Source metadata schemata should be mapped to target metadata schemata
Part of Manuel of good practices: CREATING AND MAINTAINING A NATIONAL BIBLIOGRAPHIC DATABASE FOR RESEARCH OUTPUT has been presented.

In order to improve reusability of metadata, the system could be a data provider and could export metadata to some Service Provider(s).

Source metadata schemata should be mapped to target metadata schemata.
Summary

- Part of Manuel of good practices: CREATING AND MAINTAINING A NATIONAL BIBLIOGRAPHIC DATABASE FOR RESEARCH OUTPUT has been presented.
- In order to improve reusability of metadata, the system could be a data provider and could export metadata to some Service Provider(s).
- Source metadata schemata should be mapped to target metadata schemata.
Thank you for your attention!!!
If you have any questions, please do not hesitate to
  ask me during the school
  contact me by email - dragan.ivanovic@uns.ac.rs