System for modelling rulebooks for the evaluation of scientific-research results. Case study: Serbian Rulebook

Siniša Nikolić, Valentin Penca, Dragan Ivanović
University of Novi Sad/Faculty of Technical Sciences/Department of Computing and Automatics, Novi Sad, Serbia
{sinisa_nikolic, valentin_penca, chenejac}@uns.ac.rs

Abstract—The paper presents an example of a system for storing information about rulebooks which are used for evaluation of scientific-research results of researchers. System is based on the CERIF data model and it was implemented as extension of the current CRIS system of the University of Novi Sad (CRIS UNS), where it is actively used by research community. Case study for modelling rulebook which was issued by the Ministry of Education and Science of the Republic of Serbia was presented in this paper.

I. INTRODUCTION

Evaluation in scientific-research domain is necessary as in all other areas. Today, we are witnesses of many efforts, aimed at objective and efficient evaluation of scientific-research results. Evaluation in scientific-research domain [1] is a process based on critical analysis of information and data which leads to a judgment of merit. Main output of scientific-research activity is scientific-research result as: journal paper, monograph, etc. Evaluation of scientific-research results is done for the purpose of: an election of a researcher in scientific and teaching positions within the scientific-research institutions; ranking of the researchers and the scientific institutions; financing of the scientific projects, etc.

Nowadays, there are number of research management systems for storing scientific research information. Those research management systems take an important role in the development of science [2]. Those systems usually contain information on publications, researchers and research institutions. An appropriate example of such system is the Current Research Information System (CRIS) that is based on the Common European Research Information Format (CERIF) standard.

Construction of an information system is necessary for efficient evaluation of the scientific-research data [3]. CRIS research management systems represents a good starting point for development of a system for evaluation of scientific results, because their metadata can be used for evaluation of scientific-research results of researchers in accordance with national, regional and international rulebooks. Information about the authors, scientific publications and their relations are already stored in CRIS system, so there is no need to create the system again from the scratch. Since the University of Novi Sad has a system for storing data on scientific research publications named CRIS UNS, it is possible to carry out an extension of the existing system and create subsystem which would enable defining rulebooks used in the process of evaluation of researcher's results.

II. RELATED WORK

There are many ways in which researcher can contribute to the development of science. He can actively participate in the management of research institutions/organizations, be an organisational/programme committee member in the realization of scientific events (conferences, workshops, seminars...), advisor of PhD thesis, member of scientific project, author/editor of publication etc. In practice, the evaluation mainly relays on publication (promotion of new scientific ideas). Publications can be evaluated individually (prone to subjective decisions) or as a part of larger publication. Evaluating the source (e.g. journal, conference) in which the paper is published is practical and objective, thus many evaluations rulebooks are based on this principle.

Three characteristic approaches can be distinguished in the evaluation process [4]:

- Expert group (commission) which evaluates the results based on the defined rules.
- Usage of bibliometric indicators (impact factor, h-index, citations ...).
- Combination of the previous two, for example expert group which in its final score, also can take into account the value of bibliometric indicators.

The author of the paper [5] made a comprehensive analysis by comparing the first and the second approach and revealed their positive and negative sides. A combination of an expert group and the bibliometric indicators is probably the best approach for evaluation of research results.

CERIF [6] is a standard that describes physical data model and XML messages formats enabling interoperability between CRIS research management systems. Model CERIF [7] provides a representation and exchange of a various types of scientific research data, and can be expanded and adapted to different needs. Papers [8], [9] describe an extension which uses a formalized Dublin Core to describe the published researcher results.

In the paper [10] the CERIF compatible data model based on the MARC 21 format is presented. An example for extension of the CERIF model for the needs of storage of the bibliometric indicators is presented in [11]. A CERIF-based schema for encoding of research impact in a structured way is stated in [12]. The paper [13] describes the CERIF data model extension which was created in...
order to satisfy the requirements of IST World portal (ist-world.dfki.de). Examples of CERIF based CRIS systems are: SICRIS(sicris.izum.si), Cristin (www.cristin.no), HunCRIS (nkri.info.omikk.bme.hu/HunCRIS_eng.htm), Pure system of Royal Holloway (pure.rhul.ac.uk/portal), IST Wotl portal, etc. Usage of CRIS system for the evaluation of researcher's results in Slovenia is described in [14]. Authors of [15] present an approach how a journal can be evaluated based on bibliometric indicators and the CERIF data model. Furthermore, in [16] an example of CRIS service for journals and journal articles evaluation is given.

In the paper [17] a CERIF compatible research management system CRIS UNS is presented, which can be accessed at http://www.cris.uns.ac.rs. Currently, the system stores over 10500 records of scientific publications (papers published in journals, conference proceedings, monographs, technical solutions and patents etc.). CRIS UNS system is under development since 2008 at the University of Novi Sad in the Republic of Serbia. Former development of that system covered implementation of the system for entering metadata about scientific research results [18]. Later phases in the development of CRIS UNS system included integration of various extensions that relay on CERIF model. Extension of CERIF that incorporates a set of metadata required for storing theses and dissertations in the University of Novi Sad is defined in [19]. In [20] a CRIS search profile based on SRU/W standard was created that enables a unified and semantically rich search of stored records. The paper [21] propose a CERIF data model extension which is used as a basis for future development of the CRIS UNS module for evaluation of scientific research results. System for modelling evaluation rulebooks which is a subject of this paper strongly relays on that model. The model itself was changed in a certain manner and it will be explained in details further in the paper.

III. EVALUATION ACCORDING TO THE SERBIAN RULEBOOK

The process of evaluation of scientific research is prescribed by law in the Republic of Serbia. Evaluation process is based on RuleBook on Procedure and Aspects of Evaluation and Quantitative Expression of Scientific Research Results (www.mnp.gov.rs/images/content/nauka/pravna_akta/PRAVILNIK_o_zvanjima.pdf) which in 2008 was regulated by the Ministry of Education, Science and Technological Development. The Ministry is using the previously mentioned Rulebook for funding scientific programs in Serbia. Also, many scientific institutions use the same rulebook for election of researchers into appropriate teaching and science positions.

At the University of Novi Sad a modified version of the previously mentioned Rulebook of the Ministry is applied. Researchers at the University are obligated to independently evaluate their own performance and to submit lists of their scientific results that are categorized according to the stated rulebook. In that manner, possible outcome may be that the same publication is evaluated differently by different co-authors of the paper. Problem with the different categorizing of the same results can be explained by stating that the researchers are not familiar with the rulebook for evaluation and that they often lack complete information which is necessary for the evaluation process. The previous problem can be overcome by using information support systems for evaluation that aims to institutionalize this process. The main idea of using information system is that researchers enter data about their scientific-research results by themselves and that some commission (group of experts) evaluates those results according to some rulebook. Creation of the mentioned information system should allow the evaluation of scientific results by different rulebooks.

Serbian rulebook set down the list of researcher's results (entity types) that are subjected to evaluation. Entity types that can be evaluated by Serbian rulebook are publications (journals, conference proceeding, monographs, paper published listed publications and thesis and dissertations), patents and technical solutions. A part of Serbian rulebook relays on principle that evaluation of individual publication can be accomplished trough evaluation of a publication source. So, a paper published in journal, conference proceeding or monograph is evaluated according to evaluation value of journal, conference or monograph (results source). The Rulebook prescribes the classification of researcher's results into various types (results' types with appropriate codes) which are organised in two hierarchical levels. That classification is used in evaluation of researcher's results and results sources. Hereinafter, classification for the entity types journal and paper published in journal will be explained. According to rulebook, journal (result source) can initially be classified into two main types of results: International journals (code: M20) and National journals (code: M50). Classification on lower level of hierarchy will be explained for category International journal. International journal is further subdivided into four types of results: Leading journal of international importance (code: M21), Outstanding journal of international importance (code: M22), Journal of international importance (code: M22), Specially verified international journal (code: M24). All researcher's results are categorised regarding the researcher's role in foundation of that scientific-research result. Currently, there are two distinct roles: Author and Editor (applicable to publications) of results. If we observe journal classified as M22, the authority of the paper in that journal is categorised as result type Paper published in the outstanding international journal (code: M22). Editorial of the journal of M22, shall be categorised as Editorial of outstanding international journal (code: M28).

In the Serbian rulebook quantitative measure is prescribed for every type of result in compliance to group of sciences to which the researcher belongs. There are three science groups: Mathematics and Natural Sciences (SG1), Technical and Technological Sciences (SG2), Social Sciences (SG2). Table 1 presents a part of hierarchy of types of results with Quantitative measures.

Serbian rulebook by its structure is not a unique example of rulebook for evaluating researcher's data. In the following text some similar rulebooks will be presented. Ministry of Civil Affairs of Bosnia and Herzegovina in 2012 presented a rulebook which coincides to the large extent with the Serbian rulebook. Values of entities: entity types, results' types and science groups of Bosnia and Herzegovina rulebook can be mapped to Serbian rulebook. Values of codes for results' types are different but quantitative measures are alike. Montenegro has similar rulebook as Serbia, with the
differences that there are no codes for results' types and there is no science groups for which researcher's results can get quantitative value (only one general science group). Also, Montenegro's rulebook includes two additional researcher roles: mentor and committee member. Ministry of Science, Education and Sports of Croatia in 2013 presented a rulebook that sets down additional values for researcher's roles (committee member, coordinator, collaborator...) and entity types (projects, plenary lectures...). Science groups in that rulebook are more refined, with the total number of seven. Also, every science group has its own set of results' types with their distinct codes.

IV. SYSTEM FOR MODELLING RULEBOOKS BASED ON THE CERIF DATA MODEL EXTENSION

The main purpose of the research presented in this paper was to develop an extension for modelling rulebooks in the existing scientific-research information system CRIS UNS. Data model for representing rulebooks was developed as an extension of the CERIF model. Relaying on CERIF will enable potential interoperability and simple implementation of proposed extension within all those systems that support CERIF model. Implementation of that model will be explained in the Section (Software architecture and implementation).

Data model and system architecture are represented by UML (Unified Modelling Language) diagrams. The CASE tool PowerDesigner was used for modelling. Data model is represented with physical diagram, while system’s architecture is represented by deployment and component diagrams.

A. Requirements specification

Based on the analysis of the Rulebook of the Republic of Serbia, requirements of the Evaluation Commissions and the authors' experiences in the implementation of information systems, a list of demands that must be met in the system implementation is stated. The list is established to support the definition of all rulebooks which are very similar to the rulebook of Republic of Serbia in their structure:

- Enable defining of the basic rulebook information (code, name, short description, start and end date of rulebooks' application).
- Enable linking of the physical document that contains the original text of the Rulebook to its electronic representation in the information system.
- Enabling hierarchical linking of rulebooks. Hierarchical linking should be provided regarding the fact that a number of scientific institutions are basing their rulebooks on the rulebook of the Ministry.
- Create a classification of types of results and allow defining their hierarchical linking.
- Separate the results' types classification from the rulebook and organize the classification in a manner that they belong to a certain results' group. This is justified by the fact that a large number of rulebooks does not have its own classification of scientific-research results, so the classification is adopted (identical or modified) from other rulebooks. The advantage of this organization is to define classification of results' types only once (sometimes over 100 results' types may be defined), and to use that classification for all different rulebooks which support it (by selecting a particular results' group, the rulebook shall take classification from that group).
- Create classifications of scientific disciplines/groups as independent units in the information system (regardless the Rulebook).
- Enable defining of quantitative values for results' types within a specific scientific discipline/group.
- Define researchers' role classification and the classification of entity types as independent units of the information systems and allow linking of the rulebooks to the appropriate roles (creation of lists for the researchers' roles which are supported by this rulebook) and the entity types (creation of lists for entity types which are supported within the rulebook in focus).
- Enable defining of sets of allowed values of results' types for the supported entity types.
- Provide mapping of different results' types regarding the role of researchers and types of entities. The mapping should also take into account the possibility of determining the results' type in accordance to the source of the scientific results (for publication).
- Supporting multilingualism for all entities that are prescribed in the Rulebook.

B. Expanding the CERIF model for the needs of storing data about rulebooks

The initial idea of modelling a rulebook as an extension of the CERIF model, is taken from the paper [21]. An
expansion of the CERIF model is defined, in a manner to match the specification requirements defined in the previous section. In addition, the model parts will be explained in more details:

- The entity `RuleBook` holds the basic information about the rulebook. Multilingual input values for the name and a brief description of the rulebook are supported with the entities `RuleBookName` and `RuleBookDescr`.

- CERIF entity `cfMedium` is used to present the digital document that contains the original text of the Rulebook. Binding the rulebook with the digital document is done via the `RuleBook_Medium` entity.

- CERIF semantic layer has been used to define the disciplines / groups, the researchers' roles, entity types that are subject of evaluation, results' types and results' groups. For each of these entities a scheme is defined (an instance of the entity `cfClassScheme`) and classes which represent the corresponding values of the mentioned entities (entity instance `cfClass`). The link between the group and the results' types which belong to the group and also the hierarchical organization of the results are achieved by using CERIF entity `RuleBook_Class`. Multilingual input values for the classification is achieved by using CERIF entity `cfTerm`.

- `RuleBook Class` entity is defined for the purposes of classification of the rulebooks. Linking of the rulebooks with the certain results' group is achieved by this entity.

- Entities `RuleBook_ResultsType`, `RuleBook_EntityType` and `RuleBook_ResearchersRole` will contain information on the classification of types of results, entity types and the researchers' role in the frame of the rulebook (creation of list's classifications which are supported by the Rulebook).

- A list of quantitative measures for various types of results in different science groups is defined through entity `ResultsTypeMeasure`.

- `EntityResultsType` entity shall hold all permitted combinations of allowed values of results' types for the supported entity types.

- Determination of the results' types classification in relation to the researchers' role is provided by the entity `ResultsTypeMapping`. Classification is determined regarding: the researcher's role in the forming of the scientific-research results (e.g. author, editor ..), type of entity (e.g. journal, paper), source entity type (e.g. journal), classification for the observed entity type (e.g. paper published in the leading journal of international importance, outstanding paper published in the journal of international importance...) and the observed source entity type (leading journal of international importance, outstanding journal of international importance...).

The extension of the CERIF model for defining the rulebook within CRIS research management system, is presented at Figure 1.

![Figure 1. Physical diagram of extended CERIF model.](image)

Diagram contains entities depicted with different border style. Thin line is used to indicate the original components inherited from the CERIF data model, while the bolded line components are those which are added for needs of the proposed model and do not exist in the CERIF data model.

C. Software architecture and implementation

The application’s architecture is a multi-layer client-server, based on a set of open-source software components, mostly written in `Java` programming language. CRIS UNS system for entering metadata about scientific research results uses MARC 21 format [10] to store some metadata information. The system for modelling rulebooks is integrated within the existing CRIS UNS system for management of scientific research data by extending some existing components described in [18]. UML deployment and component diagram of the system are shown respectively on Figure 2 and Figure 3. Thin line is used to indicate the original components of the CRIS UNS system, while the dashed line is used to indicate the modified components.

**Client-Web Browser.** A standard web browser can be used as a client application. The communication between client and server-side is done via HTTP.

**Application Server** component describes server side of application. Interface component was modified by creating new `xhtml` (web pages) and `ManagedBean` instances (Java classes which control execution of application) for the purpose of enabling browse, create, update, view and delete function of all entities necessary for the modelling of the rulebooks. Text server was
changed to include rules for searching and retrieving metadata of rulebooks. In component DTO&MARC21, new Java classes (Data Transfer Objects - DTOs, object representations of a MARC 21 records, converters DTO to/from MARC21) that support representation of rulebooks and accompanying classifications (sciences groups, results' groups, results' types, ...) are created. Database manger component was modified to include Java classes that enable CRUD (create, read, update, delete) operations for extension of CERIF model are added.

**Database-MySQL DBMS.** Database was altered by creating new tables which will store rulebook data.

From this point further, the implemented system related to entering data about rulebooks is described in more details. The part of system for storing classification data for sciences groups, results' groups and results' types will also be presented. The other parts of the implemented system, which are related to researcher's roles and entity types are similar to the part associated with sciences groups.

Figure 4 presents the interface for managing science groups, where all records are presented in the table. User can enter a new record by clicking on the button , delete the selected record by clicking , view record detail information or edit record data . Clicking the button opens the form for adding a science group (Figure 5). The user is obligated to enter unique classification code and name and to define the main language of classification textual attributes. Optional values are description, dates and translations. The same form but in different operation mode is reused for actions view and edit record data (create/read/update form - CRU form). All values for records are entered in Serbian script.

The interface for managing results' groups is similar to the interface of science groups. The CRU form of science group differs from the other CRU classification forms, because it contains an addition for managing results' types. Management of results' types is only possible through the management of results' groups. To depict hierarchical organisation of results' types all records are represented in the tree interface component.

Management of rulebook can be viewed in Figure 6. Rulebooks that are created on the basis of other rulebook are presented as his child elements. A part of CRU form for entering basic data is shown on Figure 7. The user is obligated to enter unique code, name, dates, results' group and to define the main language of textual attributes. Optional values are description, translations and attaching digital document containing the rulebook text. Principle for managing all advance data can be shown on example for quantitative measures for various types of results in different science groups (Figure 8). Panel Quantitative measures of Results types contains button Adding new entities, Save changes in the table and a table. Columns of the table represent the entities (different classification) which are combined in the creation of the table row. Every row has an action button remove . Form for adding new records in the table is designed to contain combination of elements that represent columns in the table.
Efficient evaluation of the scientific-research data is only possible when evaluation process is relaying on an information system. The process of evaluation of researcher’s results is usually in accordance with some rulebook. So, it is essential to enable modelling of rulebooks inside the information system.

The article presents a system for storing and managing rulebooks for evaluation of researcher’s scientific-research results. The system is based on the extension of CERIF data model, and it is integrated in the existing CRIS UNS system at a crossroads, “in Proceedings of the 6th International Conference on Current Research Information Systems, University of Kassel, Kassel, 2002, pp. 33–44.

Figure 6. Interface for managing rulebooks.

Figure 7. Part of CRU rulebook form for entering basic data

Figure 8. Example of rulebook’s advanced data

I. CONCLUSION

ACKNOWLEDGMENT

Results presented in this paper are part of the research conducted within the Grant No. III-47003, Ministry of Science and Technological Development of the Republic of Serbia.

REFERENCES


