Towards Culture-sensitive Extensions of CRISs: Gender-based Researcher Evaluation

Miloš Savić, Mirjana Ivanović, Miloš Radovanović, and Bojana Dimić Surla

University of Novi Sad, Faculty of Sciences,
Department of Mathematics and Informatics
Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia
{svc,mira,radacha,bdimic}@dmi.uns.ac.rs

Abstract. Current research information systems (CRISs) offer great opportunities for extraction of useful and actionable knowledge based on various data analysis techniques. However, many of these opportunities have not been explored in depth, especially in culture-sensitive areas such as gender-based evaluation of researchers. In this paper, we present GERBER, a methodology and accompanying tool for performing gender-based analysis of CRIS data. The tool enables the extraction of co-authorship networks, computation of various author metrics, and statistical comparison of male and female researchers. Functionality of GERBER is demonstrated on data extracted from the CRIS of the University of Novi Sad (UNS). We also present a plan to integrate GERBER into CRIS UNS in order to facilitate continuous gender-based researcher evaluation. Experiences obtained during such integration will enable us to propose more general methodological guidelines and APIs for culture-sensitive extensions of CRIS systems and standards.

Keywords: research information systems, culture-sensitive extensions, gender-based evaluation

1 Introduction

Measuring gender differences in research and teaching productivity has been a topic of interest for researchers and decision makers for many decades [17]. The basic generalization found in the literature is that male faculty outperform female faculty [5, 8] (for a more comprehensive list, please see [17] and references therein). However, more recent studies have shown that this is by no means due to some inherent superiority of one gender over the other. Xie and Shauman [34] conducted four large, nationally representative, cross-sectional surveys spanning several decades, observing that differences in research productivity declined over the period 1969–1993. Similarly, Gander [17] concludes that, although analysis of data at first glance supports the generalization, further scrutiny reveals that patterns of employment and distribution of funding are actually the main causes of the observed differences in productivity. Once the analysis is adjusted, it is revealed that female faculty have significant research productivity [17].

Current research information systems (CRISs) contain scientific production data and thus offer great opportunities for research evaluation. One such information system for storing and managing data about scientific research activity at the University of Novi Sad (UNS), Serbia – CRIS UNS was developed following the recommendations of the non-profit organisation euroCRIS [12]. CRIS UNS provides a comprehensive list of publications of researchers affiliated to UNS and enables automated evaluation of individual UNS researchers and institutions. In this paper we present GERBER (*GEndeR-Based Evaluation of Researchers*) – a methodology and an accompanying tool to perform gender-based analysis of data exported from CRIS UNS. The main functionalities provided by the tool are: (1) extraction of a co-authorship network that represents collaborations between authors contained in exported data, (2) computation of various metrics that reflect productivity, competency, collaborativity and social importance of individual researchers and (3) statistical comparison of male and female researchers based on non-parametric statistical tests. To demonstrate applicability of the tool we performed gender-based evaluation of researchers employed at UNS-PMF – University of Novi Sad Faculty of Sciences.

The rest of the paper is structured as follows. The overview of related research works is given in Section 2. Section 3 outlines the data preparation methods for gender-based analysis of researchers. The following section describes the proposed methodology for gender-based researcher evaluation and the accompanying GERBER tool. The analysis conducted using GERBER is presented in Section 5. Integration of analytical tools into the CRIS UNS system is discussed in Section 6. In the last section we give the conclusions and directions for future work.

2 Related Work

In this section we will review relevant related work comprising of studies of gender differences in scientific productivity (Section 2.1) and give the background of the CRIS UNS information system for storing and managing data about scientific research (Section 2.2).

2.1 Scientometric Analysis of Gender in Research

The comprehensive multi-decade study by Xie and Shauman [34], besides observing the decline of gender differences in research productivity, also correlates gender productivity differences with gender differences in personal characteristics, structural positions, and marital status, implying that gender differences in research productivity stem from gender differences in structural locations, and as such respond to the secular improvement of women's position in science.

Contemporary data in the field of social science shows that not only did gender differences disappear in the younger generations of researchers, but that if some differences exist, it is the young female researchers that outperform their male peers [2]. In educational psychology, on the other hand, although females are gaining ground in terms of primary and secondary article authorship and journal editorial board membership, this increase does not keep pace with the male-female ratio in organizational memberships [15]. Also, in the industrial and organizational psychology there are significant gender differences with respect to publication output and career courses [21].

In the domains of science, engineering and technology, the trends are also varying, with the overall impression that participation and performance of women improved in recent times. Within Spanish natural resources and chemistry scientists, no significant differences in productivity were found between genders within professional categories, but the outliers with the highest production were for the most part male [6]. In nano science and technology, female researchers are scarce in number, but perform equally in terms of scientific production and impact [31]. A comprehensive study involving the DBLP database of computer science publications ranging from 1936 to 2010 also indicates a low percentage of women in this field, albeit a steadily rising one [7]. The authors observed that men publish more than women, but attributed this to the fact that the average research life of men is longer.

Studies of gender differences in scientific productivity also produce varying results in different countries. In Croatia [29], within the studied young research population females are somewhat less productive than males which is in line with the observations by Xie and Shauman [34] discussed earlier. In Italy [1], there is also evidence of higher overall male productivity, but with difference smaller than reported in a large part of the literature, confirming an ongoing tendency towards decline. Russia [28], on the other hand, still exhibits strong gender disparity, which can also be said for Turkish social sciences [27].

The above constitutes only a small representative sample of studies dealing with gender differences in scientific production and teaching. A comprehensive meta-study of scientific literature on women in science and higher education, considering almost 1500 articles, shows continued growing interest in the topic, featuring more than 3000 authors, 67 countries, and 86 research areas [9].

2.2 CRIS UNS

The starting point in developing CRIS UNS was creating a well-structred and comprehensive metadata set for describing scientific results, as well as researchers and institutions. Paper [19] proposed the metadata model based on the MARC 21 library standard and compatible with CERIF (Common European Research Information Format). The CERIF data model provides a very rich and well-structured set of metadata. The core of CERIF are three basic entities Person, Project and OrganisationUnit, and three result entities ResultPatent, ResultPublication and ResultProduct [13]. The structure of CERIF enabled the development of the information system in which authors are uniquely identified and connected to their results, institutions and projects. The metadata set for some entities was further enriched by introducing the MARC 21 format of bibliographic data for presenting publications and the MARC 21 format of authority data for presenting authors [32]. The described model was the basis for developing the information system CRIS UNS for tracking research activity at the University of Novi Sad [25].

Speaking of research activity, one of the main purposes for developing the CRIS UNS system was providing automated evaluation of scientific results, researchers and institutions, which has become extremely important. Paper [20] proposed an extension of CERIF by data for evaluation of published scientific results. The extension is based on the CERIF semantic layer that enables classification of entities and their relationships by different classification schemas. The rules for evaluation proposed by the academic regulatory bodies were implemented within CRIS UNS and exposed as a service for evaluation of scientific results.

The architecture of CRIS UNS with its rich matadata model and module for evaluation of scientific results provided the environment for developing various techniques and tools for business analysis and gathering important information used by institutional management. These tools included a module for creating periodical reports on research activity and different analyses.

3 Preparing Data for Gender Analysis

Generally speaking, there were two options for preparing data for gender analysis: the first one was exporting all relevant data to a format suitable for analysis tools, and the second one was providing the interface for retrieving relevant data directly from the CRIS UNS database. In this paper we opted for the first option, and discuss the second option in Section 6.

The architecture of the CRIS UNS and adoption of international standards for presenting research data provided an appropriate environment for interoperability with other systems [18] including ontology-based integration [10].

We used the existing interoperability architecture of CRIS UNS for developing a module for exporting publications and authors to XML documents. Exporting publications and researchers consisted of several steps:

1. selecting researchers for the given institution,
2. for each researcher, select all her/his publications,
3. if the result is not already processed, i.e. stored in the resulting XML file, store the result metadata depending on its type (journal paper, conference paper, monograph, paper monograph, etc.),
4. for each publication included in the previous step, select all authors,
5. store each author to the authors XML file if she/he is not already there.

4 GERBER Tool

GERBER is a standalone tool implemented in Java that performs gender-based analysis of data exported from the CRIS UNS system. The tool consists of three modules: Data Loader, Author Metric and Gender Analyzer. The architecture of GERBER is shown in Figure 1.

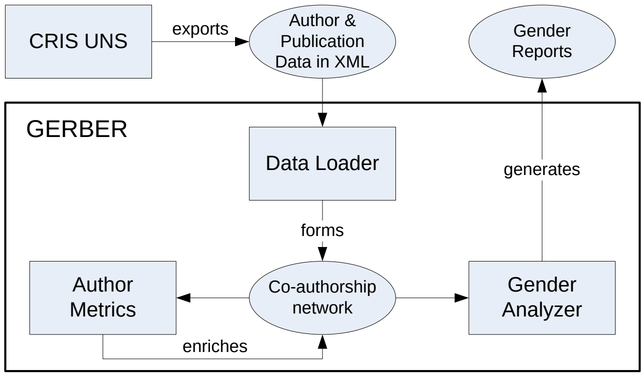

Fig. 1. The architecture of GERBER.

4.1 Data Loader

The Data Loader (DL) module parses two XML files that contain the data exported from the CRIS UNS information system and forms the co-authorship network of researchers appearing in the data. The first XML file contains demographic information about all UNS-PMF researchers and their direct external collaborators (researchers not affiliated to UNS-PMF). Each author is described by an XML element which includes unique author identifier, author name, date of birth, institution to which the author is affiliated, organizational unit within the institution, academic rank, and gender. The second XML file contains the data about publications which are authored by UNS-PMF researchers. Each publication is described by an XML element which consists of the following information: unique publication identifier, the complete list of author identifiers, publication year, title, publication type (journal, conference, monograph, etc.), information about publication venue and the categorization of the publication according to the rule book prescribed by the Serbian Ministry of Education, Science and Technological Development.

Co-authorship networks are undirected graphs showing collaboration between researchers [30]. The nodes of a co-authorship networks represent different researchers, and two researchers are connected by a undirected link if they co-authored at least one publication together. A co-authorship network can be viewed as an undirected, weighted and attributed graph $G = (V, E)$ where V is a set of attributed nodes and E is a set of weighted links. Namely, nodes have attributes that describe different demographic characteristics of authors. Also, author metrics are stored as node attributes. The strength of collaboration between two researchers is commonly quantified by one of three different weighting schema: the normal weighting scheme (the number of joint publications) [3], Newman's weighting scheme [26] which takes into account the number of authors per publication, and Salton's weighting scheme [23] which is a normalized variant of the normal scheme.

The extraction of a co-authorship network from CRIS-based data is a straightforward task since each author has an unique identifier which is used in publication records, and consequently there are no name disambiguation problems. The DL module forms the co-authorship network in three stages:

1. In the first stage the set of nodes is formed. The DL module iterates through the list of author XML elements and for each element creates one attributed node in the network.
2. The second stage forms the set of links and a decentralized inverted index which maps authors to their publications. Namely, each node in the network contains the list of publications written by the corresponding author. The DL module iterates through the list of publication XML elements and performs the following:
   - Connects each two authors of publication p by an undirected link.
   - For each author a of publication p, adds p in the list of publications a authored.
3. The third stage determines weights of links according to the three previously mentioned weighting schemes. Each link in the network is visited and for authors connected by the link weights are computed considering their lists of publications.

4.2 Author Metrics

The Author Metrics (AM) module enriches nodes of the co-authorship network formed by the Data Loader module with metrics that reflect author productivity, collaborativity, social importance and characteristics of ego networks. Table 1 shows the complete list of metrics computed by the AM module.

Table 1. The list of author metrics computed by the AM module.

Metric	Abbreviation	Metric category
Productivity, normal count	PRO-N	Productivity
Productivity, fractional count	PRO-F	Productivity
Productivity, straight count	PRO-S	Productivity
Serbian Research Competency Index	CI	Productivity
Degree centrality	DEG	Collaborativity
Local degree centrality	LDEG	Collaborativity
External degree centrality	EDEG	Collaborativity
Betweenness centrality	BET	Social importance
Closeness centrality	CLO	Social importance
Clustering coefficient	CC	Cohesiveness
Co-author gender disbalance	CGD	Gender diversity

The AM module implements three commonly used schemes to evaluate the productivity of researchers which are known as normal counting, fractional (adjusted) counting, and straight counting [22]. Let p be a publication written by n

researchers. In the normal counting scheme all authors of p receive equal credit for p, exactly one point. The straight counting scheme gives all the credit (one point) only to the first author of p. The fractional counting procedure assigns credit equal to $1/n$ to each of n authors of p. The AM module also computes the Serbian research competency index according to the categorizations prescribed by the rule book of the Serbian Ministry of Education, Science and Technological Development.

The collaborativity of a researcher can be quantified by its degree centrality in the co-authorship network. The degree centrality of author a is equal to the number of links incident to a. Each author in CRIS-exported data can be classified either as local or external. Namely, local authors are researchers affiliated to institution(s) covered by CRIS. In our case, local authors are researchers employed at UNS-PMF, while external authors are their collaborators from other institutions. Consequently, we can derive two other local centrality measures: local degree centrality (the number of local co-authors) and external degree centrality (the number of external co-authors).

The AM module also implements global metrics of social importance suitable for undirected graphs: betweenness centrality [16], and closeness centrality [4]. The betweenness centrality of node (author) a is the extent to which a is located on the shortest paths connecting two arbitrary selected nodes in the network. If a large fraction of shortest paths contain a, then a can be viewed as an important node of the network in sense that it has a vital role to the overall connectivity of the network. If the network has a clustered or community organization, then nodes with high betweenness centrality tend to be located at the intersections of communities, which means that they connect together different cohesive, homophilic social groups. The closeness centrality of node a is inversely proportional to the cumulative distance between a to other nodes in the network. Nodes with high closeness centrality can be considered as socially important since they are in proximity to a large number of other nodes.

The last category of author metrics computed by the AM module are metrics related to characteristics of ego-networks. The ego network of node a in undirected graph G, denoted by $Ego(a)$, is a sub-graph of G induced by i and its nearest neighbors. The cohesiveness of ego-networks can be quantified by the clustering coefficient [33]. The clustering coefficient of node a, denoted by $CC(a)$, is the probability that two randomly selected neighbors of a are directly connected. If $CC(a) = 1$ then neighbors of a forms the most cohesive ego-network – clique. The lowest value of $CC(a)$ is equal to 0 and happens when co-authors of a have never collaborated among themselves. The last metric computed by the AM module quantifies the gender structure of ego-networks. Let $M(a)$ and $F(a)$ denote the fraction of male and female collaborators of author a. Then, the co-author gender imbalance of a is defined as $CGD(a) = |M(a) - F(a)|$. If $CGD(a) = 1$ then all collaborators of a have the same gender, while $CGD(a) = 0$ implies that a equally collaborates with male and female researchers. It is important to emphasize that metrics of social importance and metrics related to ego-networks are computed on the reduced co-authorship network that encompasses only local researchers as nodes in order to have a clear institutional boundary when interpreting those metrics.

4.3 Gender Analyzer

The Gender Analyzer (GA) module performs statistical comparison of male and female researchers considering author metrics computed by the AM module. This module performs two non-parametric statistical tests which compare metric characteristics of male and female researchers. Implemented tests are the Mann-Whitney U (MWU) test [24] and the two independent samples Kolmogorov-Smirnov (KS) test [14].

Let M be an arbitrarily selected author metric. The MWU test is used to check the null hypothesis that concrete values of M for male researchers do not tend to be systematically higher or lower than concrete values of M for female researchers. The null hypothesis is rejected if obtained $p-$value is smaller than 0.05 and in such cases we can conclude that there is a statistically significant difference between male and female researchers regarding the aspect quantified by M. To quantify the effect size of the difference we use two probabilities of superiority [11]:

1. PS_m – the probability that for a randomly selected male researcher the value of M is strictly higher than the value of M for a randomly selected female researcher.
2. PS_f – which is the opposite probability of superiority, i.e. the probability that for a randomly selected female researcher the value of M is strictly higher than the value of M for a randomly selected male researcher.

The KS test checks the null hypothesis that cumulative distributions of M for male and female researchers are not significantly different. The test relies on the maximal vertical distance between two empirically observed distributions (the D statistic). The null hypothesis is rejected if the obtained $p-$value is smaller than 0.05.

As its final output the GA module makes three types of reports:

- basic gender statistics considering organizational units covered by the CRIS system (different departments at our faculty),
- results of non-parametric statistical tests, and
- tables that contain values of the Spearman correlation coefficient between different author metrics considering male and female researchers separately.

5 Gender-based Evaluation of UNS-PMF Researchers

To demonstrate the applicability of GERBER in a real-world scenario we performed gender-based evaluation of researchers affiliated to our institution (UNS-PMF) using real data exported from our official CRIS UNS information system. The exported data covers 423 researchers employed at UNS-PMF and their

15097 publications written in collaboration with 5267 researchers not affiliated to UNS-PMF. The co-authorship network extracted from the exported data contains 34111 links, where 2859 links (8.38%) represent collaborations between UNS-PMF researchers.

Table 2 shows basic gender statistics per organizational units (departments) of UNS-PMF. A majority of UNS-PMF researchers are female (60.76% of the total number). Female researchers are in a strong majority at the Department of Biology and Ecology and the Department of Chemistry. The smallest gender gap can be observed at the Department of Mathematics and Informatics where male and female researchers are almost equally represented.

Table 2. Basic gender statistics of UNS-PMF researchers. R denotes the absolute number of researchers, while M and F are percentages of male and female researchers, respectively.

Department	R	M[%]	F[%]
Mathematics and Informatics	87	49.43	50.57
Geography	66	57.58	42.42
Biology and Ecology	118	25.42	74.58
Physics	57	56.14	43.86
Chemistry	95	24.21	75.79
Total	423	39.24	60.76

The results of statistical tests performed by the GA module of GERBER are summarized in Table 3. Although UNS-PMF male researchers on average have slightly higher values of all productivity metrics compared to UNS-PMF female researchers, application of non-parametric statistical tests revealed that there are no statistically significant gender differences regarding scientific productivity. Also, we noticed that there are strong positive Spearman correlations between different productivity metrics for both genders – the lowest value of Spearman correlations for a randomly selected pair of productivity metrics is 0.83.

Regarding social aspects of scientific collaboration, it can be observed that UNS-PMF male researchers do not tend to have more total/local/external collaborators than UNS-PMF female researchers, and vice versa. Centrality metrics also exhibit strong Spearman correlations to productivity metrics (see Table 4). Moreover, the external degree centrality stronger correlates to productivity metrics compared to the local degree centrality which means that external collaborations have a stronger impact to productivity of both UNS-PMF male and female researchers.

From the data presented in Table 3 it can be observed that the null hypothesis of both non-parametric statistical tests were rejected for the betweenness centrality metric, but not for the closeness centrality metric. This means that there are statistically significant differences between UNS-PMF male and fe-

Table 3. The results of statistical comparison of UNS-PMF male and female researchers. $\langle M \rangle$ and $\langle F \rangle$ denote the average values of corresponding metric for male and female researchers, respectively. U is the value of the Mann-Whitney test statistic, MWU-p denotes the p-value of the MWU test, PS_m and PS_f are male and female probabilities of superiority, respectively. D is the value of the Kolmogorov-Smirnov test statistic and KS-p denotes the p-value of the KS test. Bold p values indicate that the null hypothesis of the test is rejected.

Metric	$\langle M \rangle$	$\langle F \rangle$	U	MWU-p	PS_m	PS_f	D	KS-p
PRO-N	82.72	65.40	20741	0.63	0.51	0.48	0.07	0.71
PRO-F	25.52	17.85	19975.5	0.27	0.53	0.47	0.10	0.29
PRO-S	25.83	19.33	20754	0.64	0.50	0.47	0.09	0.40
CI	132.05	101.40	20056	0.30	0.53	0.47	0.10	0.23
DEG	51.24	44.98	20440	0.47	0.52	0.47	0.06	0.78
LDEG	13.91	13.26	21194	0.91	0.49	0.48	0.05	0.95
EDEG	37.33	31.72	20301.5	0.40	0.52	0.47	0.07	0.66
BET	579.38	389.05	18564.5	**0.02**	0.55	0.42	0.14	**0.04**
CLO	0.30	0.31	18666	0.23	0.46	0.53	0.12	0.10
CC	0.49	0.55	18545	**0.02**	0.42	0.55	0.13	**0.04**
CGD	0.39	0.41	19004	0.35	0.45	0.51	0.12	0.10

Table 4. The values of the Spearman correlation coefficients between collaborativity and productivity metrics.

	Male researchers		Female researchers	
	PRO-N	CI	PRO-N	CI
DEG	0.86	0.79	0.87	0.82
LDEG	0.66	0.59	0.7	0.64
EDEG	0.88	0.82	0.88	0.82

male researchers considering their social importance. Namely, UNS-PMF male researchers do not tend to be dominant in the core of the co-authorship network, but they more frequently appear as bridges that connect different, highly cohesive research groups. Statistically significant gender differences can be also observed for the clustering coefficient: ego-networks of UNS-PMF female researchers tend to be slightly more cohesive than ego-networks of UNS-PMF male researchers. This suggests that UNS-PMF female researchers tend to stimulate their unconnected collaborators to work together more often compared to UNS-PMF male researchers. Having in mind that UNS-PMF male researchers tend to have higher betweenness centrality we can conclude the following:

- UNS-PMF male researchers tend to be more important for the cohesion of our institution at the macro scale – more often they connect different research groups, but

– UNS-PMF female researchers tend to be more important for the cohesion of our institution at the micro scale – more often they connect researchers from the same research group which previously have not collaborated.

6 Integration of Analytical Tools into CRIS

The main subject in the future work in this area is integration of GERBER and similar analytical tools with CRIS UNS. This can be done by applying a service oriented architecture in which CRIS UNS exposes services to the analytical tools. These services will include operations for obtaining relevant data that are inputs for analysis.

Instead of dividing the analysis in two independent stages, (1) export of all relevant data to XML documents and (2) loading data from these documents, the better solution is to load data directly from the CRIS UNS database through appropriate services. Some benefits of the proposed approach are:

– retrieving the updated states of the entities (as CRIS UNS is in constant use by authors who enter their publications, the latest state can be obtained only by real-time access),
– obtaining only data that are relevant for the analysis, and omitting redundant elements that will influence memory usage,
– exposing the complete CRIS UNS metadata set for analysis, including the attributes that are not recognised as relevant in time of export in the current solution.

As for the concrete technology and implementation of the services there are basically two solutions: implementing the web service with WSDL and SOAP and exchanging XML documents, or using the REST architectural style and exchanging either XML documents or JSON objects. Although there are some advantages of SOAP-based services in terms of tools support and type safety, REST services become very popular these days mostly due to ease of implementation based on the HTTP protocol.

The suggested architecture is shown in Figure 2. GERBER (or some other analytical tool) accesses CRIS UNS data through the REST API. The complete set of operations remains to be determined but some of the operations are shown in the figure. These are the following operations: (1) retrieve all researchers for the given institution, (2) retrieve all publications for the given researcher, and (3) retrieve all publications entered after the given date. The third operation will improve the efficiency of an analytical tool such as GERBER because instead of loading all publications, we can load only those entered after the date of last access.

The architecture presented in Figure 2 with a precisely defined set of operations is applicable for any analytical tool and any research information system. In other words, if we define the complete set of operations for the REST API, and extend GERBER to load data through that API, any other research information system that implements the API can use GERBER for gender analysis.

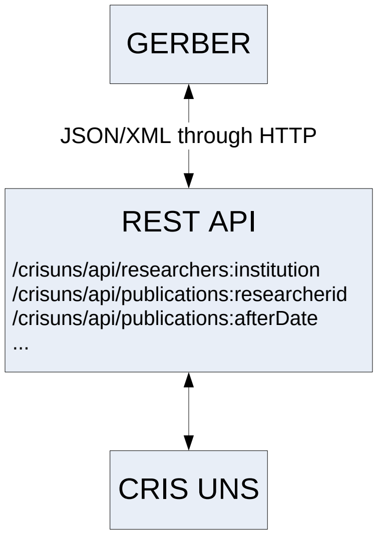

Fig. 2. GEBER and CRIS UNS integration through the REST API.

7 Conclusions and Future Work

As the main contribution of the paper we described the methodology and accompanying tool (GERBER) for gender-based evaluation of researchers. Using GERBER we analyzed gender differences for researchers employed at our faculty (UNS-PMF) relying on data exported from our official CRIS UNS system. The analysis showed that there are no statistically significant gender differences considering productivity and collaborativity of UNS-PMF researchers. On the other hand, gender differences can be observed with respect to the role of researchers in the process of improvement of institutional cohesion.

In future work we plan to improve GERBER in two ways: (1) by adding a new class of author metrics that consider weights of links in co-authorship networks, and (2) by including other culture-sensitive attributes such as age and academic rank in gender-based evaluation of researchers.

Finally, we plan to fully integrate GERBER into CRIS UNS as an analytic service. Full integration of GERBER into CRIS UNS will enable continuous gender-based evaluation of researchers employed at the University of Novi Sad. On the basis of the previously mentioned integration we will be in the position to propose general methodological and technical guidelines (APIs) for culture-sensitive extensions of CRISs.

Acknowledgments. The authors thank the Ministry of Education, Science and Technological Development of the Republic of Serbia for support through project no. OI174023, "Intelligent techniques and their integration into wide-spectrum decision support," and for additional support in cooperation with the Slovenian Research Agency through bilateral project no. 451-03-3095/2014-09/43, "Culture-sensitive aspects in data technologies."

References

1. Abramo, G., D'Angelo, C.A., Caprasecca, A.: Gender differences in research productivity: A bibliometric analysis of the Italian academic system. Scientometrics 79(3), 517–539 (2009)
2. van Arensbergen, P., van der Weijden, I., van den Besselaar, P.: Gender differences in scientific productivity: A persisting phenomenon? Scientometrics 93(3), 857–868 (2012)
3. Batagelj, V., Cerinšek, M.: On bibliographic networks. Scientometrics 96(3), 845–864 (2013)
4. Bavelas, A.: Communication patterns in Task-Oriented Groups. The Journal of the Acoustical Society of America 22(6), 725–730 (1950)
5. Bayer, A.E.: Teaching faculty in academe: 1972–73. ACE Research Reports 8, American Council on Education, Washington, DC, USA (1973)
6. Bordons, M., Morillo, F., Fernández, M.T., Gómez, I.: One step further in the production of bibliometric indicators at the micro level: Differences by gender and professional category of scientists. Scientometrics 57(2), 159–173 (2003)
7. Cavero, J.M., Vela, B., Cáceres, P., Cuesta, C., Sierra-Alonso, A.: The evolution of female authorship in computing research. Scientometrics 103(1), 85–100 (2015)
8. Cole, J.R.: Fair Science: Women in the Scientific Community. The Free Press, New York, USA (1979)
9. Dehdarirad, T., Villarroya, A., Barrios, M.: Research on women in science and higher education: A bibliometric analysis. Scientometrics 103(3), 795–812 (2015)
10. Dimić Surla, B., Segedinac, M., Ivanović, D.: A BIBO ontology extension for evaluation of scientific research results. In: Proceedings of the Fifth Balkan Conference in Informatics. pp. 275–278. ACM (2012)
11. Erceg-Hurn, D.M., Mirosevich, V.M.: Modern robust statistical methods: an easy way to maximize the accuracy and power of your research. The American Psychologist 63(7), 591–601 (2008)
12. euroCRIS: euroCRIS Current Resesarch Information Systems, `http://www.eurocris.org/`
13. euroCRIS: CERIF 1.3 Full Data Model (FDM) Introduction and Specification (2012), `http://eurocris.org/Uploads/Web%20pages/CERIF-1.3/Specifications/CERIF1.3_FDM.pdf`
14. Feller, W.: On the Kolmogorov-Smirnov limit theorems for empirical distributions. The Annals of Mathematical Statistics 19(2), 177–189 (1948)
15. Fong, C.J., Yoo, J.H., Jones, S.J., Torres, L.G., Decker, M.L.: Trends in female authorships, editorial board memberships, and editorships in educational psychology journals from 2003 to 2008. Educational Psychology Review 21(3), 267–277 (2009)
16. Freeman, L.C.: A set of measures of centrality based on betweenness. Sociometry 40, 35–41 (1977)
17. Gander, J.P.: Faculty gender effects on academic research and teaching. Research in Higher Education 40(2), 171–184 (1999)
18. Ivanović, D., Ivanović, L., Dimić Surla, B.: Multi-interoperable CRIS Repository. Procedia Computer Science 33, 86–91 (2014)

19. Ivanovic, D., Surla, D., Konjovic, Z.: CERIF compatible data model based on MARC 21 format. The Electronic Library 29(1), 52–70 (2011)
20. Ivanović, D., Surla, D., Racković, M.: Journal evaluation based on bibliometric indicators and the CERIF data model. Computer Science and Information Systems 9(2), 791–811 (2012)
21. König, C.J., Fell, C.B., Kellnhofer, L., Schui, G.: Are there gender differences among researchers from industrial/organizational psychology? Scientometrics 105(3), 1931–1952 (2015)
22. Lindsey, D.: Production and citation measures in the sociology of science: The problem of multiple authorship. Social Studies of Science 10(2), 145–162 (1980)
23. Lu, H., Feng, Y.: A measure of authors centrality in co-authorship networks based on the distribution of collaborative relationships. Scientometrics 81(2), 499–511 (2009)
24. Mann, H.B., Whitney, D.R.: On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics 18(1), 50–60 (1947)
25. Milosavljevic, G., Ivanovic, D., Surla, D., Milosavljevic, B.: Automated construction of the user interface for a CERIF-compliant research management system. The Electronic Library 29(5), 565–588 (2011)
26. Newman, M.E.J.: Who is the best connected scientist? A study of scientific coauthorship networks. In: Ben-Naim, E., Frauenfelder, H., Toroczkai, Z. (eds.) Complex Networks, Lecture Notes in Physics, vol. 650, pp. 337–370. Springer Berlin Heidelberg (2004)
27. Ozel, B., Kretschmer, H., Kretschmer, T.: Co-authorship pair distribution patterns by gender. Scientometrics 98(1), 703–723 (2014)
28. Paul-Hus, A., Bouvier, R.L., Ni, C., Sugimoto, C.R., Pislyakov, V., Larivière, V.: Forty years of gender disparities in Russian science: A historical bibliometric analysis. Scientometrics 102(2), 1541–1553 (2015)
29. Prpić, K.: Gender and productivity differentials in science. Scientometrics 55(1), 27–58 (2002)
30. Savić, M., Ivanović, M., Radovanović, M., Ognjanović, Z., Pejović, A., Jakšić Krüger, T.: The structure and evolution of scientific collaboration in Serbian mathematical journals. Scientometrics 101(3), 1805–1830 (2014)
31. Sotudeh, H., Khoshian, N.: Gender differences in science: The case of scientific productivity in Nano Science & Technology during 2005–2007. Scientometrics 98(1), 457–472 (2014)
32. The Library of Congress: MARC Standards, `http://www.loc.gov/marc/`
33. Watts, D.J., Strogatz, S.H.: Collective dynamics of "small-world" networks. Nature 393, 440–442 (1998)
34. Xie, Y., Shauman, K.A.: Sex differences in research productivity: New evidence about an old puzzle. American Sociological Review pp. 847–870 (1998)