Malaysian Journal of Library & Information Science, Vol. 29, no. 2, August, 2024:93-106

Automated gender identification vs. manual review: Gender assignment of editorial board members in Information and Library Science Journals

Yiming Liu[1,2*], Adolfo Alonso-Arroyo[1,2], Rafael Aleixandre-Benavent[1,3], and Juan-Carlos Valderrama-Zurián[1,2]

[1]Unidad de Investigación e Información Social y Sanitaria (Grupo UISYS), Universitat de València, València, Unidad asociada al Instituto Interuniversitario de Investigación Avanzada sobre Evaluación de la Ciencia y la Universidad (INAECU) UC3M-UAM, SPAIN

[2]Departament d'Historia de la Ciència i Documentació. Facultat de Medicina i Odontologia, Universitat de València, SPAIN

[3]Ingenio (CSIC-Universitat Politècnica de València), SPAIN

e-mail: Yiming.Liu@uv.es* (corresponding author); Adolfo.Alonso@uv.es; Rafael.Aleixandre@uv.es; Juan.Valderrama@uv.es

ORCID ID: Y.Liu: 0009-0006-3592-0175
A.Alonso-Arroyo: 0000-0002-5084-2818
R.Aleixandre-Benavent: 0000-0002-6678-8844
J.C. Valderrama-Zurián: 0000-0001-5787-6853

ABSTRACT

In recent years, computer programmes that automatically assign gender based on a person's name and country of affiliation have been increasingly used in gender studies of authors of scientific publications. This study aims to compare the results generated by the automated genderize.io programme with those obtained through manual gender identification. To assess the accuracy of genderize.io, the gender of Editorial Board Members (EBMs) from 84 journals in the field of Information and Library Science was analysed. The comparison revealed discrepancies: genderize.io incorrectly classified 80 out of 1,419 men as women, and 124 out of 2,580 women were misidentified as men. Additionally, genderize.io classified the gender of 123 EBMs as unknown. While the manual method achieved a 99.15 percent accuracy rate, genderize.io had a slightly lower accuracy of 91.51 percent. There was, however, strong agreement between the two methodologies (Cohen's Kappa = 0.829, p < 0.001). Genderize.io exhibited a 7.71 percent inaccuracy rate, particularly underestimating the number of women. The study concludes that while automated software like genderize.io is effective for large-scale analyses and useful for library and information professionals, manual review is recommended for smaller studies to ensure higher accuracy.

Keywords: Gender assignment; Automated gender identification; genderize.io*;* Editorial board members; Library and Information Science

INTRODUCTION

In recent years, gender studies on authors of scientific publications have increasingly relied on automated gender assignment programmes that use individuals' first and last names for identification. Some programmes enhance accuracy by incorporating additional data, such as country of origin. Among the most commonly used tools are *Gender-API*, *NamSor*, *genderize.io*, *Wiki-Gendersort*, *Face++*, *SSA*, *IPUMS* and *Sexmachine* (Sebo, 2021a; 2021b; 2022; Wais, 2016; Karimi et al., 2016). However, several limitations in using these programmes have been identified. For instance, when analysing Chinese names in Pinyin format, the accuracy of correct gender assignments is often low (Sebo, 2022). Such errors are also prevalent in the assignment of Asian names in general (Jadidi et al., 2017). Another issue arises when assigning country codes to authors based on their institutional affiliation. Authors may be affiliated with institutions in countries different from the country of origin associated with their name, leading to misidentification or failure to assign gender (Sebo, 2021a). Additionally, some programmes, such as *genderize.io*, require manual revision to remove diacritical marks, special characters, and second first names when retrieving data from bibliographic databases (Sebo, 2021b).

Among the various gender assignment programmes, *genderize.io* has emerged as a popular choice, as evidenced by its presence in the scientific literature. Between 2016 and the end of 2022, at least 34 works indexed in the Web of Science and 41 in Scopus have utilized *genderize.io*. This software, which can assign gender to 114,541,298 name variants across 242 countries or regions, has been applied to professionals in various disciplines, including Information Science & Library Science (Sebo, 2021a; 2021b), Sociology (Heiberger, 2022), Engineering (Narasimhan et al., 2016; Hartzler et al., 2021; Lund & Shamsi, 2023), and Medicine (Gadek et al., 2023; Schlein et al., 2022; Gottlieb et al., 2021; Nguyen et al., 2021). These studies have either compared the performance of gender detection tools or validated the results through manual Internet searches.

This study addresses a critical gap in research on automated gender classification tools, particularly their accuracy compared to manual methods within academic publishing, specifically in editorial contexts. While *genderize.io* is increasingly employed to assign gender based on names, limited research has assessed its reliability, especially in fields such as Library and Information Science (LIS). Existing studies using automated tools often neglect potential inaccuracies in gender identification, which can distort data and lead to misguided policy decisions aimed at addressing gender disparities in editorial board composition.

The primary objective of this study is to compare the data obtained from the genderize.io programme on the gender assignment of editorial board members of Information Science & Library Science journals with the data obtained manually. This comparison holds considerable importance as gender data is vital for analysing disparities in academic publishing. It is imperative to ensure that automated tools yield reliable data, as this is essential for research examining the gender balance among editorial board members, which can subsequently impact policies and practices within the realm of academic publishing. By guaranteeing the reliability of automated tools, researchers and institutions can more effectively monitor and tackle gender imbalances in editorial and academic positions.

Therefore, the primary objective of this study is to compare the gender assignment results of *genderize.io* with those obtained through manual verification for editorial board members in Information Science and Library Science journals. This comparison aims to assess the reliability of automated tools in accurately identifying gender within academic publishing. Given the crucial role of precise gender data in shaping policy decisions related to gender equity, this study underscores the need for caution when relying solely on automated tools. It is imperative to ensure that automated tools yield reliable data, as this is essential for research examining the gender balance among editorial board members, which can subsequently impact policies and practices within the realm of academic publishing.

METHOD

To assess the validity of the *genderize.io* programme, the study collected data on Editorial Board Members (EBMs) from 84 journals in the Information Science & Library Science category listed in the Journal Citation Reports (JCR) - Social Science Citation Index (SSCI), Web of Science 2020 edition. The EBMs' data were compiled into a Microsoft Excel spreadsheet in CSV format, including first names and institutional country codes. Prior to automatic gender assignment, the names were standardised by removing accents and diacritical marks and by retaining only the first name in cases of compound names.

The gender of the EBMs was then determined using the *genderize.io* program. In parallel, each name was manually reviewed by consulting various sources, including institutional or personal websites, curricula, and academic networks such as ResearchGate, Academia.edu, and Google Scholar. A flowchart illustrating the gender analysis process for the EBMs is provided in Figure 1. Appendix 1 provides a detailed list of the 84 journal titles, including the number and percentage of EBMs classified by gender using both automated and manual methods. It highlights the discrepancies between gender assignments from the *genderize.io* programme and manual verification, with shaded areas indicating the number of members (male, female, or unknown) whose gender classification differed between the two methods for each journal.

A descriptive analysis of the data was conducted, and Cohen's Kappa index was calculated to assess the level of agreement between the genders assigned by *genderize.io* and those determined manually. Additionally, error metrics from Wais (2016), including errorCoded, errorCodedWithoutNA, naCoded, and errorGenderBias, along with the weightedError metric from Santamaría & Mihaljević (2018), were applied. These metrics are recommended for reproducing and evaluating different gender analysis approaches to ensure optimal accuracy and reliability. The distribution of gender errors by country, according to the percentage assigned to each country, is also provided. In this case, "assigned percentage" refers the proportion of correct gender classifications provided by *genderize.io*, which is referred to as "accuracy" by the programme.

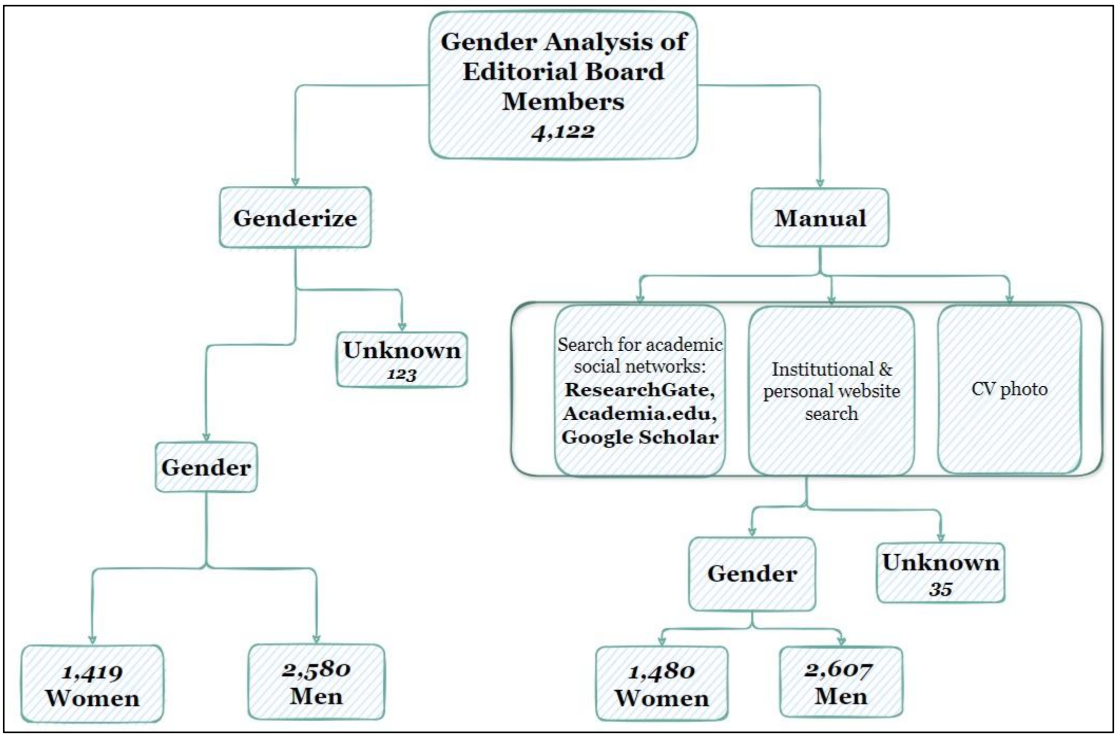

Figure 1: Flowchart of the Gender Analysis Process for Editorial Board Members

RESULTS

The total number of EBMs is 4,122. Comparing the results from *genderize.io* with those obtained manually, it is observed that 80 women (5.64%) out of 1,419 identified as female by *genderize.io* were incorrectly classified as male by the manual verification (see Table 1, highlighted in orange). On the other hand, 124 cases (4.81%) (highlighted in orange) out of the 2,580 defined by *genderize.io* as male, were female. Therefore, the match rate for female was 93.38 percent (1,325 correctly identified), while for males it was 94.84 percent (2,447 correctly identified). In 123 cases (2.98%) (Table 1), *genderize.io* assigned the gender as unknown. Of these, 99 cases (80.49%) were due to the program's inability to determine the gender, while 24 cases (19.51%) were unresolved due to incomplete information, either missing name details (n=16) or country information (n=8). Notably, 70 of the 99 unidentified cases (70.71%) were Asian names (predominantly from Chinese, Japanese, Korean, and Indian origins).

Table 1: Distribution of Women and Men Identified by *genderize.io* and Manual Identification

,

		Gender by *genderize.io*			**Total**
		Female	**Male**	**Unknown (G)**	
Gender by manual identification	**Female**	1,325 (93.37%)	124 (4.80%)	31 (25.20%)	1,480 (35.90%)
	Male	80 (5.63%)	2,447 (94.84%)	80 (65.04%)	2,607 (63.24%)
	Unknown (M)	14 (0.98%)	9 (0.34%)	12 (9.75%)	35 (0.84%)
Total (*genderize.io)*		1,419	2,580	123	4,122

Unknown (G): signatures to which *genderize.io* assigned "unknown" because it could not identify the gender.
Unknown (M): signatures to which manual identification could not assign gender.

Based on the country of affiliation, 51 Asian names matched their country of origin, while 19 names were associated with institutions in western countries (primarily the United Kingdom (UK), the United States (US), Australia, or Denmark) and were therefore not identified by the programme. During the manual review, it was not possible to assign gender to 35 EBMs (0.85% of the total).

The results indicate that manual review would improve accuracy by approximately 8 percentage points, achieving a 99.15% correct assignment rate (1,480 women and 2,607 men) compared to 91.51% accuracy from *genderize.io* (1,325 women and 2,447 men). The agreement between the two methods was high, with a Cohen's Kappa index of 0.829 ($p<0.001$), demonstrating excellent concordance.

In terms of quality metrics (Table 2), *genderize.io* shows a fraction inaccuracy of 7.71% (errorCoded) and a proportion of non-classified names of 2.72% (naCoded). When excluding "unknowns," the misclassification rate is 5.13% (errorCodeWithoutNA). The gender bias metric shows a -1.11% error, indicating a potential underestimation of female representation. The weighted error, with a weight of 0.2, is 5.66%.

Table 2: Quality Metrics for Gender Classification Using *Genderize.io*.

	errorCoded	**errorCodedWithoutNA**	**errorGenderBias**	**naCoded**	**weightedError$_{0.2}$**
genderize.io	0.0771	0.0513	-0.0111	0.0272	0.0566
Formula	$\frac{f_m+m_f+m_u+f_u}{m_m+f_m+m_f+f_f+m_u+f_u}$	$\frac{f_m+m_f}{m_m+f_m+m_f+f_f}$	$\frac{m_f-f_m}{m_m+f_m+m_f+f_f}$	$\frac{m_u+f_u}{m_m+f_m+m_f+f_f+m_u+f_u}$	$\frac{f_m+m_f+w*(m_u+f_u)}{m_m+f_m+m_f+f_f+w*(m_u+f_u)}$

f_m: females's misclassification; m_m: males correct classification; m_f: males' misclassification; f_f: females correct classification; m_u: males' non-classification; f_u: females' non-classification; w: 0.2, as in the methodology of Santamaría & Mihaljević (2018).

Therefore, despite an acceptable level of agreement with manual assignment (Kappa index of 0.829), the tool's accuracy in classifying women is somewhat lower than for men. This discrepancy could affect its reliability in studies where gender representation is a critical factor. Furthermore, the weighted error obtained, confirms the presence of a margin of error that could influence the overall accuracy of the gender analysis.

Table 3 presents the 204 errors among males and females assigned by *genderize.io* according to their country of affiliation, together with the percentage of corrected assignation according to the programme. Names affiliated to the US are the ones with the highest percentage of errors, with 85 cases, mainly because half of the names with the US code have Asian affiliation and the programme assigns them erroneously. Regarding the 36 errors from Chinese names, it has been observed that *genderize.io* has assigned names such as Adel, Dan and Juan, which were female, as male. Another common error that is incorrectly assigned by *genderize.io* occurs in names that can designate both men and women, such as Robin, Laurence and Michele. Even if *genderize.io* assigns it correctly according to the country code, the person may have another nationality. For example, the name "Andrea" is classified as male when affiliated with Italy but as female when associated with Spain. The error analysis reveals that *genderize.io* incorrectly assigned gender with 100% accuracy to 10 professionals (4.9%). The most frequent errors occurred when the programme's accuracy ranged between 50-59%.

Table 3: Error Analysis of *Genderize.io* Classification based on Country of Affiliation

Country / region Code of Affiliation																						% of correction by Genderize *	Number (% out of 204)
AU	CA	CN	DE	DK	FR	GB	IL	IN	IT	JP	KR	MY	NL	NZ	SE	SG	TH	TW	US	VN	ZA		
2	4	19	2			6				1						6		1	15	1		50%-59%	57 (27.94%)
2	3	5									2	1	1				1	1	22			60%-69%	38 (18.63%)
	4	1				3			1		1			1		2		1	17			70%-79%	31 (15.2%)
1	2	3	1	1		4	1	1		2	1				1	2			17			80%-89%	37 (18.14%)
2	2	6			1	7									1	2			9		1	90%-99%	31 (15.2%)
1		2				1										1			5			100%	10 (4.9%)
8	15	36	3	1	1	21	1	1	1	3	4	1	1	1	2	13	1	3	85	1	1		

AU-Australia, CA-Canada, CN-China, DE-Germany, DK-Denmark, FR-France, GB-United Kingdom, IL-Israel, IN-India, IT-Italy, JP-Japan, KR-South Korea, MY-Malaysia, NL- Netherlands, NZ-New Zealand, SE-Sweden, SG-Singapore, TH-Thailand, TW-Taiwan, US-United States, VN-Vietnam, ZA-South Africa

*Percentage that *genderize.io* assigns for the correctness of the gender. 100% means that the gender is completely correct for *genderize.io*.

DISCUSSION

As the volume of scientific publications grows, so does the need to analyse gender disparity and equality within various fields to ensure fair representation. Accurate identification of the gender of authors and tracking changes over time necessitates the use of automated programs that can assign gender to researchers based on their names. One of the most widely used programmes for this purpose is *genderize.io* (Sebo, 2021b; Wais, 2016; Nguyen, Robinson & Hoover, 2021; Waldhorn et al., 2022; Batumalai et al., 2023; Sixto-Costoya et al., 2022; Wang et al., 2022).

Genderize.io reports the probability that the gender assignment made to each name it analyses is correct. Therefore, studies using this programme to assign gender to individuals usually apply a minimum probability above which they consider that the assignment can be considered valid. For example, in a study by Waldhorn et al. (2022) on trends in women leadership of oncology clinical trials, a minimum probability of 60% was chosen - a threshold that was also applied by Nguyen et al. (2021) in another study investigating changes in the distribution of women as first authors in pharmacy practice journals. In a study analysing the gender of authors in the *Journal of Medical Imaging and Radiation Oncology* (Batumalai et al., 2023) and another examining the gender of editorial boards in veterinary sciences journals (Wang et al., 2022), thresholds exceeding 95% were used. However, in the present study, applying a threshold of 90% to 99% detected 31 errors in gender assignment, and a threshold of 100% identified 10 errors.

The quality metrics for *genderize.io* in this study yielded lower percentages compared to those reported by Santamaría & Mihaljević (2018) in terms of fraction inaccuracy, proportion of non-classified names, and weighted error. While *genderize.io* in this study tends to underestimate the number of females, Santamaría & Mihaljević's study found an underestimation of males. These discrepancies may be attributed to the differences in sample size analyzed between these two studies.

The existence of these errors, even when very high probability thresholds of correct assignment are applied, raises the need to manually (humanly) check the assignments made by the programme. Thus, comparing *genderize.io* assignments with manual checks has revealed that neutral names, which can be either masculine or feminine depending on the country, often cannot be accurately assigned a gender. For example, the name "Juan" can be a male Spanish name or a female Chinese name; the name "Jaime", which is a female Anglo-Saxon name, can also be a male Spanish name; and the name Dan, which is a female name in Chinese, is a male name in Anglo-Saxon. Conversely, determining the gender of names such as Angappa, Atreyi, Xuemeng, Zhaochun, Piyya, Thayanan, Mingwen, Bhuva, Rui, Jianwei, Xueqi, and Taemin from Asian cultures, as well as Anglo-Saxon names such as Laurence, Laurie, Robin, Jan, Kristin, Daniele, Fran, Adel, Neal, Ciaran, and Toni, is challenging due to their neutral nature.

To improve the accuracy of automated gender assignment tools, the following considerations could be taken into account. Tools should have comprehensive and representative datasets that include names from diverse cultures, genders and countries, taking into account minority language names and special script characters and accents used by some languages. In addition, the tools should incorporate as much additional information as possible to facilitate identification, such as institutional affiliations and areas of research. Moreover, they should incorporate machine learning algorithms that improve their performance as they process more data. It would also be desirable that errors detected in some studies and reported in publications could be taken into account in these tools, thus improving their performance. Moreover, given the availability of various programmes for automatic gender assignment, applying multiple tools simultaneously could provide a more comprehensive comparison of results. However, manual verification remains essential to ensure the highest accuracy in gender assignment. Additionally, current artificial intelligence systems can significantly aid in resolving ambiguous cases.

The limitation of this study includes the potential for the programme to incorrectly assign gender to rare or uncommon names due to insufficient data in its database. Additionally, the programme may not accurately detect ambiguity in certain names and lacks the capability to assign non-binary gender. Furthermore, conducting the manual search with a single researcher could have led to fewer accurate gender assignments compared to having two researchers perform the task.

CONCLUSIONS

The software genderize.io correctly assigns gender in 91.51% of cases, making it a useful tool for LIS professionals engaged in studies analysing the gender of large datasets with thousands of authors. This effectiveness is particularly valuable when manual annotation is impractical due to the dataset's size. Using *genderize.io* (or other commercial APIs) with a high threshold can be a good approximation for observing trends in authorship in terms of gender and country of origin. However, in smaller studies, in order to reduce as much as possible, the margin of error, it would be worthwhile to carry out a manual review recording to the method proposed in this study. This will make it possible to identify and correct the errors that the gender identification software is currently unable to support and to reduce the margin of error to below 1%. In conclusion, despite advances in automatic gender assignment, human review is necessary to correct possible errors and ensure accurate gender assignment. The inclusion of human review and continuous improvement of assignment algorithms of automated methods are important aspects to mitigate potential biases and limitations.

Incorrect gender assignment in scientific publications can have significant implications. It can lead to the invisibility of certain researchers, as misclassified authorships may result in their work being undervalued or overlooked, contributing to inequality. This misclassification may also discourage individuals from underrepresented genders from engaging in academic activities, exacerbating their invisibility and hindering their career advancement. Additionally, such errors can negatively impact an institution's reputation by presenting a skewed gender representation of its members.

The results of this study have important implications for gender studies, the field of scientometrics, and future LIS researchers. It facilitates more precise research on gender disparities by improving the accuracy of gender distribution data in academic output, collaboration patterns, and scientific impact across disciplines. Reliable gender data is crucial for enhancing the precision of research on gender diversity and equity in LIS, as inaccuracies can distort findings and affect interpretations of gender representation in scholarly communication and editorial practices. Additionally, these findings can guide LIS researchers in selecting appropriate gender classification tools and underscore the need for continuous improvement and validation of these tools. Recognising the limitations of automated tools may encourage the development of more comprehensive and transparent methods for assessing gender representation.

ACKNOWLEDGEMENT

Betlem Ortiz Campos. Beneficiary of the call for Technical Support Staff Grants. Ministry of Science and Innovation. State Research Agency. Co-financed by the European Union. PTA2021-019882-I for invaluable technical and documentary support.

CONFLICT OF INTERESTS

The authors declare that they have no competing interests, including no conflicts of interest

AUTHOR CONTRIBUTION

Conceptualization: [all authors], Methodology: [Y.Liu & A.Alonso-Arroyo], Formal analysis and investigation: [Y.Liu, R.Aleixandre-Benavent & J.C.Valderrama-Zurián], Writing - original draft preparation: [Y.Liu, A.Alonso-Arroyo & R.Aleixandre-Benavent]; Writing - review and editing: [all authors]

REFERENCES

Batumalai, V., Kumar, B., & Sundaresan, P. (2023). Trends in gender of first and senior authors of articles published in JMIRO. *Journal of Medical Imaging and Radiation Oncology*, *67*(2), 179-184. https://doi.org/10.1111/1754-9485.13492.

Gadek, L., Dammann, C., Savich, R., Mmuo-Oji, C., Barrera, L., Gallagher, P. G., & Machut, K. (2023). Gender analysis of Journal of Perinatology authorship during COVID-19. *Journal of Perinatology*, *43*, 518-522. https://doi.org/10.1038/s41372-022-01551-x.

Gottlieb, M., Krzyzaniak, S. M, Mannix, A., Parsons, M., Mody, S., Kalantari, A., Ashraf, H., & Chan, T. M. (2021). Sex distribution of editorial board members among emergency medicine journals. *Annals of Emergency Medicine*, *77*(1), 117-123. https://doi.org/10.1016/j.annemergmed.2020.03.027.

Hartzler, A. L., Leroy, G., Daurelle, B., Ochoa, M., Williamson, J., Cohen, D., & Stipelman, C. (2021). Comparison of women and men in biomedical informatics scientific dissemination: Retrospective observational case study of the AMIA Annual Symposium: 2017–2020. *Journal of the American Medical Informatics Association (JAMIA)*, *28*(9), 1928-1935. https://doi.org/10.1093/jamia/ocab097.

Heiberger, R. H. (2022). Applying machine learning in sociology: How to predict gender and reveal research preferences. *KZfSS - Cologne Journal of Sociology and Social Psychology*, *74*(1), 383-406. https://doi.org/10.1007/s11577-022-00839-2.

Jadidi, M., Karimi, F., Lietz, H., & Wagner, C. (2017). Gender disparities in science? Dropout, productivity, collaborations and success of male and female computer scientists. *Advances in Complex Systems*, *21*(03n04), 1750011. https://doi.org/10.1142/S0219525917500114.

Karimi, F., Wagner, C., Lemmerich, F., Jadidi, M., & Strohmaier, M. (2016). Inferring gender from names on the Web: A comparative evaluation of gender detection methods. *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 53-54). https://doi.org/10.1145/2872518.2889385.

Lund, B. D., & Shamsi, A. (2023). Women authorship in library and information science journals from 1981 to 2020: Is equitable representation being attained? *Journal of Information Science, 49*(5), 1335-1343. https://doi.org/10.1177/01655515211050026.

Narasimhan, V. L., Bhargavi, G. V., & Sultana, A. (2016). A subjective examination of implicit root stereotypes of STEM disciplines. In M. E. Auer, & K. S. Kim (Eds). *Engineering Education for a Smart Society. World Engineering Education Forum & Global Engineering Deans Council 2016* (pp. 225-236). Springer Cham.

Nguyen, E., Robinson, R., & Hoover, R. M. (2021). Women as first authors in key pharmacy journals: Analysis by publication type. *Journal of the American Pharmacists Association*, *61*(1), e26-e29. https://doi.org/10.1016/j.japh.2020.08.037.

Santamaría, L., & Mihaljević, H. (2018). Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science*, 4, e156. https://doi.org/10.7717/peerj-cs.156.

Sebo, P. (2021a). Performance of gender detection tools: A comparative study of name-to-gender inference services. *Journal of the Medical Library Association (JMLA)*, *109*(3), 414-421. https://doi.org/10.5195/jmla.2021.1185.

Sebo, P. (2021b). Using genderize.io to infer the gender of first names: How to improve the accuracy of the inference. *Journal of the Medical Library Association (JMLA)*, *109*(4), 609-612. https://doi.org/10.5195/jmla.2021.1252.

Sebo, P. (2022). How accurate are gender detection tools in predicting the gender for Chinese names? A study with 20,000 given names in Pinyin format. *Journal of the Medical Library Association (JMLA)*, *110*(2), 205-211. https://doi.org/10.5195/jmla.2022.1289.

Schlein, S. M., Pollock, N. W., Polukoff, N. E., Brown, A. B., Byrne, A., & Keyes, L. E. (2022). Gender equity in membership, leadership, and award recognition in the Wilderness Medical Society. *Wilderness & Environmental Medicine, 33*(3), 275-283. https://doi.org/10.1016/j.wem.2022.04.006.

Sixto-Costoya, A., Alonso-Arroyo, A., Castelló-Cogollo, L., Aleixandre-Benavent, R., & Valderrama-Zurián. J. C. (2022). Gender presence on the editorial boards of journals in the Women's Studies subject category. *Women's Studies International Forum*, 93, 102617. https://doi.org/10.1016/j.wsif.2022.102617.

Wais, K. (2016). Gender prediction methods based on first names with genderizeR. *The R Journal*, *8*(1), 17-37. https://doi.org/10.32614/RJ-2016-002.

Waldhorn, I., Dekel, A., Morozov, A., Alon, E. S., Stave, D., Tsrooya, N. B., Schlosser, S., Markel, G., Bomze, D., & Meirson, T. (2022). Trends in women's leadership of oncology clinical trials. *Frontiers in Oncology*, *12*, 885275. https://doi.org/10.3389/fonc.2022.885275.

Wang, A., Dunlop, R., Allavena, R., & Palmieri, C. (2022). Gender representation on journal editorial boards in the field of veterinary sciences. *Research in Veterinary Science*, *148*, 21-26. https://doi.org/10.1016/j.rvsc.2022.05.001.

APPENDIX

Appendix 1. Library and Information Science Journals, Editorial Board Members and Classification by Gender Using *genderize.io* and Manual Assignment

Journal Title	No of EBMs	*genderize.io*						Manual assignment					
		M*	% of M	F*	% of F	Unk*	% of Unk	M*	% of M	F*	% of F	Unk*	% of Unk
African Journal of Library Archives and Information Science	10	8	80.0	2	20.0	0		8	80.0	1	10.0	1	10.0
Aslib Journal of Information Management	28	14	50.0	13	46.4	1	3.6	13	46.4	14	50.0	1	3.6
Canadian Journal of Information and Library Science	25	13	52.0	12	48.0	0		13	52.0	11	44.0	1	4.0
College & Research Libraries	20	7	35.0	13	65.0	0		6	30.0	14	70.0	0	
Data Base for Advances in Information Systems	46	33	71.7	12	26.1	1	2.2	31	67.4	15	32.6	0	
Data Technologies and Applications	26	15	57.7	11	42.3	0		15	57.7	10	38.5	1	3.8
Electronic Library	36	16	44.4	18	50.0	2	5.6	17	47.2	18	50.0	1	2.8
Ethics and Information Technology	33	26	78.8	7	21.2	0		25	75.8	8	24.2	0	
European Journal of Information Systems	54	33	61.1	19	35.2	2	3.7	32	59.3	22	40.7	0	
Government Information Quarterly	62	43	69.4	16	25.8	3	4.8	47	75.8	15	24.2	0	
Health Information and Libraries Journal	28	10	35.7	18	64.3	0		7	25.0	21	75.0	0	
Informacao & Sociedade-Estudos	25	11	44.0	13	52.0	1	4.0	11	44.0	14	56.0	0	
Informacios Tarsadalom	15	14	93.3	0		1	6.7	14	93.3	0		1	6.7
Information & Culture	34	24	70.6	10	29.4	0		22	64.7	12	35.3	0	
Information & Management	146	107	73.3	29	19.9	10	6.8	115	78.8	30	20.5	1	0.7
Information and Organization	70	44	62.9	26	37.1	0		45	64.3	24	34.3	1	1.4
Information Development	18	14	77.8	4	22.2	0		13	72.2	4	22.2	1	5.6
Information Processing & Management	114	83	72.8	26	22.8	5	4.4	87	76.3	27	23.7	0	
Information Research-An International Electronic Journal	70	35	50.0	35	50.0	0		36	51.4	34	48.6	0	
Information Society	52	34	65.4	16	30.8	2	3.8	35	67.3	17	32.7	0	
Information Systems Journal	104	65	62.5	35	33.7	4	3.8	67	64.4	37	35.6	0	
Information Systems Research	72	52	72.2	16	22.2	4	5.6	56	77.8	16	22.2	0	
Information Technology & Management	53	46	86.8	7	13.2	0		46	86.8	7	13.2	0	
Information Technology & People	102	52	51.0	47	46.1	3	2.9	57	55.9	44	43.1	1	1.0

Journal Title	No of EBMs	*genderize.io*						Manual assignment					
		M*	% of M	F*	% of F	Unk*	% of Unk	M*	% of M	F*	% of F	Unk*	% of Unk
Information Technology and Libraries	13	7	53.8	6	46.2	0		7	53.8	6	46.2	0	
Information Technology for Development	60	34	56.7	23	38.3	3	5.0	37	61.7	23	38.3	0	
International Journal of Computer-Supported Collaborative Learning	66	35	53.0	29	43.9	2	3.0	34	51.5	32	48.5	0	
International Journal of Geographical Information Science	65	45	69.2	18	27.7	2	3.1	47	72.3	18	27.7	0	
International Journal of Information Management	209	145	69.4	55	26.3	9	4.3	153	73.2	54	25.8	2	1.0
Investigacion Bibliotecologica	5	3	60.0	2	40.0	0		3	60.0	2	40.0	0	
Journal of Academic Librarianship	16	7	43.8	8	50.0	1	6.3	7	43.8	9	56.3	0	
Journal of Computer-Mediated Communication	99	62	62.6	36	36.4	1	1.0	61	61.6	38	38.4	0	
Journal of Documentation	21	14	66.7	7	33.3	0		13	61.9	7	33.3	1	4.8
Journal of Enterprise Information Management	54	43	79.6	10	18.5	1	1.9	45	83.3	7	13.0	2	3.7
Journal of Global Information Management	92	62	67.4	25	27.2	5	5.4	68	73.9	21	22.8	3	3.3
Journal of Global Information Technology Management	81	59	72.8	20	24.7	2	2.5	60	74.1	21	25.9	0	
Journal of Health Communication	64	35	54.7	29	45.3	0		33	51.6	31	48.4	0	
Journal of Information Science	15	10	66.7	5	33.3	0		11	73.3	4	26.7	0	
Journal of Information Technology	70	45	64.3	24	34.3	1	1.4	48	68.6	22	31.4	0	
Journal of Informetrics	22	17	77.3	3	13.6	2	9.1	19	86.4	3	13.6	0	
Journal of Knowledge Management	138	104	75.4	32	23.2	2	1.4	108	78.3	30	21.7	0	
Journal of Librarianship and Information Science	35	14	40.0	18	51.4	3	8.6	14	40.0	20	57.1	1	2.9
Journal of Management Information Systems	70	57	81.4	11	15.7	2	2.9	58	82.9	12	17.1	0	
Journal of Organizational and End User Computing	86	59	68.6	21	24.4	6	7.0	64	74.4	21	24.4	1	1.2
Journal of Scholarly Publishing	8	7	87.5	1	12.5	0		6	75.0	1	12.5	1	12.5
Journal of Strategic Information Systems	80	50	62.5	29	36.3	1	1.3	52	65.0	28	35.0	0	
Journal of the American Medical Informatics Association	77	44	57.1	32	41.6	1	1.3	43	55.8	34	44.2	0	
Journal of the Association for Information Science and Technology	83	61	73.5	20	24.1	2	2.4	58	69.9	25	30.1	0	

Journal Title	No of EBMs	*genderize.io*						Manual assignment					
		M*	% of M	F*	% of F	Unk*	% of Unk	M*	% of M	F*	% of F	Unk*	% of Unk
Journal of the Association for Information Systems	121	81	66.9	38	31.4	2	1.7	83	68.6	38	31.4	0	
Journal of the Australian Library and Information Association	17	8	47.1	9	52.9	0		7	41.2	10	58.8	0	
Journal of The Medical Library Association	60	18	30.0	41	68.3	1	1.7	14	23.3	46	76.7	0	
Knowledge Management Research & Practice	37	30	81.1	5	13.5	2	5.4	30	81.1	7	18.9	0	
Knowledge Organization	27	12	44.4	13	48.1	2	7.4	14	51.9	13	48.1	0	
Learned Publishing	31	19	61.3	12	38.7	0		18	58.1	13	41.9	0	
Library & Information Science Research	19	6	31.6	13	68.4	0		4	21.1	15	78.9	0	
Library and Information Science	9	2	22.2	6	66.7	1	11.1	2	22.2	4	44.4	3	33.3
Library Collections Acquisitions & Technical Services	15	5	33.3	7	46.7	3	20.0	5	33.3	7	46.7	3	20.0
Library Hi Tech	71	49	69.0	21	29.6	1	1.4	46	64.8	23	32.4	2	2.8
Library Journal	22	7	31.8	14	63.6	1	4.5	6	27.3	16	72.7	0	
Library Quarterly	41	11	26.8	30	73.2	0		10	24.4	31	75.6	0	
Library Resources & Technical Services	12	1	8.3	10	83.3	1	8.3	1	8.3	10	83.3	1	8.3
Library Trends	13	7	53.8	6	46.2	0		7	53.8	6	46.2	0	
Libri - International Journal of Libraries and Information Studies	40	19	47.5	20	50.0	1	2.5	19	47.5	20	50.0	1	2.5
Malaysian Journal of Library & Information Science	9	5	55.6	2	22.2	2	22.2	4	44.4	5	55.6	0	
MIS Quarterly	73	50	68.5	17	23.3	6	8.2	52	71.2	21	28.8	0	
MIS Quarterly Executive	55	36	65.5	19	34.5	0		35	63.6	20	36.4	0	
Online Information Review	61	36	59.0	21	34.4	4	6.6	37	60.7	23	37.7	1	1.6
Portal-Libraries and The Academy	34	14	41.2	20	58.8	0		12	35.3	21	61.8	1	2.9
Profesional de la Información	60	49	81.7	10	16.7	1	1.7	50	83.3	10	16.7	0	
Qualitative Health Research	91	34	37.4	55	60.4	2	2.2	26	28.6	65	71.4	0	
Reference & User Services Quarterly	1	0		1	100.0	0		0		1	100.0	0	
Reference Services Review	29	13	44.8	16	55.2	0		12	41.4	17	58.6	0	
Research Evaluation	26	18	69.2	8	30.8	0		19	73.1	7	26.9	0	
Restaurator - International Journal for the Preservation of Library and Archival Material	13	5	38.5	8	61.5	0		5	38.5	7	53.8	1	7.7
Revista Española de Documentación Científica	24	13	54.2	11	45.8	0		13	54.2	11	45.8	0	
Scientist	37	12	32.4	25	67.6	0		11	29.7	26	70.3	0	
Scientometrics	89	67	75.3	20	22.5	2	2.2	65	73.0	24	27.0	0	

Journal Title	No of EBMs	*genderize.io*						Manual assignment					
		M*	% of M	F*	% of F	Unk*	% of Unk	M*	% of M	F*	% of F	Unk*	% of Unk
Serials Review	33	10	30.3	23	69.7	0		9	27.3	24	72.7	0	
Social Science Computer Review	26	18	69.2	7	26.9	1	3.8	17	65.4	9	34.6	0	
Social Science Information Sur Les Sciences Sociales	24	21	87.5	3	12.5	0		20	83.3	4	16.7	0	
Telecommunications Policy	59	48	81.4	10	16.9	1	1.7	49	83.1	10	16.9	0	
Telematics and Informatics	25	13	52.0	9	36.0	3	12.0	12	48.0	12	48.0	1	4.0
Transinformacao	33	18	54.5	14	42.4	1	3.0	19	57.6	14	42.4	0	
Zeitschrift fur Bibliothekswesen und Bibliographie	13	7	53.8	6	46.2	0		7	53.8	6	46.2	0	
Total	**4,122**	**2,580**	**62.6**	**1,419**	**34.4**	**123**	**3.0**	**2,607**	**63.2**	**1,480**	**35.9**	**35**	**0.8**

* M=Male; F=Female; Unk=Unknown