Analyzing gender disparities in STEAM: A Case Study from Bioinformatics Workshops in the University of Granada⋆

Elena Ruiz[1], Coral del Val[1], Rafael Alcalá[1], Alberto Fernández[1], Carlos Cano[1], Waldo Fajardo[1], and Jesús Alcalá Fdez[1]

Department of Computer Science and Artificial Intelligence, University of Granada, E-18071, Granada, Spain

Abstract. Bioinformatics is an interdisciplinary area that has raised a high interest for both academia and corporations in recent years. This rising area combines knowledge and skills from Bio and Science, Technology, Engineering, Arts and Mathematics (STEM) areas. One of the advantages of the synergy between these two work areas is that it offers an opportunity for closing the traditional STEM's gender gap. Despite this opportunity and the significance and wide application of bioinformatics field, this topic has still not gained enough visibility in the graduate programs for the Bio Bachelor Degrees at the University of Granada. This has motivated the organization of an annual "Educational Workshop on Bioinformatics" at the University of Granada by the Department of Computer Science and Artificial Intelligence. Results of the analysis of the first two editions of this workshop show a great interest on the topic by the university community at all levels (e.g. undergraduate and graduate students, teachers and researchers) without significant distinction among genders at global level. When analyzing student group, women did show a higher interest on the subject. However, this interest was not reflected in the higher university strata (teachers and researchers), which represents a glimpse of the spanish general current situation on the area.

Keywords: Bioinformatics · STEM · gender gap

1 Introduction

The Europe 2020 strategy [6, 7] set one of its targets in increasing participation in tertiary education with special emphasis in gender equality in Science, Technology, Engineering and Mathematics (STEM) skills. Unfortunately, despite more than two decades of initiatives, the under-representation of women in STEM careers is still a gap to close. The last published EU-28 statistics [10] show

---

that thought the number of women reached the 54.3% of all tertiary students, and 53.5% for Masters degrees, these numbers are extremely low in the field of STEMs careers were three quarters of the students are male [11]. Thus, there are still many initiatives working, at different levels and with different scopes, for gender balance in science, technology, engineering and mathematics [5, 14, 27, 13, 26, 8]. In the intersection between STEM and Bio and Health Degrees is where there is a higher growth of new careers, employment possibilities and higher salaries.

Bioinformatics is an interdisciplinary area that fits in this intersection, being STEM skills crucial in order to expand the use of biological, medical, behavioral or health data including their acquisition, storage, organization, archive, analysis, or visualization of such data [1, 24]. Bioinformatics is a relatively young field that brings together areas such as Biology, Biochemistry, Biotechnology, Medicine, Pharmacy mostly dominated by female students with areas such as Computer Science, Mathematics, Data Science, Physics; mostly dominated by male students [10]. It is interesting to notice that at the beginning of bioinformatics, when it was not called so, there were a large proportion of women involved. Among them one of the first true founders of Bioinformatics, the chemist Margaret Dayhoff [17]. However, recent works [3, 4] shows gender disparity in bioinformatics research publications authorships. This underrepresentation of women in high-quality research publications and conferences is also present in other scientific areas, even on those where women constitute over 50% of graduates and hold a great amount of faculty and research positions [2, 21, 12, 18]. Bioinformatics is reaching its peak [15, 19] bringing in the market a great number of job opportunities for scientists with computer science expertise and bio-knowledge [22], but the number of women accessing to them is still low. Unfortunately, there is a shortage of appropriate bioinformatics courses for undergraduate bio-students. In the European Community most of the courses are at Master level [20], which due to their specialization are not suitable for undergraduate students.

In accordance with these facts, the Dpt. of Computer Sciences teaching panel in the Bachelor Degrees of Biology, Biochemistry and Biotechnology at the University of Granada (UGR) (Spain) considered the necessity of developing an annual Educational Workshop on Bioinformatics. Even though it was originally intended for Bio Bachelor Degrees, we wanted to organize an open event. This way the workshop was publicized as an open free activity for all the Bachelor and Master students related to the Bio, Bio-Health and STEM areas. The main aim was to stress the importance and application of STEM skills knowledge for their future, and give visibility to careers and professional profiles in the field for individuals with their backgrounds. The initiative was focused on the knowledge transfer from experts in the field and potential employers. The final objective was to make clear the importance and applicability of the informatics to undergraduate students, and how this knowledge can influence their job opportunities. Students were provided with information for decision on the

building of their extra-curricular itinerary and they were encouraged to carry out research activities and company practices in this direction.

However, since these Bio majors present a higher enrollment of females (60% at the UGR) opposite to the STEM ones, we were also very interested to know if there was a gender imbalance in bioinformatics' interest at the tertiary educational level, as it seems to be universal across all aspects of the scientific itinerary. This work analyzes the opinions of the attendants to understand the current situation of bioinformatics teaching area and the interest of the different university collectives from a gender perspective. We evaluate both attendance data and evaluation tests results. From this analysis we can conclude there is a high interest by undergraduate and post-graduate students, interest shared by both females and males without significant differences. Nevertheless, the scenario was just the opposite when focusing on the higher university professionals (teachers and researchers), which represents a glimpse of the spanish general current situation on the area.

The remainder of this work is organized as follows. In Sec. 2 we present the material and methods used to set up the workshop. Then, Sec. 3 presents the analysis of results related to attendance, participants' background and knowledge acquisition from a gender perspective. A thorough discussion on the lessons learned throughout this work is carried out in Sec. 4. Finally, Sec. 5 summarizes and concludes the work.

2 Material and Methods

In this section we address several aspects related to the organization of the workshop as its program (Sec. 2.1), the structure of the quality and satisfaction survey designed by the organizers to assess the impact of the workshop (Sec. 2.2), and the test designed to assess the knowledge acquired by the participants during the workshop (Sec. 2.3).

2.1 Workshop Program

For the two first editions, the structure has been broadly the same. The workshop was two days long and consisted of plenary talks and a concluding round-table where all participants were given the opportunity to contribute and send their doubts to the speakers and organizers. The inscription was online and completely free. The schedule consisted of 5 consecutive hours each of the two days with a half-hour break. The first half-hour of the first day was devoted to the opening, followed by four plenary talks. The second day consisted of three additional plenary talks. All plenary talks comprised an hour of duration. Finally, during the last hour and a half of the workshop, a round-table was held. Among the speakers' profiles we could find members from both industry and academia. Their academic training included both computer science and biology, but also other degrees such as engineering, physics and medicine. While in the 1st edition the content of the lectures was mainly focused on genomics and epigenomics,

and the application of bioinformatics to personalized medicine; for the second edition the covered topics were extended to include, among other, bioinformatics applications in ecology or lectures about the usefulness of bioinformatics to get a job in the software industry.

2.2   Quality and Satisfaction Evaluation

Once the workshop was finished, all participants received an e-mail with a link to a satisfaction survey which fit in with Likert approach [16]. For the second edition the satisfaction survey was updated to include new questions that allow us to get a more accurate idea about participants' opinions. Additionally, the participants were asked to which universities collective they belonged, namely: undergraduate students, graduate students, professors, researchers, or others.

2.3   Participants' Evaluation

In both editions the workshop was recognized for the Biology Degree students with one ECTS credit under the condition of realizing a test to evaluate the knowledge acquired. In the second edition, Biotechnology Degree students were offered the same option. The test was formed, in both cases, by 10 questions directly related to the content of the plenary talks, each of the questions with four different options.

3   Results and Analysis

This section presents the analysis of the results of the two first editions of the workshop (2017 and 2018) including the analysis of the audience (Sec. 3.1 and 3.2) and of the results for the knowledge evaluation survey (Sec. 3.3).

3.1   Workshop Attendance

As mentioned in Sec. 2, the registration for the event took place online. Due to the fact that students could get a participation certificate, attendance to the sessions both days was controlled to keep record of the true audience. The analysis from the online registrations (371 for the first edition - 477 for the second one) showed that there were no significant differences in the percentages of males (47.7% - 48.6%) and females (52.3% - 51.4%) registered in the workshop. These numbers remains almost the same when looking to the attendance data: females (52% - 50%) and males (48.1% - 50%). It is worth mentioning, that because registration was free many people register themselves just in case they could attend. In 2017 a total of 234 people participate both days what makes a 63% of the registered people distributed equally between females (32.6%) and males (30.2%). In 2018 the total number of people participating both days rose to 274, i.e. a 57.5% of the registrations distributed once more equally between females (29.37%) and males (28.94%). This means a 17% increase in attendance

for the second edition of the workshop. Despite this increase in the number of attendants, the number of people who attended both days supposes a lower percentage of the total amount of registrations than in the first edition. The registration period was longer for the second edition than for the first one, then it makes sense that the number of people who registered themselves just in case they could attend but finally couldn't, were higher for this second edition. Moreover, during the first edition the 82.69% of the attendants to the session of the first day attended again the second one. For the second edition the datum is very similar, 83.03% of the attendants to the session of the first day attended again the second one.

The great interest of this topic among the UGR community has been evidenced by two facts: the great number of inscriptions and the high level of real attendance both days. In fact, it is very impressive the significant percentage of attendants to the first day that decided to participate the second independently of their gender. Furthermore, the increases in both inscriptions and attendants for the second edition lay bare that this interest is not going down. Most of the attendants were students who have to deal with very full class schedules and heavy workloads, having in many cases both morning and afternoon classes.

3.2 Participants' Background Diversity

Noteworthy, although the workshop was originally designed for undergraduate students in Biological Sciences Degrees, the interest arisen among other university collectives exceeded our expectations. When analyzing attendance data of the first workshop by group, students of all levels were clearly the major group of attendants (76%) with a slightly higher attendance of female students (56.83%) when compared to male students (43.17%). Performing the same analysis of the data from the second workshop we find that students remain being the main group of attendants (83.93%) equally distributed between females (49.29%) and males (50.71%). The remaining minority group of attendants was integrated by PDI (researchers and professors), and others (known gender, unknown role). In the PDI collective, which is an indicator of the current employment situation of women in leading roles at the University, we saw that this parity disappeared and turned into a significant imbalance of females (33.33% in the first workshop - 28.57% in the second workshop) against males (66.67% - 71.43%). Although these numbers can not be taken as representative due to the fact that the workshop was oriented to the students and not to the faculty (as shown in Fig. 1).

Moreover, when looking in detail into the distribution of students by type (see Fig. 2), we found around 67% of the participants to be undergraduate students and about 33% postgraduate students. These results show a much higher diversity of participants than expected, and corroborate the interest of the subject even in more highly-qualified participants, which constitute an important part of the total attendance.

To get a better picture of the interest in bioinformatics in the different educational stages by gender we looked at the gender distribution by degrees'
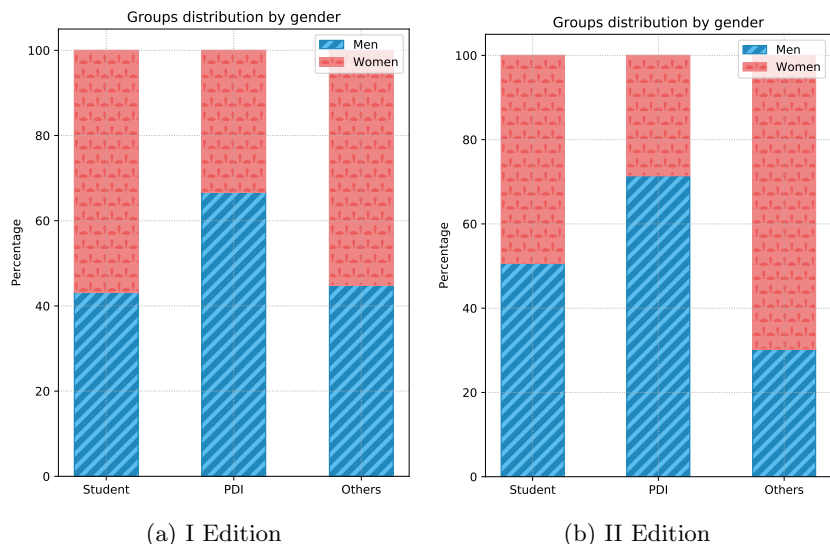
(a) I Edition   (b) II Edition

Fig. 1: Distribution of each group of participants (student, PDI, others) by gender
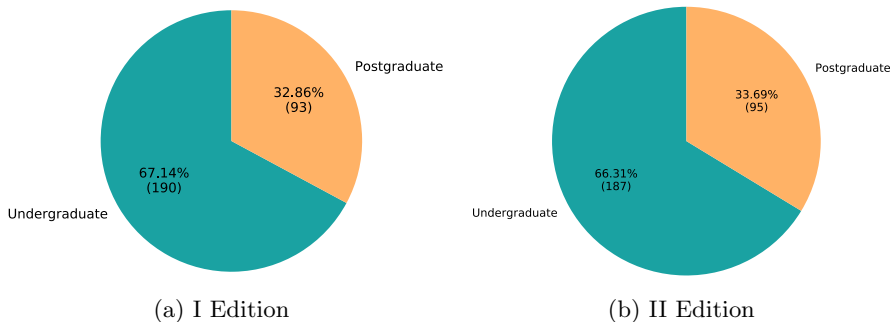

(a) I Edition   (b) II Edition

Fig. 2: Ratio of undergraduate and postgraduate participants that forms the student group

and postgraduate courses' knowledge areas. Figure 3 shows this type of distribution among participants who are coursing both undergraduate and postgraduate studies.

We observe that at the undergraduate level the majority of the students belong to Bio Degrees, representing more than 80% for both editions of the workshop, followed by STEM Degrees (between 9 and 14%) and Bio-Health Degrees (less than 5%). In the case of Bio Degrees we found that the attendance is almost equally distributed between genders but with a slight advantage of females (around 53%). This percentage difference might be only a reflex of the percent-

age of enrollments in these degrees, which round 55-60% females versus 45-40% males. This female advantage significantly increase when looking at Bio-Health Degrees students, specially in the first edition where the few undergraduate that attended were all females. In any case, since the number of Bio-Health undergraduate attendants are very reduced this results may be not so significant. Just the opposite is the case of undergraduate STEM Degrees where the percentage of female attendants (20% or less) is significantly lower than the percentage of males (80% minimum). Thought the numbers of STEM undergraduate students are also quite reduced it is pretty interesting that we do observed the typical gender imbalance expected from the enrollment imbalance in STEM Degrees at the University of Granada. The female enrollment in engineering degrees at the UGR is under 20%, being the minimum a 9.63% in the case of Computer Science [25]. The greater numbers of undergraduate students in the Bio Degrees in comparison to Bio-Health Degrees and STEM Degrees can be explained through the focus of the workshop more than on the interest aroused. This Workshop was mainly advertised in the Biology, Biochemistry and Biotechnology Degrees.

Figure 3 also contains analogous information for postgraduate participants. In the first edition the post-graduates background diversity was noticeably greater than in undergraduate students. While postgraduate students from Biosciences Masters were the majority (56%), we found a significant student participation increase compared to undergraduates from Bio-Health Masters such as Medicine (17%) and STEM Masters such as Computer Science (7%). This tendency is significant for the female sector in Medicine (20.2% vs. 8%) and STEM (4% vs. 3%) postgraduate degrees. Although not so clearly, this upward trend can also be appreciated in second edition data in both Bio-Health (4.81% of undergraduates and 11.58% of postgraduates) and STEM (13.3% of undergraduates and 15.8% of postgraduates). In the female sector Bio-Health goes from representing 6.74% of total female undergraduates to 17% of total female postgraduates. Results point that students from areas initially distant to biology consider bioinformatics as a meaningful field in their academic formation when their level of qualification increases.

These results encourage us to continue organizing activities to promote and integrate this discipline into the academic training of more and more students, opening their minds to new horizons both academic and labor. Helping to break the traditional gender stereotypes associated with the different disciplines that integrate bioinformatics, specially showing to Bio and Bio-Health Degrees female students the advantages of acquiring STEM skills but also showing to STEM students the interesting possibility of applying their knowledge to the Bio and Bio-Health fields.

3.3   Assessment of knowledge acquisition

As we have mentioned before, a test exam to assess the knowledge acquisition during the workshop had to be realized by the students to obtain the recognition of one ECTS credit. As it has been also mentioned, during the first workshop only Biology Degree students had the opportunity to get this ECTS credit but
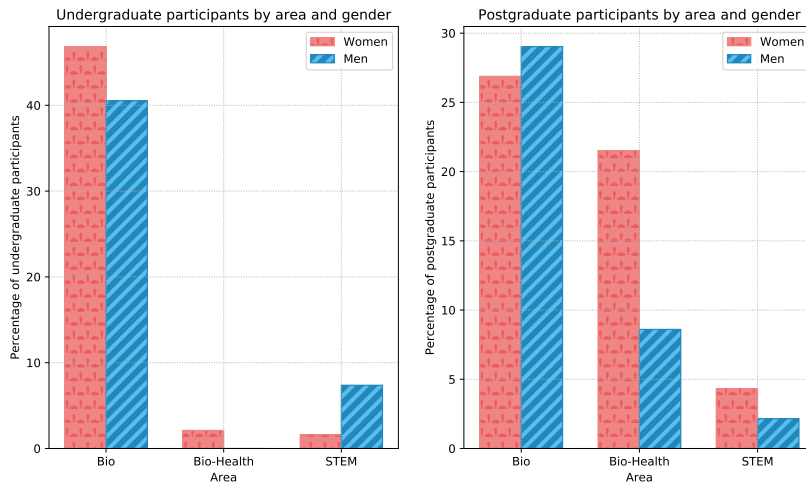
(a) I Edition
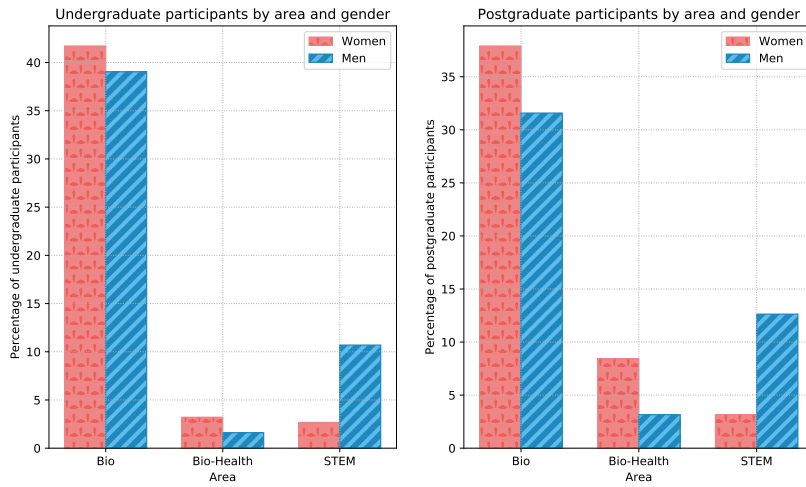
(b) II Edition

Fig. 3: Distribution of students who attended to the workshop by knowledge area of their degrees and gender

for the second workshop Biotechnology Degree students were also offered this option. A total of 64 undergraduate students realized the test the first year, 35 females and 29 males. This number rose to 91 undergraduate students the second year, 43 females and 48 males. The results show the same median for males and females both years. But while in the first case the range is more concentrated in the highest scores for females (8-10 over 10 possible points) and males scores present a higher dispersion (6-10); this situation is reversed in the

second case where females scores present a higher dispersion (6-10) and males results are more concentrated in the highest results (9-10), as shown in Fig. 4. As we can see, both women and men show an equal average interest in the subject, obtaining both of them quite high mean scores.

4 Discussion

In this work we analyzed the impact of the first two editions of the annual Educational Workshop on Bioinformatics at the UGR on the attendants with special emphasis from the gender viewpoint. A large number of participants were involved, and the attendance even increased in the second edition. The collaboration shown in the satisfaction survey, more than 80% from the total, implied a great interest on the topic of Bioinformatics within the UGR community at all university levels. This great support had no significant distinction among genders at global level.

When analyzing the participation of undergraduate students, we observed almost equal interest of both females and males in Bioinformatics in the "Bio-" degrees. Furthermore, the evaluation results suggested that men and women are equally fitted to learn about the subject. In the case of postgraduate degrees, data showed an increase in the amount of females in Bio-Health Degrees and, in a lesser ratio in STEM Degrees, who were interested in the topic.

To the light of our data we may conclude that women are initially more attracted by Bio related Degrees than by Computer Science. A "turning point" appears during the latter courses and subjects. It is only when they learn about new Next Generation Techniques (NGT) and their impact in Biology, Biochemistry, Biotechnology and personalized medicine [23], when their interest in developing STEM skills grows significantly. This way, we may find a greater percentage of postgraduate female students getting information and formation in the area of Bioinformatics.

We find that women show in early university stages the same interest and fitness capacity as males in STEM skills applied to their study field. An analysis of the gender data from our Bioinformatics workshop shows much greater gender equity than in STEM degrees and closer to Biology degrees where the number of female enrollments reaches the 60%. Understanding which factors influence this behavior could provide useful insight when designing interventions to help narrow the gender gap in STEM degrees [9].

5 Conclusion

In this paper, we have reviewed the information on gender equality from the first and second editions of the Educational Workshop of Bioinformatics at the UGR. To do so, we have first described the organization and structure of the workshop, in order to understand the conditions under study. Specifically, we have defined the workshop program, the quality and satisfaction evaluation and the participants' evaluation.
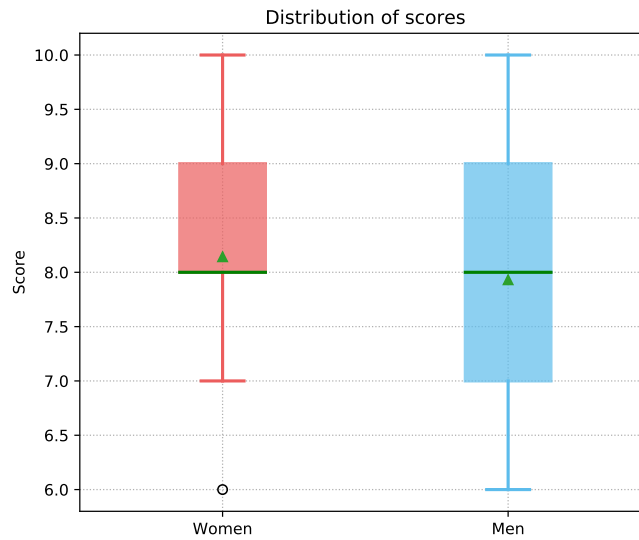
(a) I Edition
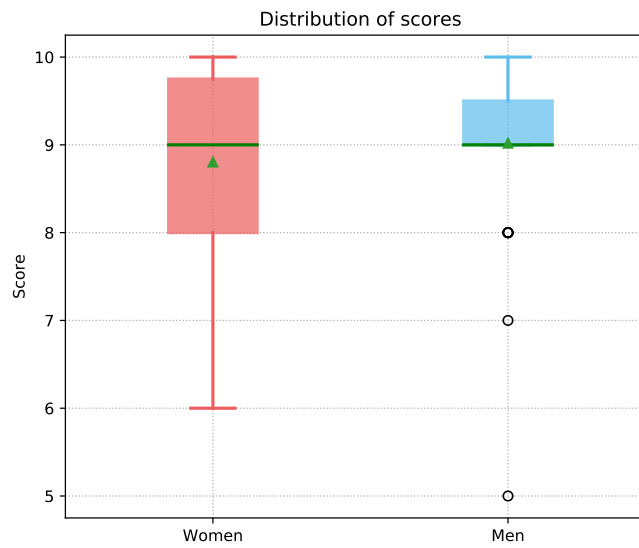
(b) II Edition

Fig. 4: Distributions of the scores obtained in the knowledge acquisition test by gender. On the left (red) distribution among female participants and on the right (blue) among male ones. Green line represents median and a triangle the mean

Our analysis were divided into three different aspects, namely the workshop attendance, the participants' background diversity, and the assessment of

knowledge acquisition. In this sense, we have contrasted the results between both editions showing a significant improvement in the general results for the latter one.

Finally, we have posed a discussion on the gender disparities within the STEM areas, in particular from the viewpoint of Bio- degrees and Computer Science. From the data collected from our two workshops, we have concluded that the topic of Bioinformatics implies a very interesting synergy between both areas, narrowing the "gender gap" in general STEM. In accordance with the former, our future efforts in academia must be oriented towards the development of additional informative sessions for students. This may allow them to acknowledge the significance of acquiring relevant skills such as NGT, scripting tools, and novel programming languages, among others.

References

1. Arabnia, H.R., Tran, Q.N.: Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology: Algorithms and Software Tools. Morgan Kaufmann (2015)
2. Bendels, M., Mller, R., Brueggmann, D., Groneberg, D.: Gender disparities in high-quality research revealed by nature index journals. PLoS ONE **13**(1), 1–21 (2018). https://doi.org/10.1371/journal.pone.0189136
3. Bonham, K.S., Stefan, M.I.: Gender disparity in computational biology research publications. bioRxiv (2017). https://doi.org/10.1101/070631, `http://biorxiv.org/content/early/2017/04/13/070631`
4. Bonham, K., Stefan, M.: Women are underrepresented in computational biology: An analysis of the scholarly literature in biology, computer science and computational biology. PLoS Computational Biology **13**(10), 1–12 (2017). https://doi.org/10.1371/journal.pcbi.1005134
5. Can, Y.W.: Yes, women can! au féminin pluriel, `https://yeswomencan.ca/`
6. Commission, E.: Europe 2020, `http://ec.europa.eu/europe2020/index_en.htm`
7. Commission, E.: Gender equality as a crosscutting issue in horizon 2020, `https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/interim_evaluation_gender_long_final.pdf`
8. Commission, E.: Guidance to facilitate the implementation of targets to promote gender equality in research and innovation, `https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/KI-07-17-199-EN-N.pdf`
9. Commission, E.: She figures 2015, `https://ec.europa.eu/research/swafs/pdf/pub_gender_equality/she_figures_2015-final.pdf`
10. Eurostat: Gender statistics at regional level, `http://ec.europa.eu/eurostat/statistics-explained/index.php/Gender_statistics_at_regional_level`
11. Eurostat: Tertiary education statistics, `http://ec.europa.eu/eurostat/statistics-explained/index.php/Tertiary_education_statistics`
12. Ford, H., Brick, C., Blaufuss, K., Dekens, P.: Gender inequity in speaking opportunities at the american geophysical union fall meeting. Nature Communications **9**(1), 1–6 (2018). https://doi.org/10.1038/s41467-018-03809-5
13. University of Granada, University of Jaén, P.d.O.U.: Seré ingeniera, `http://sereingeniera.ugr.es/`
14. Hypatia: Hypatia project, `http://www.expecteverything.eu/hypatia/`

15. Kang, J., Park, S., Venkat, A., Gopinath, A.: Quantitative analysis of the trends exhibited by the three interdisciplinary biological sciences: biophysics, bioinformatics, and systems biology. Journal of microbiology & biology education **16**(2), 198 (2015)
16. Likert, R.: A technique for the measurement of attitudes. Archives of psychology (1932)
17. Masic, I.: The most influential scientists in the development of medical informatics (13): Margaret belle dayhoff. Acta Informatica Medica **24**(4), 299 (2016)
18. Nittrouer, C.L., Hebl, M.R., Ashburn-Nardo, L., Trump-Steele, R.C., Lane, D.M., Valian, V.: Gender disparities in colloquium speakers at top universities. Proceedings of the National Academy of Sciences **115**(1), 104–108 (2018)
19. Ouzounis, C.A.: Rise and demise of bioinformatics? promise and progress. PLoS Comput Biol **8**(4), e1002487 (2012)
20. Pevzner, P., Shamir, R.: Computing has changed biologybiology education must catch up. Science **325**(5940), 541–542 (2009)
21. Sasor, S., Cook, J., Duquette, S., Loewenstein, S., Gallagher, S., Tholpady, S., Chu, M., Koniaris, L.: Scholarly activity in academic plastic surgery: the gender difference. Journal of Surgical Research **229**, 332–336 (2018). https://doi.org/10.1016/j.jss.2018.04.031
22. Scilico, L.: Bioinformatics job, `http://www.bioinformatics.org/jobs/`
23. Serratì, S., De Summa, S., Pilato, B., Petriella, D., Lacalamita, R., Tommasi, S., Pinto, R.: Next-generation sequencing: advances and applications in cancer diagnosis. OncoTargets and therapy **9**, 7355 (2016)
24. for Stony Brook, T.C.: Stony brook university libraries, `http://guides.library.stonybrook.edu/bioinformatics`
25. UGR: University of granada degrees enrollment, `http://transparente.ugr.es/pages/estudiantes#__doku_matriculas`
26. University, P.F.: Wisibilízalas, `https://www.upf.edu/web/wisibilizalas`
27. WISE: Wise campaign, `https://www.wisecampaign.org.uk/`