STATE OF GENDER EQUALITY IN AND BY ARTIFICIAL INTELLIGENCE

J. David Patón-Romero^{1,2}, Ricardo Vinuesa³, Letizia Jaccheri¹ and Maria Teresa Baldassarre⁴

¹Department of Computer Science, Norwegian University of Science and Technology (NTNU). Sem Sælands Vei 7, 7034 Trondheim, Norway

²Department of IT Management, Simula Metropolitan Center for Digital Engineering (SimulaMet). Pilestredet 52, 0167 Oslo, Norway

³FLOW, Engineering Mechanics, KTH Royal Institute of Technology. SE-100 44 Stockholm, Sweden ⁴Department of Informatics, University of Bari "Aldo Moro" (UniBa). Via Edoardo Orabona, 4, 70126 Bari, Italy

ABSTRACT

When talking about sustainability, we usually think that it is only about safeguarding the environment; nothing is further from reality. Of course, the environment is a crucial component of sustainability and our survival, but it is important to recall that the society and the economy play important roles in this regard, and without the interconnection and development of these three perspectives it will not be possible to achieve sustainable progress. The Sustainable Development Goals (SDGs) established by the United Nations (UN) defend this idea and address the main challenges that humanity faces. One of these challenges is gender equality, which is identified in the perspective of social sustainability through SDG 5. Gender equality is a very complex and difficult challenge to address due to the great cultural diversity of our society. Thus, achieving this goal will require laying a solid foundation and working together by combining very different fields of knowledge. In this sense, Artificial Intelligence (AI) is one of the fields that is currently having the greatest impact and relevance for the development of new technologies and for the advancement of numerous areas. This growing evolution of AI demonstrates that its repercussions at the social level must be analyzed and addressed in such a way that AI becomes a positive asset for sustainability and, in this particular case, for gender equality.

For all these reasons, this study aims to analyze the current state of the art and collect the existing knowledge in the fields of AI and gender equality, by conducting a Systematic Mapping Study (SMS). The obtained results and findings have allowed us to identify the most relevant advances in this regard, as well as the gaps and drawbacks that currently exist and on which we must urgently focus to address gender equality both in and by AI. In the same way, these findings demonstrate the limited joint development of both fields, but also indicate an increase in the relevance and the number of proposals that these fields are receiving in recent years.

KEYWORDS

Artificial Intelligence; Gender Equality; Social Sustainability; Literature Review; Literature Analysis

1. INTRODUCTION

Sustainability has become a key actor for the development and advancement of civilization. However, many times it is only interpreted as an environmental characteristic and other perspectives such as social and economic sustainability are ignored (Purvis et al., 2019). These three perspectives go hand in hand and must be addressed together, since, for example, it is not possible to aim to achieve sustainable development by and for the environment if changes are not conducted in society itself to support it.

According to the findings identified by (Harari, 2018), sustainability and Artificial Intelligence (AI) are two of the biggest challenges faced by humanity. From Information Technology (IT), AI has become one of the most relevant and innovative fields (Carleton et al., 2020; Menzies, 2019). The unstoppable progress that AI is experiencing demonstrates the importance of pursuing AI applications that can help to achieve a sustainable development and use in this regard (Nishant et al., 2020; Vinuesa et al., 2020). Thus, it is essential to relate both fields and work to achieve sustainability in and by AI. However, the focus has always been on the relationship between AI and IT in general with energy/climate neutrality (i.e., environmental sustainability) and marginally on its relationship with social and economic sustainability (Becker, 2015). Some of these aspects are discussed by (Vinuesa & Sirmacek, 2021).

Thus, this study aims to shed light into the issue and show the current relationship between AI and social sustainability, focusing on gender equality, which is one of the Sustainable Development Goals (SDGs) established by the 2030 Agenda (United Nation, 2015) which is having more focus and relevance regarding social sustainability (Rosa, 2017).

Current research shows that women are underrepresented in technology research, practice, and education (Albusays et al., 2021). Likewise, gender imbalance in technology has been seen as harming the economy, as highlighted by the *European Commission* when identifying an annual productivity loss of 16 billion Euro for the European economy (European Commission, 2018). In the same way, the OECD (*Organization for Economic Co-operation and Development*) states that "greater inclusion of women in the digital economy and increased diversity bring value, both social and economic" (OECD, 2018).

Therefore, the present literature analysis through a Systematic Mapping Study in the areas of AI and gender equality will be useful, since it will allow knowing the latest knowledge and establishing the pillars that will guide the development of new and innovative research and ideas in this regard. Thanks to a greater understanding of the interplay between AI and gender equality, it will it be possible to understand the changes and challenges that exist towards achieving a sustainable development through the SDGs.

The rest of this study is organized as follows: Section 2 contains the background about 2030 Agenda, gender equality, and AI; Section 3 presents the research methodology followed to analyze the state of the art in the fields of gender equality and AI; Section 4 shows the results obtained from the analysis performed; Section 5 discusses the findings, limitations, and implications that have been reached; finally, Section 6 contains the conclusions and lines for future work in this regard. Likewise, Appendix A includes the list of primary studies selected during the analysis of the state of the art; Appendix B shows the answers to the established research questions from each of these primary studies; and Appendix C shows an overview of the topics covered in each primary study.

2. BACKGROUND

2.1 2030 Agenda & Sustainable Development Goal 5

The 2030 Agenda (United Nation, 2015) is an initiative promoted and agreed upon by the 193 Member States of the United Nations (UN) with the aim to achieve the so-called Sustainable Development Goals (SDGs). This includes a total of 17 Goals and 169 Targets that address the three pillars of sustainability (environmental, social, and economic) (Purvis et al., 2019), including areas such as climate change, economic inequality, innovation, natural resources consumption, peace, and justice, among other priorities. Likewise, for each of the Targets there is also a set of indicators that make it possible to measure the progress made in this regard (United Nations, 2017).

Among these SDGs, this study aims to focus on Goal 5 (Gender Equality), one of the Goals belonging to the field of social sustainability. The main purpose of this Goal is to "achieve gender equality and empower all women and girls", for which it establishes the following 9 Targets (United Nation, 2015):

- 5.1 End all forms of discrimination against all women and girls everywhere.
- 5.2 Eliminate all forms of violence against all women and girls in the public and private spheres, including trafficking and sexual and other types of exploitation.
- 5.3 Eliminate all harmful practices, such as child, early and forced marriage and female genital mutilation.
- 5.4 Recognize and value unpaid care and domestic work through the provision of public services, infrastructure and social protection policies and the promotion of shared responsibility within the household and the family as nationally appropriate.
- 5.5 Ensure women's full and effective participation and equal opportunities for leadership at all levels of decision-making in political, economic and public life.
- 5.6 Ensure universal access to sexual and reproductive health and reproductive rights as agreed in accordance with the Programme of Action of the International Conference on Population and Development and the Beijing Platform for Action and the outcome documents of their review conferences.
- 5.a Undertake reforms to give women equal rights to economic resources, as well as access to ownership and control over land and other forms of property, financial services, inheritance and natural resources, in accordance with national laws.
- 5.b Enhance the use of enabling technology, in particular information and communications technology, to promote the empowerment of women.
- 5.c Adopt and strengthen sound policies and enforceable legislation for the promotion of gender equality and the empowerment of all women and girls at all levels.

It is important to highlight that the 2030 Agenda identifies two types of Targets within the SDGs: 1) "outcome" Targets (i.e., circumstances to be attained), labeled by numbers; and 2) "means of implementation" Targets, labeled by lower case letters.

2.2 Gender Equality & Artificial Intelligence

In order to achieve gender equality, as defended by SDG 5 of the 2030 Agenda, actions must be conducted in all areas of knowledge. AI has proven to be an increasingly important actor in the development of new and innovative systems used by all levels of society (Lu et al., 2018). That is why it is vital that the entire life cycle of these systems is committed to achieving a better society and, therefore, gender equality must play an important role in this regard.

In general terms, it could be said that the main goal of AI in social sustainability is "the study and practice of design, build and use of AI systems with a positive impact on the society". However, when relating the terms of gender equality and AI, there is no clear criterion or definitions per se. To establish this relationship, the definitions and same logic as that used for the terms *Green by IT* and *Green in IT* (idea proposed in (Erdélyi, 2013)) will be followed, which defend sustainability in and by IT. Thus, in this case, we are faced with two perspectives:

- Gender by AI: in which AI provides the necessary tools to achieve gender equality through different contexts (i.e., AI as an enabler).
- *Gender in AI*: in which AI itself produces a negative impact on gender equality (e.g., lack of balance during the development of a system) and, therefore, said impact must be reduced (i.e., AI as a producer).

3. RESEARCH METHODOLOGY

A Systematic Mapping Study (SMS) is a research method used to collect, analyze, and categorize existing information from a specific context. In this study, the guidelines established by (Kitchenham, 2007) have been followed, adopting also the lessons learned for the data extraction and analysis identified by (Brereton et al., 2007), and considering examples of application of SMSs such as (Petersen et al., 2008). Thus, the characteristics established during the planning stage are shown below, as well as how the execution stage was conducted.

3.1 Planning Stage

3.1.1 Research Questions

The main goal of this study is to inspect the current state and existing relationship between the fields of AI and gender equality. In this way, it is intended to collect and categorize all the information in this regard and identify the gaps that exist in order to develop new research proposals. To do this, the research questions (RQs) shown in Table 1 have been established.

Table 1. Research questions

Research question	Motivation
RQ1. What kind of studies exist on AI and gender	Determine the type, number of publications, and trend
equality?	over recent years in relation to AI and gender equality.
RQ2. What gender equality Targets are addressed in	Determine what gender equality Targets are addressed
and by AI?	in/by AI to identify possible opportunities and threats.
RQ3. What kind of AI proposals exist to address	Determine the AI proposals that exist to address
gender equality?	gender equality to identify trends and possible gaps in
	or by AI.

3.1.2 Search Strategy

As a strategy for search the relevant studies and information, the *Scopus* database will be used. To this end, we decided to conduct a general search and a search for each of the Targets identified by the SDG 5 of the 2030 Agenda (United Nation, 2015) (i.e., 10 searches). In this way, specific terms of each Target can be addressed in more detail and the identification of studies in this regard is facilitated. Thus, Table 2 shows the search strings that will be used. As can be seen, these search strings are divided into two main parts (the two contexts within the scope of this study).

Table 2. Search strings

Scope	Search string
General	("Artificial Intelligence" OR AI) AND (Gender OR "Women rights" OR "Social sustainability"
	OR "SDG 5")
Target 5.1	("Artificial Intelligence" OR AI) AND ((Women OR Girls OR Gender) AND Discrimination)
Target 5.2	("Artificial Intelligence" OR AI) AND ((Women OR Girls OR Gender) AND (Violence OR
	Exploitation OR Trafficking))
Target 5.3	("Artificial Intelligence" OR AI) AND (((Women OR Girls OR Gender) AND "Harmful
	practices") OR ((Child OR Early OR Forced) AND Marriage) OR "Genital mutilation")
Target 5.4	("Artificial Intelligence" OR AI) AND ("Care work" OR "Domestic work" OR "Social protection
	policies" OR "Shared responsibility")
Target 5.5	("Artificial Intelligence" OR AI) AND ((Women OR Girls OR Gender) AND ("Equal
	opportunities" OR Participation OR Leadership))
	("Artificial Intelligence" OR AI) AND ((Sexual OR Reproductive) AND (Health OR Rights))
	("Artificial Intelligence" OR AI) AND ((Women OR Girls OR Gender) AND Equal* AND Rights)
	("Artificial Intelligence" OR AI) AND ((Women OR Girls) AND Technology)
Target 5.c	("Artificial Intelligence" OR AI) AND ((Women OR Girls OR Gender) AND (Equal* OR
	Empower*))

These search strings will be applied to the title, abstract and keywords of the studies. Likewise, publications from 2010 and onwards will be considered, since it has been during the last decade when, mainly, the area of gender equality has had its momentum.

3.1.3 Selection Criteria

All the documents and information collected through the searches will be analyzed considering the title, abstract, and keywords of each one. This will determine which studies will be included for a more detailed analysis. To do this, on the one hand, those studies that meet the following inclusion criteria will be considered for further analysis:

- I1. Studies in English dealing with AI and gender equality.
- **I2.** Studies published between 2010 and 2021 in journals, conferences, and/or workshops, with peer review process.

On the other hand, the studies that meet any of the following exclusion criteria will be automatically discarded:

- **E1.** Discussion or opinion studies, as well as those that are only available as abstract or presentation.
- **E2.** Duplicate studies (in which case will be considered the most complete and recent).
- **E3.** Studies whose main contribution is not related to AI and gender equality, or where AI and gender equality are not related to each other.

In the same way, the snowballing effect (Wohlin, 2014) will be followed, so the documents referenced in the considered studies will also be evaluated for their possible inclusion.

3.1.4 Quality Assessment Criteria

One of the most critical points to obtain representative and relevant results and references for future research is the quality assessment of the studies. To do this, the following issues have been established that will be analyzed following a scoring system of three values (-1, 0, +1), generating a quality result for each study between -4 and +4:

- a. The study presents a detailed description on how AI can contribute to gender equality. *Yes* (+1); *Partially* (0); *No* (-1).
- b. The study validates the proposal or idea that it defends. *Empirically validated* (+1); *Theoretically validated* (0); *Not validated* (-1).
- c. The study has been published in a relevant journal¹/conference². *High ranking (+1); Medium ranking (0); Low ranking or not indexed (-1).*
- d. The study has been cited by other authors in publications. *More than 5 cites* (+1); *Between 1 and 4 cites or recently published* (0); *Not cited* (-1).

3.1.5 Data Extraction

A series of answers have been established for each RQ (cf. Table 3). In this way, the same data extraction criteria will be applied to all studies, facilitating their analysis and categorization.

Table 3. Classification schema

Research question	Answers		
RQ1. What kind of studies exist on AI and	a. State of the art a	lation	
gender equality?*	b. Proposal d. Others		rs
RQ2. What gender equality Targets are	a. Target 5.1	d. Target 5.4	g. Target 5.a
addressed in and by AI?**	b. Target 5.2	e. Target 5.5	h. Target 5.b
	c. Target 5.3	f. Target 5.6	Target 5.c
RQ3. What kind of AI proposals exist to	a. Gender by AI		
address gender equality?	b. Gender in AI		

^{*}The answers to RQ1 follow the idea of the example of (Petersen et al., 2008).

3.1.6 Synthesis Methods

A both quantitative and qualitative synthesis of data will be conducted related to the answers to the RQs and the quality evaluations, respectively. These syntheses will be represented by tables and/or graphs with the results in a matter of numbers and/or percentages, as well as bubble plots to analyze how the RQs are related through their answers.

3.2 Execution Stage

In order to apply the protocol established during the planning stage, three main phases have been followed during the execution stage:

• **First phase.** Based on the identification of potential studies. To do this, first, after performing the 10 searches applying the search strings (cf. Table 2) on the *Scopus* database, 3,558 studies were obtained. Then, the selection criteria were applied to these studies, considering the abstract of each one, and 169 potential studies were obtained.

^{**}The answers to RQ2 have their origin in the Targets of the SDG 5 from the 2030 Agenda (United Nation, 2015).

¹ Following the Journal Citation Reports (JCR): https://jcr.clarivate.com/

² Following the GII-GRIN-SCIE Conference Rating: https://scie.lcc.uma.es/

- **Second phase.** Oriented to the identification of primary studies, by means of which the selection criteria were applied again, but this time on the complete content of each one of the 169 potential studies. As a result, 29 primary studies were obtained.
- Third phase. This last phase is dedicated to the compilation of results, for which the characterization of the primary studies was performed through the answers to the RQs and the obtaining of the main findings, as well as the quality assessment of said studies.

4. RESULTS

The general results obtained after the execution of the SMS are shown below, answering each of the established RQs. It is important to highlight that, related to these results, Appendix A includes the list of references of primary studies, while Appendix B contains a summary table with the mapping of the answers to the RQs of each of these primary studies.

4.1 RQ1. What Kind of Studies Exist on AI and Gender Equality?

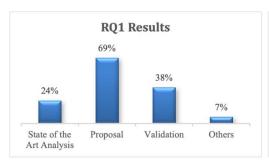
The main objective of this RQ is based on identifying the type of studies that currently exist in AI and gender equality. In this way, it is possible to determine which are the most relevant studies when, e.g., conducting a new proposal in this regard or considering a specific proposal that is validated for the application of a case.

Our results (represented in Figure 1) show that about 24% of the studies found (S05, S08, S09, S11, S17, S18, and S21) are based on or contain some analysis of the state of the art on the field that concerns us.

Likewise, 69% of the studies (20 in total) deal with specific proposals to address gender equality from some point in and by AI. However, of all these proposals only 11 are validated by some empirical case (S01, S02, S04, S10, S13, S16, S22, S23, S27, S28, and S29).

It is also important to highlight 2 studies (S03 and S06), which are based on experiments aimed at assessing the impact of gender in and by AI and remove some biases in this regard.

Finally, related to this RQ, it is also important to analyze the evolution of the studies over the last few years. Figure 2 shows how this progression has been, through which it can be seen that mainly in the last 2 years there has been a boom in publications in this regard. This is because it is in recent years when more efforts are beginning to be made to achieve the objectives of the 2030 Agenda (United Nation, 2015).



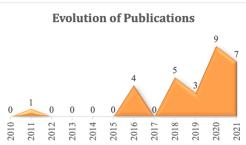


Figure 1. Results for the RQ1 (percentage of studies in each of the four categories)

Figure 2. Evolution of the number of publications on gender equality and AI in recent years

4.2 RQ2. What Gender Equality Targets are Addressed in and by AI?

This RQ is the main one of the present study, since its objective is to analyze and map the studies according to the Target(s) of the SDG 5 from the 2030 Agenda (United Nation, 2015) that address. In this way, it is possible to identify which are the Targets that are usually dealt with in and by AI, as well as the possible Targets that are not yet covered and need development in this regard.

From the results obtained (represented in Figure 3), there are only studies that address 5 of the 9 Targets established. Targets 5.3, 5.4, 5.5, and 5.a are not covered in any of the studies, so there is no evidence on possible developments in AI that address gender equality in the specific contexts of these Targets (it will be discussed later in detail).

Regarding the Targets that are covered, Target 5.1 is the one with the most development in the AI area, since 52% of the studies (15 in total) base their objective on addressing the context of this one. Likewise, following a decreasing order, Target 5.2 is found in 34% of the studies (10), Target 5.6 in 17% (5), Target 5.b in 17% (5), and Target 5.c in 7% (2).

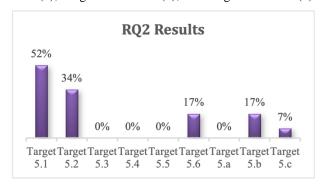


Figure 3. Results for the RQ2 (percentage of studies addressing each of the Targets within SDG 5)

4.3 RQ3. What Kind of AI Proposals Exist to Address Gender Equality?

The last of the RQs is focused on knowing the studies that are oriented both to achieve gender equality in different areas through the use of AI (*Gender by AI*) and to implement measures that help achieve better gender equality issues in AI itself (*Gender in AI*). In this regard, the results show equality in both perspectives, since 18 studies deal with *Gender by AI* and 13 with *Gender in AI*.

First, the studies S01, S02, S03, S04, S05, S06, S07, S10, S12, S13, S14, S15, S16, S23, S24, S25, S26, and S29 include some proposal or characteristics related to the context of *Gender by AI*. Second, *Gender in AI* is addressed in the studies S03, S05, S08, S09, S11, S17, S18, S19, S20, S21, S22, S27, and S28. And, finally, it should be noted that the studies S03 and S05 consider both perspectives.

Thus, 62% of the studies are framed in the perspective of *Gender by AI*, while 45% deal with *Gender in AI*.

5. DISCUSSION

5.1 Principal Findings

The main objective of this study is based on knowing the state of the art and the current relationship between the fields of gender equality and AI. In this way, it is intended to identify important aspects and gaps that help to develop new innovative ideas in this regard. Therefore, after analyzing in detail the obtained results, the following observations can be made:

- Focus on Gender by AI. Although there is a high level of equality in the number of studies that deal with Gender by AI and Gender in AI, in recent years the main focus has been the development of studies related to Gender by AI. This could be due to the greater ease of developing a specific technology to address an aspect related to gender equality (such as, e.g., an AI device that detects dangerous situations for women) than, e.g., change the business/management processes that guide organizations when developing new AI proposals so that they follow a set of best practices that respect gender equality. Undoubtedly, the latter is more complicated, because it is not only necessary to understand both fields to develop useful and applicable best practices for most contexts, but also a high number of practical cases and the involvement of external actors who allow validation of these practices are needed. Therefore, the fact that the focus is currently on Gender by AI is an issue that can generate a lot of controversy, since the "cart is being put before the horse" and the question arises as to whether AI proposals that help gender equality can actually be developed, when in AI itself and in all the processes that surround it (i.e., the basis) there is no such equality.
- Inequality when dealing with the Targets. It is a normal result that there is a difference in the number of studies when addressing different contexts, but, in this case, the difference is quite large. Thus, to improve understanding, we can talk about three groups:
 - Advanced development: the Target 5.1 is the only one in this group, since it is the
 most addressed by the studies found. This is because it is a fairly general Target

- whose objective is "end all forms of discrimination against all women and girls everywhere" (United Nation, 2015). In general terms, the main purpose of SDG 5 of the 2030 Agenda is the one that defines this Target and that is why most studies tend to focus on it. However, we must not forget that there are other Targets with more specific purposes and that they must be addressed urgently.
- o *Medium development:* the Targets 5.2, 5.6, 5.b, and 5.c (from highest to lowest incidence) are found in this group. It is always relevant to find evidence that supports, in this case, the specific context of each of the Targets. However, the evidence is quite scarce, and it is necessary to continue developing new ideas that address these Targets, as well as improving the current ones. From a practical point of view, following the evidence found, these Targets can be addressed in a simple way in and by AI. For example, Target 5.2 aims to "eliminate all forms of violence against all women and girls in the public and private spheres, including trafficking and sexual and other types of exploitation" (United Nation, 2015), which has been shown to be easily addressed through specific AI proposals that analyze various situations in order to identify potential dangers affecting the integrity of women and girls. In general terms, we are on the right path, but we must not get lost and continue dedicating efforts to develop proposals in and by AI in these Targets.
- Null development: the Targets 5.3, 5.4, 5.5, and 5.a are not addressed by any study. This may be due to the difficulty of identifying an idea in and/or by the AI that effectively and efficiently addresses the specific contexts of these Targets. For example, regarding the Target 5.3 ("eliminate all harmful practices, such as child, early and forced marriage and female genital mutilation" (United Nation, 2015)), it is difficult to understand that AI can do something about it, since it deals with behaviors rooted in certain cultures and the application field of AI may not be sufficient to adequately address such a cultural change. However, the Target 5.a ("undertake reforms to give women equal rights to economic resources, as well as access to ownership and control over land and other forms of property, financial services, inheritance and natural resources, in accordance with national laws" (United Nation, 2015)) can be applied in the field of AI through, e.g., systems that analyze the profiles of certain candidates for obtaining economic resources of different kinds, without any type of bias related to the gender of each person. For this reason, all the Targets of the SDG 5 from the 2030 Agenda are equally important, but it is necessary and transcendental to guide and begin to dedicate efforts to develop these Targets that have not yet been explored, in order to propose ideas in and by AI that allow progress in gender equality in this regard.
- Low number of practical cases. When developing a proposal, it is essential to conduct practical cases that validate it and demonstrate its applicability, quality, effectiveness, and efficiency when addressing the objective for which it was developed. However, of the 20 proposals identified through the SMS, only 11 (i.e., 55%) have been validated. This supposes a too low number of validated proposals, since all or the vast majority should have been applied in some practical case, showing that they are complete and serious proposals. For this reason, it is extremely important that any development conducted in and/or by AI to, in this case, address some Target(s) of SDG 5 from the 2030 Agenda, be accompanied by a practical application and validation that demonstrates its applicability, effectiveness, and efficiency in this regard.

• Lack of joint development of both fields. Although there is evidence on the development of ideas in and by AI to address the specific context of gender equality, there is very little. This, together with the analysis of the progression of publications over the last few years, demonstrates the novelty of this field and the growing interest in conducting new research in this regard. It is very important to closely follow this progression, as well as analyze the new studies that arise and develop new ideas that contribute to this research field so important and necessary to achieve gender equality, improving the field of social sustainability and complying with the provisions of the 2030 Agenda defined by the UN (United Nation, 2015).

5.2 Limitations

Although an attempt has been made to design this study to avoid or mitigate the possible limitations (such as the case of performing a general search and 9 searches for each of the Targets, with the aim to find studies with very specific terminology for certain contexts), there are always limitations that can affect when identifying and analyzing results and findings.

It should not be forgotten that the present analysis comes from the perspective of the authors and may not be interpreted in the same way by other researchers or professionals in the field. Likewise, it is possible that certain literature on the field has been overlooked, or even that some more recent evidence on the studies found has not yet been published at the time of the SMS execution. Consequently, to mitigate the risks in this regard, several authors analyzed and interpreted the data and results obtained here, contributed to the final consensus, thus reducing the bias among each other.

5.3 Implications

The development of this study has a high implication and significance for the fields in which it is found. As has been shown, there are few studies that put the fields of AI and gender equality in common. Thus, thanks to this study, not only the state of the art in this regard is identified, but also the gaps and possible lines of research that improve existing studies or that address new and innovative ideas not considered until now.

In Section "5.1 Principal Findings" a discussion has been conducted in which different lines of future work/research are identified. These can be used by researchers who are in the fields of AI, IT in general, gender equality and social sustainability. Therefore, this study is a necessary starting point and the demonstration of the importance of the fields that concern us, which will attract new researchers and professionals to the development of new proposals with the goal of achieving gender equality in and by AI.

6. CONCLUSIONS AND FUTURE WORK

The increasing use of technology and AI by a wide range of people around the world shows that they must be driven by and for the whole of society, avoiding gender, culture, religion, and other kinds of discrimination. However, women and other vulnerable and discriminated minorities are underrepresented in this regard and the progress to get around this situation is slow and scant (Adams & Khomh, 2020; Albusays et al., 2021).

That is why this study is focused on analyzing the state of the art in the fields of AI and gender equality, through a Systematic Mapping Study (SMS). On the one hand, AI is becoming a fundamental field for the development of new and innovative technologies, so it is vital that it represents a positive asset for sustainability (Harari, 2018; Nishant, 2020). On the other hand, gender equality must be addressed in all fields of knowledge without exception, and, above all, it needs a boost with new ideas and proposals in the field of technology and AI (European Commission, 2018; OECD, 2018).

Through the results presented here, not only the current state in this regard has been identified, but also a series of problems and gaps that must be addressed. The novelty of this work has been demonstrated, due to the small number of studies in this area, as well as the large increase in studies and the growing importance that these fields are taking in recent years. Therefore, it is necessary to continue with this momentum and address the gaps that exist by developing proposals and empirical validations that cover the different specific contexts of the Targets identified by the SDG 5 of the 2030 Agenda (United Nation, 2015) both in and by AI.

Thus, as future work, we are conducting new studies on gender equality in different areas related to technology, such as IT processes (Patón-Romero et al., 2023) and entrepreneurship in the IT sector (Wiken Wilson & Patón-Romero, 2022), with the aim to identify different points of view and links that help together to develop new and better proposals to address gender equality in this regard. Likewise, we also intend to develop a framework of best practices that establish the bases for the development, validation, evaluation, and improvement of proposals for both *Gender by AI* and *Gender in AI*. In this way, we want to facilitate and promote these fields both at the research and professional level in organizations. Similarly, this type of practices can serve to further expand and improve (from the perspective of social sustainability) frameworks already developed and successfully applied in organizations around the world, as is the case of the "Governance and Management Framework for Green IT" (Patón-Romero et al., 2017; Patón-Romero et al., 2019; Patón-Romero et al., 2021).

A society unable to change will not generate any progress. Let us be the change our society needs, promoting new and inclusive ideas for all humankind.

ACKNOWLEDGEMENT

This work is result of a postdoc from the ERCIM "Alain Bensoussan" Fellowship Program conducted at the Norwegian University of Science and Technology (NTNU). This research work is also part of the COST Action - European Network for Gender Balance in Informatics project (CA19122), funded by the Horizon 2020 Framework Programme of the European Union. Likewise, RV acknowledges the financial support of the Swedish Research Council (VR).

REFERENCES

Adams, B. & Khomh, F., 2020. The Diversity Crisis of Software Engineering for Artificial Intelligence. *IEEE Software*, 37(5), pp. 104-108.

Albusays, K., Bjorn, P., Dabbish, L., Ford, D., Murphy-Hill, E., Serebrenik, A., & Storey, M., 2021. The Diversity Crisis in Software Development. *IEEE Software*, 38(2), pp. 19-25.

- Becker, C., Betz, S., Chitchyan, R., Duboc, L., Easterbrook, S. M., Penzenstadler, B., Seyff, N., & Venters, C. C., 2015. Requirements: The Key to Sustainability. *IEEE Software*, 33(1), pp. 56-65.
- Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M., 2007. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, 80(4), pp. 571-583.
- Carleton, A. D., Harper, E., Menzies, T., Xie, T., Eldh, S., & Lyu, M. R., 2020. The AI Effect: Working at the Intersection of AI and SE. *IEEE Software*, 37(4), pp. 26-35.
- Erdélyi, K., 2013. Special factors of development of green software supporting eco sustainability. *Proc. of IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY 2013)*. Subotica, Serbia, pp. 337-340.
- European Commission, 2018. Women in the Digital Age. SMART 2016/0025. European Commission, Brussels, Belgium.
- Harari, Y. N., 2018. 21 Lessons for the 21st Century. Random House, New York, NY, USA.
- Kitchenham, B., 2007. Guidelines for Performing Systematic Literature Reviews in Software Engineering (Version 2.3). EBSE Technical Report. Keele University, Keele, UK.
- Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S., 2018. Brain Intelligence: Go Beyond Artificial Intelligence. *Mobile Networks and Applications*, 23(2), pp. 368-375.
- Menzies, T., 2019. The Five Laws of SE for AI. IEEE Software, 37(1), pp. 81-85.
- Nishant, R., Kennedy, M., & Corbett, J., 2020. Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management*, 53, pp. 102104.
- OECD, 2018. Bridging the Digital Gender Divide. Include, Upskill, Innovate. Organization for Economic Co-operation and Development, Paris, France.
- Patón-Romero, J. D., Baldassarre, M. T., Piattini, M., & García Rodríguez de Guzmán, I., 2017. A Governance and Management Framework for Green IT. *Sustainability*, 9(10), pp. 1761.
- Patón-Romero, J. D., Baldassarre, M. T., Rodríguez, M., & Piattini, M., 2019. A Revised Framework for the Governance and Management of Green IT. *Journal of Universal Computer Science*, 25(13), pp. 1736-1760.
- Patón-Romero, J. D., Baldassarre, M. T., Rodríguez, M., & Piattini, M., 2021. Governance and Management of Green IT. In Calero, C., Moraga, M. A., & Piattini, M. (eds.), 2021. Software Sustainability. Springer Nature Switzerland AG, Cham, Switzerland, pp. 243-267.
- Patón-Romero, J. D., Block, S., Ayala, C., & Jaccheri, L., 2023. Gender Equality in Information Technology Processes: A Systematic Mapping Study. Proc. Of Future of Information and Communication Conference (FICC 2023). San Francisco, CA, USA.
- Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M., 2008. Systematic Mapping Studies in Software Engineering. *Proc. of 12th Int. Conference on Evaluation and Assessment in Software Engineering (EASE 2008)*. Bari, Italy, pp. 68-77.
- Purvis, B., Mao, Y., & Robinson, D., 2019. Three pillars of sustainability: in search of conceptual origins. *Sustainability Science*, 14(3), pp. 681-695.
- Rosa, W., 2017. Goal 5. Achieve Gender Equality and Empower All Women and Girls. *In Rosa, W. (eds.)*, 2017. A New Era in Global Health: Nursing and the United Nations 2030 Agenda for Sustainable Development. Springer Publishing Company, New York, NY, USA, pp. 301-307.
- United Nations, 2015. Transforming Our World: The 2030 Agenda for Sustainable Development. Seventieth Session of the United Nations General Assembly. Resolution A/RES/70/1. United Nations, New York, NY, USA.
- United Nations, 2017. Global indicator framework for the Sustainable Development Goals and targets of the 2030 Agenda for Sustainable Development. *Seventy-first Session of the United Nations General Assembly*. Resolution A/RES/71/313. United Nations, New York, NY, USA.

- Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Nerini, F. F., 2020. The Role of Artificial Intelligence in Achieving the Sustainable Development Goals. *Nature Communications*, 11(1), pp. 1-10.
- Vinuesa, R. & Sirmacek, B., 2021. Interpretable deep-learning models to help achieve the Sustainable Development Goals. *Nature Machine Intelligence*, 3(11), pp. 926.
- Wiken Wilson, A. & Patón-Romero, J. D., 2022. Gender Equality in Tech Entrepreneurship: A Systematic Mapping Study. *Proc. of 3rd Workshop on Gender Equality, Diversity, and Inclusion in Software Engineering (GE@ICSE 2022)*. Pittsburgh, PA, USA.
- Wohlin, C., 2014. Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering. *Proc. of 18th Int. Conf. on Evaluation and Assessment in Software Engineering (EASE 2014)*. London, UK, pp. 1-10.

APPENDIX A. PRIMARY STUDIES SELECTED

Table 4. References of the primary studies

Reference

- S01 Hossain, N., Ovi, J. H., Tasnim, S., Islam, N., & Zishan, S. R., 2021. Design and development of Wearable multisensory smart device for human safety. *Proc. of 2021 IEEE International IOT*, *Electronics and Mechatronics Conference (IEMTRONICS 2021)*. Toronto, Canada, pp. 1-7.
- S02 Islam, R., Keya, K. N., Zeng, Z., Pan, S., & Foulds, J., 2021. Debiasing Career Recommendations with Neural Fair Collaborative Filtering. *Proc. of Web Conference 2021 (WWW '21)*. Ljubljana, Slovenia, pp. 3779-3790.
- S03 Winkle, K., Melsión, G. I., McMillan, D., & Leite, I., 2021. Boosting Robot Credibility and Challenging Gender Norms in Responding to Abusive Behaviour: A Case for Feminist Robots. *Proc. of Companion* of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21 Companion). Boulder, CO, USA, pp. 29-37.
- S04 Miranda, J. A., Canabal, M. F., Gutiérrez-Martín, L., Lanza-Gutierrez, J. M., Portela-García, M., & López-Ongil, C., 2021. Fear Recognition for Women Using a Reduced Set of Physiological Signals. Sensors, 21(5), pp. 1587.
- S05 Guevara-Gómez, A., de Zárate-Alcarazo, L. O., & Criado, J. I., 2021. Feminist perspectives to artificial intelligence: Comparing the policy frames of the European Union and Spain. *Information Polity*, 26(2), pp. 173-192.
- S06 Tomalin, M., Byrne, B., Concannon, S., Saunders, D., & Ullmann, S., 2021. The practical ethics of bias reduction in machine translation: Why domain adaptation is better than data debiasing. *Ethics and Information Technology*, 23, pp. 1-15.
- S07 Siristatidis, C., Stavros, S., Drakeley, A., Bettocchi, S., Pouliakis, A., Drakakis, P., Papapanou, M., & Vlahos, N., 2021. Omics and Artificial Intelligence to Improve In Vitro Fertilization (IVF) Success: A Proposed Protocol. *Diagnostics*, 11(5), pp. 743.
- S08 Cirillo, D., Catuara-Solarz, S., Morey, C., Guney, E., Subirats, L., Mellino, S., Gigante, A., Valencia, A., Rementeria, M. J., Chadha, A. S., & Mavridis, N., 2020. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. NPJ Digital Medicine, 3(81), pp. 1-11.
- S09 Cernadas, E. & Calvo-Iglesias, E., 2020. Gender perspective in Artificial Intelligence (AI). Proc. of Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'20). Salamanca, Spain, pp. 173-176.
- S10 Peña, A., Serna, I., Morales, A., & Fierrez, J., 2020. Bias in Multimodal AI: Testbed for Fair Automatic Recruitment. Proc. of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2020). Virtual, pp. 129-137.
- S11 Wellner, G. & Rothman, T., 2020. Feminist AI: Can We Expect Our AI Systems to Become Feminist?. Philosophy & Technology, 33(2), pp. 191-205.

ID Reference

- S12 Hernández-Álvarez, M. & Granizo, S. L., 2020. Detection of Human Trafficking Ads in Twitter Using Natural Language Processing and Image Processing. Proc. of Advances in Artificial Intelligence, Software and Systems Engineering at the 11th Int. Conference on Applied Human Factors and Ergonomics (AHFE 2020). New York, NJ, USA, pp. 77-83.
- S13 Mensa, E., Colla, D., Dalmasso, M., Giustini, M., Mamo, C., Pitidis, A., & Radicioni, D. P., 2020. Violence detection explanation via semantic roles embeddings. *BMC Medical Informatics and Decision Making*, 20(1), pp. 1-13.
- S14 Khatri, H. & Abdellatif, I., 2020. A Multi-Modal Approach for Gender-Based Violence Detection. Proc. of 2020 IEEE Cloud Summit. Fairfax, VA, USA, pp. 144-149.
- S15 Montiel Fernandez, Z. A., Torres Cruz, M. A., Peñaloza, C., & Hidalgo Morgan, J., 2020. Challenges of Smart Cities: How Smartphone Apps Can Improve the Safety of Women. *Proc. of 2020 4th International Conference on Smart Grid and Smart Cities (ICSGSC 2020)*. Sichuan, China, pp. 145-148.
- S16 Bhagat, P., Prajapati, S. K., & Seth, A., 2020. Initial Lessons from Building an IVR-based Automated Question-Answering System. Proc. of 2020 International Conference on Information and Communication Technologies and Development (ICTD2020). Guayaquil, Ecuador, pp. 1-5.
- S17 Sun, T., Gaut, A., Tang, S., Huang, Y., ElSherief, M., Zhao, J., Mirza, D., Belding, E., Chang, K., & Wang, W. Y., 2019. Mitigating Gender Bias in Natural Language Processing: Literature Review. Proc. of 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019). Florence, Italy, pp. 1630-1640.
- S18 Adams, R. & Loideáin, N. N., 2019. Addressing Indirect Discrimination and Gender Stereotypes in AI Virtual Personal Assistants: The Role of International Human Rights Law. Cambridge International Law Journal, 8(2), pp. 241-257.
- S19 Johnson, K. N., 2019. Automating the Risk of Bias. George Washington Law Review, 87(6), pp. 1214-1271.
- S20 Parsheera, S., 2018. A Gendered Perspective on Artificial Intelligence. Proc. of 2018 ITU Kaleidoscope: Machine Learning for a 5G Future (ITU K 2018). Santa Fe, Argentina, pp. 1-7.
- S21 Leavy, S., 2018. Gender Bias in Artificial Intelligence: The Need for Diversity and Gender Theory in Machine Learning. Proc. of 1st Int. Workshop on Gender Equality in Software Engineering (GE '18). Gothenburg, Sweden, pp. 14-16.
- S22 Sutton, A., Lansdall-Welfare, T., & Cristianini, N., 2018. Biased Embeddings from Wild Data: Measuring, Understanding and Removing. Proc. of Int. Symposium on Intelligent Data Analysis (IDA 2018). Hertogenbosch, Netherlands, pp. 328-339.
- S23 Kudva, V., Prasad, K., & Guruvare, S., 2018. Andriod Device-Based Cervical Cancer Screening for Resource-Poor Settings. *Journal of Digital Imaging*, 31(5), pp. 646-654.
- S24 Rabbany, R., Bayani, D., & Dubrawski, A., 2018. Active Search of Connections for Case Building and Combating Human Trafficking. Proc. of 24th Int. Conf. on Knowledge Discovery & Data Mining (KDD'18). London, UK, pp. 2120-2129.
- S25 Berk, R. A., Sorenson, S. B., & Barnes, G., 2016. Forecasting Domestic Violence: A Machine Learning Approach to Help Inform Arraignment Decisions. *Journal of Empirical Legal Studies*, 13(1), pp. 94-115.
- S26 Alvari, H., Shakarian, P., & Snyder, J. E. K., 2016. A Non-Parametric Learning Approach to Identify Online Human Trafficking. Proc. of 2016 IEEE Conf. on Intelligence and Security Informatics (ISI 2016). Tucson, AZ, USA, pp. 133-138.
- S27 Bolukbasi, T., Chang, K., Zou, J., Saligrama, V., & Kalai, A., 2016. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. Proc. of 30th International Conference on Neural Information Processing Systems (NIPS'16). Barcelona, Spain, pp. 4356-4364.
- S28 Vachovsky, M. E., Wu, G., Chaturapruek, S., Russakovsky, O., Sommer, R., & Fei-Fei, L., 2016. Toward More Gender Diversity in CS through an Artificial Intelligence Summer Program for High School Girls. Proc. of 47th ACM Technical Symposium on Computing Science Education (SIGCSE '16). Memphis, TN, USA, pp. 303-308.
- S29 Crutzen, R., Peters, G. Y., Portugal, S. D., Fisser, E. M., & Grolleman, J. J., 2011. An Artificially Intelligent Chat Agent That Answers Adolescents' Questions Related to Sex, Drugs, and Alcohol: An Exploratory Study. *Journal of Adolescent Health*, 48(5), pp. 514-519.

APPENDIX B. PRIMARY STUDIES MAPPING

Table 5. Answers to research questions from the primary studies

ID	RÇ	21		RQ2								RQ3		Quality		
- עו	a	b	c	d	a	b	c	d	e	f	g	h	i	a	b	score
S01		X	X			X								X		+1
S02		X	X		X							X		X		+3
S03				X	X	X						X		X	X	+4
S04		X	X			X								X		+2
S05	X				X							X	X	X	X	+1
S06				X	X									X		+3
S07		X								X				X		+2
S08	X				X					X					X	+3
S09	X				X							X			X	+1
S10		X	X		X									X		+3
S11	X				X										X	+2
S12		X				X								X		0
S13		X	X			X								X		+3
S14		X				X								X		0
S15		X				X								X		+1
S16		X	X							X				X		+2
S17	X				X										X	+3
S18	X				X								X		X	+3
S19		X			X										X	+3
S20		X			X										X	+1
S21	X				X										X	+1
S22		X	X		X										X	+3
S23		X	X							X				X		+3
S24		X				X								X		+3
S25		X				X								X		+3
S26		X				X								X		+3
S27		X	X		X										X	+4
S28		X	X									X			X	+4
S29		X	X							X				X		+4

APPENDIX C. PRIMARY STUDIES TOPICS

Table 6. Overview of topics covered in the primary studies

ID Topic

S01 Proposal and validation of a wearable AI device for safety to prevent rape, hijacking, unwanted harassment, and other dangerous situations for women.

S02 Proposal and validation of a neural fair collaborative filtering (NFCF), a practical framework for mitigating gender bias in recommending career-related sensitive items (e.g., jobs, academic concentrations, or courses of study) using a pre-training and fine-tuning approach to neural collaborative filtering, augmented with bias correction techniques.

S03 Experiment for exploring the impact of feminist robot behavior on human-robot interaction. Specifically, it considers (i) use of a social robot to encourage girls to consider studying robotics, (ii) if/how robots should respond to abusive, and anti-feminist sentiment and (iii) how robots can be designed to challenge current gender-based norms of expected behavior.

ID Topic

- S04 Proposal of a specialized fear recognition system for women based on a reduced set of physiological signals in order to prevent scenarios such as personal assaults, the abuse of children or the elderly, gender violence or sexual aggression.
- S05 Analysis of 23 official documents about AI issued by the European Union (EU) and Spain to understand how they frame the gender perspective in their discourses.
- S06 Experiment for exploring three automated strategies (downsampling, upsampling, and counterfactual augmentation) for the removal of gender-bias in Neural Machine Translation (NMT) systems.
- S07 Proposal of a tool, integrating omics data (transcriptomics and metabolomics) and artificial intelligence, to propose optimal and personalized treatment options and enhance treatment success rates for In Vitro Fertilization (IVF).
- S08 Review to highlight the available biomedical data types and the role of several AI technologies to understand sex and gender differences in health and disease, as well as to provide recommendations to ensure that these differences are accounted for in AI implementations that inform Precision Medicine.
- S09 Analysis of the degree of introduction of the gender perspective in the AI grades/studies and how to improve the gender competences of the students.
- S10 Proposal and validation of a new experimental framework (FairCVtest) on AI-based automated recruitment to study how multimodal machine learning is affected by biases (including gender discrimination) present in the training data, as well as how to prevent undesired effects of these biases.
- S11 Analysis of four types of solutions to the gender bias in AI (ignoring any reference to gender, revealing the considerations that led the algorithm to decide, designing algorithms that are not biased, or lastly, involving humans in the process).
- S12 Proposal of a Natural Language Processing and Image Processing system for detecting Twitter messages that are suspected of promoting minors for sexual services, and, therefore, related to trafficking of persons.
- S13 Proposal and validation of VIDES (Violence Detection System), that aims to provide the categorization computed through a neural network-based classifier with explanations, in order to detect episodes of violence from texts in emergency room reports (specifically against women, elderly, and children).
- S14 Proposal of a multi-modal approach to measure Gender-Based Violence Index (GBVI) by detecting the coverage of green canopies using satellite imagery in addition to sensing the level of atmospheric pollution to calculate violence occurrences before they even happen.
- S15 Proposal of a smartphone app called Circle Armored that serves as an example of the type of apps that could prevent or assist during violence situations, mainly, against women.
- S16 Proposal and validation of an automated question-answering system in the domain of sexual and reproductive health and rights.
- S17 Review of studies on recognizing and mitigating gender bias in NLP. It also discusses gender bias based on four forms of representation bias and analyzes methods recognizing gender bias, as well as the advantages and drawbacks of existing gender debiasing methods.
- S18 Analysis to critique the reproduction of negative gender stereotypes in virtual personal assistants and explore the provisions and findings within international women's rights law to assess both how this constitutes indirect discrimination and possible remedies.
- S19 Proposal to increase gender inclusion in the development of AI technologies, which will introduce important and diverse perspectives, reduce the influence of cognitive biases in the design, training, and oversight of learning algorithms, and, thereby, mitigate bias-related risk management concerns.
- S20 Proposal of a three-stage pathway towards bridging gender gap in AI. First, is to develop a set of publicly developed standards on AI, which should embed the concept of "fairness by design". Second, is to invest in R&D in formulating technological tools that can help translate the ethical principles into actual practice. Third, is to strive towards reducing gendered distortions in the underlying datasets to reduce biases and stereotypes in future AI projects.
- S21 Analysis to identify the gender bias that exists in AI, especially in machine learning, by analyzing learning bias from text (naming, ordering, biased descriptions, metaphor, and presence of women in text).
- S22 Proposal and validation of methods for measure and removing gender bias from the word embeddings in Natural Language Processing (NLP).
- S23 Proposal and validation of an algorithm for cervix image analysis to provide an instant decision during cervical cancer screening based on images acquired through an android device.

ID Topic

- S24 Proposal of an algorithm for spotting organized activities in escort advertisements.
- S25 Proposal of using a Machine Learning approach to help inform arraignment decisions in domestic violence.
- S26 Proposal that employs both data mining and semi-supervised machine learning techniques to identify the potential human trafficking related advertisements.
- S27 Proposal and validation of algorithms for removing gender bias from the word embeddings.
- S28 Proposal and validation of a pilot program to increase interest in AI, contextualize technically rigorous AI concepts through societal impact, and address barriers that could discourage girls from pursuing computer science and AI.
- S29 Proposal and validation of the use of artificial intelligence chat agent (chatbot) to answer questions about sex, drugs, and alcohol.