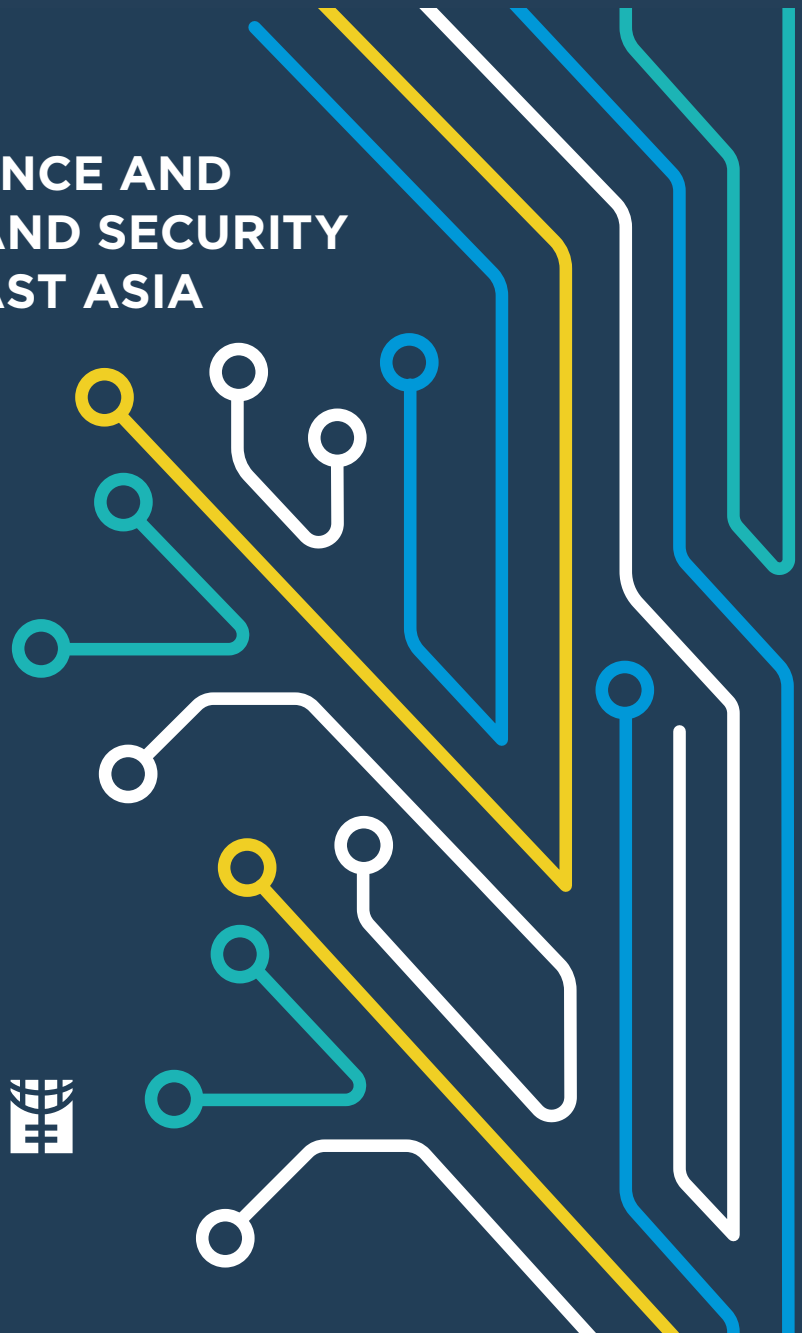


# ARTIFICIAL INTELLIGENCE AND THE WOMEN, PEACE AND SECURITY AGENDA IN SOUTH-EAST ASIA



## Background

The global mainstreaming of artificial intelligence technologies (AI) has had an important impact on the security of women. While AI can be used for peacebuilding purposes, it is used unequally across genders, with women much more impacted than men by the digital divide.

At the same time, AI systems have been shown to pose security risks to women, particularly in relation to online harms (such as cyberbullying, cyberstalking, doxxing, misogynistic hate-speech and other forms of harassment), dis- and misinformation, and privacy. As

South-East Asian countries begin to develop strategies and regulations for AI, they often have not taken into consideration gender risks in AI, which can be obfuscated under the projected economic potential of the technologies.

Investment in AI technologies is increasing in the South-East Asian region, although at an uneven pace, with some countries positioning themselves as global leaders in AI, while others having a much more limited capacity. It is projected, however, that AI will add USD 1 trillion to the gross domestic product of South-East Asian countries by 2030. In this context, understanding the impact of these technologies on the Women, Peace and Security (WPS) Agenda is critical to supporting South-East Asian countries to regulate the technologies and mitigate their risks.



## The Research

This research conducted by UN Women and the United Nations University Institute in Macau examines the opportunities and risks of AI from a WPS lens in South-East Asia, with a focus on four types of gender biases in AI which will need to be addressed before the region can fully benefit from new technological developments. These are discrimination, stereotyping, exclusion, and insecurity (as set out below). Understanding

these and mitigating them will be an important step in developing a safe and trustworthy AI ecosystem.

The objective of this research is to inform government, policy makers, international organisations, bilateral partners, the private sector and civil society organisations (CSOs) interested in AI in the region on the risks and opportunities which these technologies bear, addressing an important gap in understanding on this issue. Given existing lacunes in AI and the WPS agenda also at the global level, this research will also contribute to the global conversation on ethics and norms.

---

### Discrimination

in which AI systems provide different outputs for women than for men, based only on their gender

---

### Stereotyping

where AI systems produce outputs that promote an image of women as inferior, sexualized or hateful

---

### Exclusion

in which women face often-insurmountable barriers to participation in the development and governance of AI or that inhibit their access to its benefits

---

### Insecurity

in which AI systems threaten women's psychological or physical safety.

---

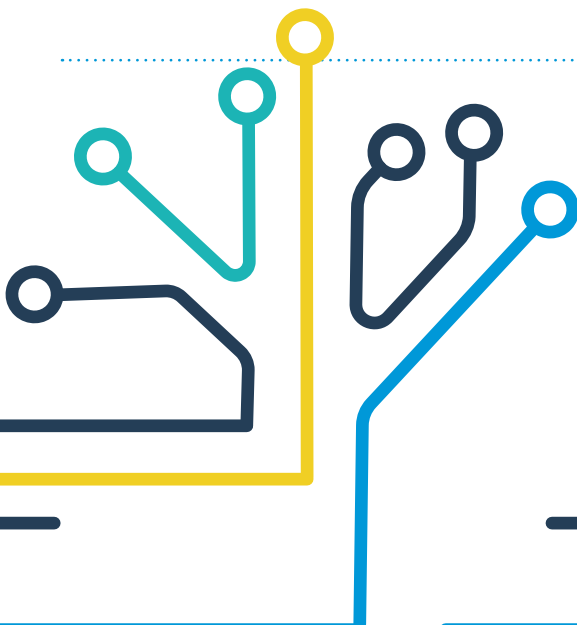




## Key Findings

The relationship between AI and WPS was categorised according to three types of AI and its applications: AI for Peace, Neutral AI, and AI for Conflict (see table). The research finds that across these categories, there are favourable and unfavourable effects of AI for gender-responsive peace and women’s agency in peace efforts.

TYPE OF AI	DEFINITION	APPLICATIONS
<i>AI for Peace</i>	AI technology is built for the purposes of responding to conflict or reducing the likelihood of future conflict.	<ol style="list-style-type: none"> <li>1. Data analysis and forecasting for conflict prevention, response and recovery</li> <li>b. Early-warning and early-response systems</li> <li>c. Migration and trafficking-in-persons systems</li> <li>d. Applications developed to directly support peace efforts</li> <li>5. Deliberative AI for large-scale digital dialogues               <ol style="list-style-type: none"> <li>f. Fact-checking and countering disinformation</li> <li>g. Chatbots providing legal and other support</li> <li>h. Mobile applications to support women CSOs and women human rights defenders</li> </ol> </li> </ol>
<i>Neutral AI</i>	AI technology is built neither for peace nor conflict, but nevertheless can impact them.	<ol style="list-style-type: none"> <li>9. Social media, news, and search recommendation systems and filter bubbles</li> <li>10. Generative AI</li> <li>11. Misinformation content</li> </ol>
<i>AI for Conflict</i>	The objective of the technology is to engage in physical or virtual conflict.	<ol style="list-style-type: none"> <li>12. Autonomous weapons systems</li> <li>13. Disinformation content (especially generated by Violent Extremist Groups)</li> <li>14. Surveillance without knowledge or consent (as in the case of human rights activists)</li> </ol>



## AI for Peace

- > Using AI for peace purposes can have many benefits, such as improving the effectiveness of conflict prevention, humanitarian, peacekeeping and post-conflict interventions; tracking evidence of human trafficking, conflict and human rights breaches; and protecting the human security of women. However, AI developed for peaceful purposes can also have unintentional risks. Primarily, these risks occur when gender impact assessments have not been conducted, or if the assessments were conducted improperly. AI tools have a tendency to have higher rates of error for women – facial recognition systems, for example, have been shown to misidentify women of colour considerably more often than white men, and therefore need to be evaluated in order to avoid bias to the extent possible.
- > There are systemic issues that place women's security at risk when an AI technology is adopted. These include threats to privacy and cybersecurity, which can lead to women's personal data being published online. Women are commonly subjected to doxxing – publishing personal data online without someone's consent – in South-East Asia, especially journalists or human rights defenders who might be challenging the status quo. This phenomenon warrants extreme precaution when collecting and creating data for AI tools, even those built for peace.
- > The dual trends of techno-solutionism and the digital divide lead to technologies being adopted that are not mature enough to be considered safe, have too many risks compared to their benefits, or do not benefit marginalised women. The digital divide is more important in South-East Asia for women who are socioeconomically disadvantaged. When AI tools do have potential benefits for women's security, such as security-protecting mobile applications for women, those that might most benefit from them are often excluded from their development.

## Neutral AI

- > Unlike AI for Peace, which tends to suffer unintended security risks to women and gender diverse persons, neutral AI often threatens women's peace and security intentionally. This is because, while the technology was not originally developed with peace or conflict in mind, it can be manipulated for purposes adverse to values upheld by the WPS agenda. On the other hand, these tools can be harnessed to support WPS objectives, notably by CSOs who use them for communications and advocacy.
- > Different actors have used social media to threaten women, and even to propagate stereotypes about women to foster conflict. Moreover, the report notes increasing concerns on the effects of new generative AI technologies in South-East Asia. It has been found that Large Language Models, such as ChatGPT, propagate false narratives after having been given misleading information 80 per cent of the time. They can also easily be prompted to create hateful or misogynistic content, including for recruitment into violent extremist groups. Similarly, image generating AI tools have a tendency to create sexualised bodies for women, even when unprompted. Advances in video generating tools are also a significant concern for deepfake pornography, which has been used to threaten the reputation and safety of women in the region.
- > Concerns were raised about the anonymity of users and the lack of accountability of social media platforms, which have resulted in online harms directed towards women being largely unchecked in South-East Asia. This is compounded by the fact that automated content moderation systems work much less effectively, if at all, on content written in South-East Asian languages, while human content moderators also rarely speak the languages.

## AI for Conflict

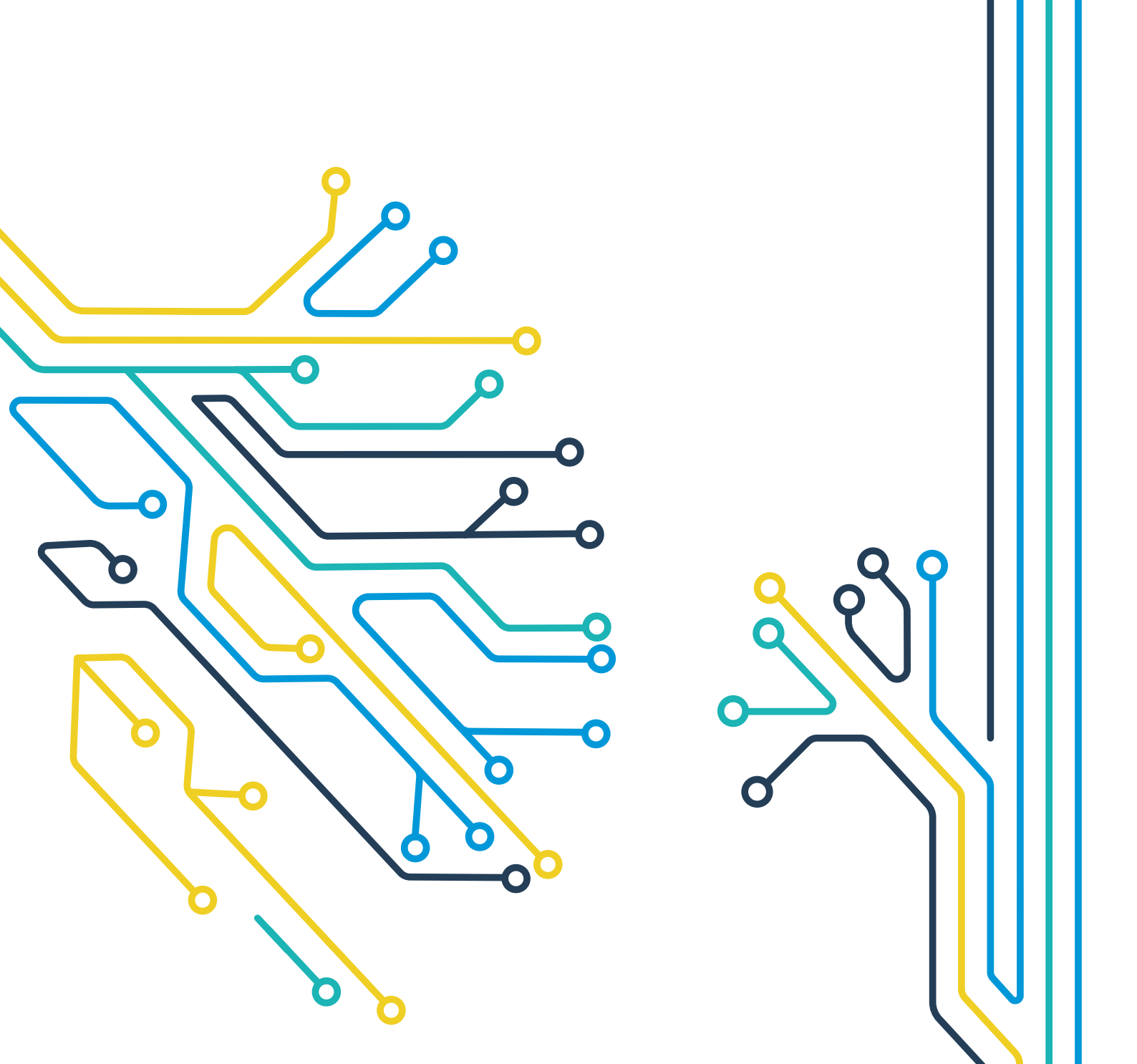
- > Using AI for Conflict has no benefits from a WPS perspective, but rather several significant risks. Automated Weapons Systems (AWS) are one of the prime examples of AI developed for usage in conflict and defence operations. In fact, AWS are developed as weapons that do not require human control, but rather have programmed target profiles that they strike when they are identified. Women's rights organisations have raised several concerns in relation to AWS, namely that they are built on AI systems that already discriminate against women, are not designed with a gender perspective, and could accidentally target civilian women due to algorithmic errors. At the moment, there is nothing preventing using these tools to sequester women, or to target people based on their sexual orientation.
- > AI systems have also been developed to surveil human rights defenders, including women human rights defenders; an increasing concern in the case of smart cities, which often include extensive video surveillance and tracking of inhabitants. This use of AI is of concern to CSOs who fear that increased investment in smart city and other digital transformation initiatives these risks into account. Surveillance can lead to harassment and control of women and can have repercussions on their safety both on and offline.
- > There are several broader opportunities for mitigation of the risks of AI to WPS and fostering the development of AI at the service of WPS. These include continued investment in hate-speech monitoring and fact-checking on social media, particularly in local languages; awareness-raising of other risks and opportunities of AI tools for WPS; conducting gender-impact assessments of peacebuilding tools, including early warning and migration management tools; and fostering multi-stakeholder dialogue on the issue in the region.



## Recommendations

There are two dimensions to improving the dynamics of AI and the WPS agenda in the region. The first is mitigating the risks of AI systems to advancing the WPS agenda, especially on social media, but also on other tools, such as chatbots and mobile applications. The second is fostering the development of AI tools built explicitly to support gender-responsive peace in line with WPS commitments. These are broken down into six distinct recommendations:

1. Mainstream WPS considerations in national, regional and global dialogues on AI governance, and vice versa.
2. Support the design of inclusive, conflict-sensitive and gender-responsive AI by ensuring that women have equal opportunities to lead and meaningfully participate in said processes.
3. Map and conduct gender- and human rights impact assessments of AI systems and draft policies and legislation, including those relevant to the advancements of peace and security.
4. Raise awareness and strengthen capacities of key stakeholders on risks and opportunities of AI tools for gender-responsive conflict prevention and peace efforts.
5. Leverage AI for hate-speech monitoring, fact-checking and countering disinformation on social media, accounting for the use of misogynistic and otherwise harmful gendered narratives.
6. Strengthen accountability mechanisms for social media companies by enhancing users' agency in choice of providers and platforms, with specific attention to inclusive and rights-based solutions.



This brief summarises the report **'Artificial Intelligence and the Women, Peace and Security Agenda in South-East Asia'** written by a team led by Eleonore Fournier-Tombs, United Nations University Centre for Policy Research (UNU-CPR), composed of JeongHyun Lee, UNU Institute in Macau; Min Yang, UNU Institute in Macau; and Preeti Raghunath, University of Monash in Malaysia. Further contributions were provided by Jingbo Huang, Serge Stinckwich and Jaimee Stuart, from UNU Macau; and Gaëlle Demolis and Alexandra Håkansson Schmidt from UN Women. (<https://doi.org/10.17605/OSF.IO/H38WZ>)

Thank you to the Government of Australia, under the Cyber and Critical Tech Cooperation Program of the Department of Foreign Affairs and Trade, and to the Government of the Republic of Korea for their support of this research and of the UN Women project entitled Women, Peace and Cybersecurity: Promoting Women's Peace and Security in the Digital World.

May 2024



Ministry of Gender Equality  
and Family



Australian Government

