Address AI's dangers and geo-political implications; integrate ethics in policies, say experts

New Delhi, May 20

Artificial Intelligence (AI) developers and policymakers need to address the dangers stemming from deepfakes and biases impacting privacy as well as the emerging technology's damaging geo-political implications and psychological consequences, according to experts.

Speaking at a webinar organised by the New Delhi-based research institute RIS on 'AI ethics and responsible AI', Ms Renata Dessallien, UN Resident Coordinator in India, said AI increasingly has geo-political implications as the emerging technology is disrupting the current distribution of power between countries, adding that this important issue was not currently getting adequate attention. Ms Dessallien said, "There is no question that AI will throw up into the air the current distribution of power. It not only about power between different groups, but also between countries and we are talking geo-politics. It is also between countries and private sector as we now have tech companies bigger than the GDP of many countries."

Addressing the same event, Professor V. Kamakoti, Department of Computer Science and Engineering, IIT Madras, talked about the dangers of AI, including certain instances where the use of the technology was not fair and responsible. "In some of the earlier deployment of AI, there was a bias in identifying criminals based on colour. This was well reported to an extent that some of the major tech companies stopped selling their face recognition software to police. So, AI has to be fair. Also, there is software such as deep-fakes which started creating a methodology where criminals can create fake videos. So, AI has to be responsible," he said. Professor Kamakoti emphasised that future of AI depends on making AI 'sane' and promoting responsible uses through interpretable and transparent AI systems. He also highlighted green AI and stressed the need to understand that AI can be localized and is an evolving technology.

According to Ms Dessallien, another crucial issue that should be given focus was how AI was dividing societies. "Look at the echo chambers that are eroding common grounds between people in the same country that used to be able to talk to each other and now hardly do so because AI algorithms are sent into these little camps and tribal groups are fed different kinds of diets in their groups and the end is a complete erosion of common ground and common ability to discuss and debate dispassionately issues of core importance to all of us," she said.

In addition, the AI also has psychological consequences, she said. "The attention economy is constantly working to try to grab and keep our attention on it as long as possible. It is making us distracted, leading to addictions not only in young people, but also in many other age groups as well," Ms Dessallien said. There are beginnings of studies showing how young people are feeling isolated, depressed, suicidal or doing acts of self-harm, she said. Therefore, these bigger AI issues that go beyond privacy and bias are equally important, but on occasions, even more important, she said.

Professor Sachin Chaturvedi, Director General, RIS, said the access, equity and inclusion framework and the humanitarian value system are important while looking at the ethical aspects of AI that have an economic imperative for the developing world. Inclusive and responsible AI is vital to ensure that no one is left behind and to achieve the UN Sustainable Development Goals by 2030, he added.

Professor Bernd Stahl, Director, Centre for Computing and Social Responsibility, De Montfort University, UK, said in order to get a handle on the ethics of AI, "it was important to not think about an individual technology, but conceptualise AI as a set of interlocking ecosystems which are driven and enabled to a certain degree by AI." He said the way to ensure that these technologies are conducive to human flourishing is to understand it from the nature of the technology as well as from the cultural and legal context.

Mr Santosh K. Misra, IAS, CEO, Tamil Nadu e-Governance Agency, said the government has an intertwining role of being the promoter of AI and also of being the regulator and user of AI. He said the state government has brought out a policy on ethical AI. It was probably the first in the country and even among the few government to do so, the official said. According to the policy, any entity rolling out AI solution in public domain will have to follow the principle of 'DEEPMAX', which stands for diversity, equity, ethics, privacy, misuse protection, accountability, and cross geography application and then decide whether it is safe to roll out, he added.

Dr Grace Eden, Assistant Professor, Indraprastha Institute of Information Technology, Delhi, Mr Ameen Jauhar, Senior Resident Fellow, Vidhi Centre for Legal Policy, Mr Rohit Satish, Consultant, NITI Aayog and Senior Fellow, Wadhwani Institute of Artificial Intelligence, Ms Vidushi Marda, Senior Programme Officer, Article 19 and Non-resident Research Analyst, Carnegie India, and Dr Krishna Ravi Srinivas, Consultant, RIS, spoke on the occasion.

Various governments and intergovernmental organisations have developed/are developing certain set of AI principles which intend to address ethical issues around AI. OECD, G20, EU, UNESCO, IEEE, WEF and many national governments have either formulated frameworks related to ethics of AI or are in the process of finalising such a framework.

The NITI Aayog document on "Responsible AI #AIforALL: Approach Document for India", seven broad principles for responsible management of AI have been identified, viz. safety and reliability; equality; inclusivity and non-discrimination; privacy and security; transparency; accountability; and protection and reinforcement of positive human values. There are considerable convergences and divergences between responsible AI principles and the AI ethics principles, proposed/implemented by various other governments and intergovernmental organisations. For example, the EU has proposed a set of regulations to govern AI and in this EU is trying to balance core values espoused by EU with applications of AI and to develop a trustworthy AI in/for Europe. (ENDS)