

ABSTRACT

The impact of artificial intelligence (AI) already goes beyond technological wonders into a social phenomenon. In this regard, there is a need to apprehend AI-related risks, such as ethical consideration and governance of AI, which have led many countries to publish their national strategies to capture the benefits and challenges that AI technologies can offer. These strategic documents provide a rich source of evidence regarding AI development and implementation. Therefore, this study focuses on answering whether G20 members have applied social responsibility regarding AI and what roles the state plays in the governance of AI through national strategy documents. By applying quantitative and qualitative content analysis, this paper comprehensively examines the national strategy documents of sixteen G20 members using Floridi and Josh Cowls's work, *A Unified Framework of Five Principles for AI in Society*, and Borrás & Edler's framework regarding the roles of the state in the transformation of socio-technical systems. This paper reports on (1) both frameworks' analytical and theoretical usefulness, (2) the *autonomy* principle gaining the lowest score, which indicates an inadequate promotion of human-centric approaches over machines/technologies and novel information regarding the fear that AI will ultimately take over humanity, (3) the three most dominant roles of the state as a *facilitator*, an *enabler of societal engagement*, and a *promoter*, (4) a distinction in approach between developed and developing countries towards the development and adoption of AI, and (5) a higher level of socially responsible AI as indicated by a close public-private collaboration. Policy implications, as well as limitations and future research, are discussed.

Keywords: artificial intelligence (AI), socially responsible AI, ethics, roles of the state