



## SUMMARY OF DISCUSSIONS FROM THE G20 AI DIALOGUE IN 2020

### Introduction

In 2020, the Saudi G20 Presidency conducted a G20 AI Dialogue (*the Dialogue*) under the Digital Economy Task Force (DETF), as part of its efforts to advance the G20 AI Principles. The Dialogue aimed to explore practical examples of why the AI Principles have an important role to play in achieving better policies for AI. Against the backdrop of the coronavirus (COVID-19) pandemic, the Dialogue took the form of a multi-stakeholder event on “Trustworthy AI in Pandemic Response” on 1 June, 2020, as well as DETF responses to short questionnaires on “AI and Health” and “AI and Education”. The Dialogue benefitted from three background documents prepared by the OECD – the “G20 Digital Economy Task Force Dialogue on Trustworthy AI in Pandemic Response: Discussion Note”, and reports on “Trustworthy Artificial Intelligence in Health” and “Trustworthy Artificial Intelligence in Education: Promises and Challenges”. Elements of the Dialogue also informed the “Examples of National Policies to Advance the G20 AI Principles: Background Report” prepared by the OECD.

This summary draws out key points and commonalities from discussions and inputs pertaining to the three elements of the 2020 Dialogue, particularly regarding the relevance and approaches of applying AI in a trustworthy manner, the associated challenges, and the potential role of government consistent with the G20 AI Principles in this context.

### Examples of uses of AI in health, pandemic responses and education

Over the course of the Dialogue, it was shown that AI is already being applied in the context of health, pandemic responses and education; but that even greater potential lies ahead for this technology to contribute to people’s well-being.

The use of **AI in health** in everyday practice still seems to be extremely limited, yet the number of potential applications is growing by the minute. Applications range from clinical settings, to biomedical research, to health system administration and management. Almost every aspect of health care delivery seems amenable to AI use and implementation. Much promise concerns the application of AI for automating diagnostic processes, clinical decision-making and a range of other clinical applications. AI tools may perform as well as – or even better than – average clinicians, in areas ranging from retinal scans to tumour detection in radiology. AI is being used in the surgical field to aggregate diverse sources of information to predict consequences of surgical decisions and to improve safety of interventions. AI tools can also affect the daily workflow of a health care practice by assigning priority based on appropriateness criteria, while another prime target for AI tools is patient record reviews (or clinical audits). Meanwhile, the biomedical research setting is proving fertile ground for AI, where it might exponentially increase the probability for new drug discovery including important drugs such as novel antibiotics and antivirals. AI models based on large and varied personal health data are a key opportunity to make precision medicine a reality. Finally, the application of AI in the “back office” of the healthcare sector (e.g. scheduling, billing, coding, intervention authorisations, payment) could have sizeable impacts, through undertaking tasks in a more efficient, accurate and unbiased fashion.

With respect to the **current COVID-19 pandemic**, AI is being used to understand SARS-CoV-2 and accelerate medical research on drugs and treatments; detect and diagnose the virus, and predict its evolution; assist in preventing or attempting to slow the virus' spread through surveillance and contact tracing; respond to the health crisis through personalised information and learning; monitor the recovery and improve early warning tools. Notably, AI algorithms contributed to the automated global early warning systems that first spotted SARS-CoV-2 (the strain of coronavirus causing COVID-19). New diagnostic tools based on deep learning are helping overwhelmed health care providers identify cases of COVID-19 using scans of patients’ lungs. Machine learning is providing insights into how old and new drugs might be effective in halting viral replication, as well as enhancing randomised controlled trials (RCTs) to test new treatments. Several governments already consider AI as an important part of their pandemic response toolkit, including applications in online medical consultation and “chat-bots”. AI is also being harnessed to help manage scarce

health resources, through data analytics and its contribution to more proactive management of the public health emergency.

AI has potential to drive innovation in the classroom as well as at the **education** system level and its diffusion and uptake in the sector may help to achieve some of the global educational targets identified by the international community in Sustainable Development Goal 4. AI applications for education are still nascent and are embodied more in solutions for individuals than for schools or governments. A key example is personalised learning, which harnesses AI to help identify teaching materials and approaches adapted to the level of individual students, and make predictions, recommendations and decisions about the next steps of the learning process. At the system and school management level, AI is now helping power predictive and diagnostic models to support decision making, for instance in analysing data to provide early warning on children at risk of dropping out of the education system. AI also helps deliver game-based assessment and simulations to test complex skills that are difficult to assess with standardised testing. AI systems may also help students with disabilities, for instance using AI-enabled virtual or augmented reality to assist students with hearing or visual impairments, or mental health issues.

### Challenges in these applications

The Dialogue revealed that deployment of AI in health, education and for pandemic responses raises new challenges for health professionals, educators and policy makers. Many of those challenges relate back to trust and to shaping a trustworthy use of AI, in line with the key objective of the G20 AI Principles. In particular, health and education systems play a critical role in people's lives and the opportunities they enjoy, making transparency, explainability and accountability of AI systems a top priority.

In the context of health and pandemics, Dialogue participants spoke of the need for policy makers to **address a privacy-transparency challenge**, in terms of respecting patients' data while also leveraging the full potential of this data to inform AI applications. Similar sentiments were raised with respect to AI in education, where personal data enhances the effectiveness of AI systems but where the collection and retention of data raises privacy concerns. Families may also be concerned that educational institutions (and employers) will use "old" data to make decisions, and in the case of both education and health, there are questions around the commercial use of data in sectors that have often had a strong public element. Lack of trust among patients, students, the public, data custodians and other stakeholders, in how data are used and protected is a major impediment to data use and sharing.

The **quality and availability of data** was also a cross-cutting theme. Most AI today is artificial narrow intelligence, designed to accomplish specific tasks on previously curated data from a single context. In the real world, data are unstandardized, populations are diverse, and biased decision-makers make mistakes that are then reflected in data. This becomes especially problematic in fast-moving crises. Large curated data sets may not exist to adequately train AI algorithms – or, alternatively, may not adequately represent people's changing behaviours. It was raised that, for a number of reasons (e.g. representativeness and breadth of input data), many applications of AI could potentially benefit from cross-border collaboration in the processing of personal data for purposes that serve the public interest, while safeguarding privacy.

The **need for capacity building and acknowledging the human element** in diffusion and uptake of AI was also considered a common challenge. Health workers, for instance, may already be stretched and suffering from "change fatigue", and the black-box nature of AI algorithms may result in either resistance from clinicians to adopt and use their predictions, or a blanket acceptance of their outputs with little critical assessment of the potential for biased and suboptimal predictions. The current health crisis has revealed a need to bridge technological-pedagogical gaps in the education sector through better preparation and development. More broadly there is a need to increase data science and other relevant skills.

With deployment of AI still at relatively early stages, especially in responding to pandemics, Dialogue participants observed there is still **limited evidence about the effectiveness of many AI solutions**. They saw scope for more R&D and evaluation to build a stronger evidence base, especially

in these human-focused sectors where the need for trustworthiness, both in the technology and its use, is high. Importantly, there are significant differences across countries in their preparedness for and ability to achieve greater uptake of AI in health and education and for pandemic responses.

### **The role of government in advancing trustworthy AI**

The Dialogue highlighted the role of the G20 AI Principles in promoting responsible stewardship of trustworthy AI, from responding to crises to improving outcomes in health and education. Prominent threads in discussion were the importance of data and associated data infrastructure to fuel AI, and the need for countries to develop digital infrastructure for connectivity, raise human capacities, boost R&D and build trust. These issues mirror the principles and policy recommendations of the G20 AI Principles.

A number of **suggested roles for government** consistent with the G20 AI Principles were voiced over the course of the Dialogue, including:

- Promoting multi-stakeholder dialogue, to facilitate coherent responses to common challenges, share vital information and data, and draw in all stakeholders to ensure AI adds value to society.
- Seeking interoperability and collaboration across countries, especially to avoid widening digital divides.
- Bringing greater clarity to trade-offs, especially if some principles (e.g. privacy) are given higher priority than others (e.g. transparency) in developing AI applications under pressure.
- Supporting by open-sourcing government data, and working to improve data quality, interoperability and access in a secure way through better data governance.
- Setting an appropriate regulatory environment, including one that facilitates testing of new approaches and business models.
- Encouraging innovation, so that AI can make an even stronger contribution to tackling future crises and delivering high-quality, effective and resilient health and education services.
- Boosting monitoring and impact assessment of AI applications, especially those rolled out at speed in crises, to help inform future policy directions and build trust with populations.

There may be an opportunity for policymakers to set a trajectory for trustworthy AI that helps countries better face the next unknown health, environmental or other global challenge, including through building stronger and more effective health and education sectors. There seems to be a clear role for governments in shaping this trajectory, building on the G20 AI Principles that provide a shared vision of AI that advances innovation and respects human rights and democratic values.