ACADEMIES FOR GLOBAL INNOVATION AND DIGITAL ETHICS (AGIDE)

A PROJECT UNTER THE AUSPICES OF THE AUSTRIAN ACADEMY OF SCIENCES AND IN COOPERATION WITH ACADEMIES OF SCIENCES WORLDWIDE

EXECUTIVE SUMMARY

The digital transformation has brought about an unprecedented degree of global interconnectedness, accompanied by increasing efforts to formulate universal ethical guidelines for dealing with emerging digital technologies. The relative ease with which countries around the world seem to agree on universal action-guiding principles of digital ethics along the lines of „fairness", „transparency" and „accountability" seems to contrast sharply with the vast differences in technology adoption that we see around the world, and also the vast differences in attitudes towards technology. These differences can influence the salience of ethical concerns in different settings. The project AGIDE, which stands for „Academies for Global Innovation and Digital Ethics", seeks to embrace this diversity of perspectives on digital ethics, exploring where differences lie and how these differences might be conceptualised beyond existing stereotypes.

In order to get a clearer picture of the key issues, similarities and differences, the starting point was over 75 qualitative interviews with expert voices from around the world, as well as deep dives in the course of three workshops held in April, June and October 2023. Taking a „situated" approach and considering local knowledge contexts, the results were then analysed and synthesised by an international working group comprising representatives from 10 academies of sciences on six different continents, as well as other eminent experts from around the world.

The AGIDE project found that there is a remarkable consistency in core values (such as „justice", „dignity" or „privacy") across different regions of the world. The fact that agreement has been reached on common principles and guidelines, such as the UNESCO Recommendations on the Ethics of Artificial Intelligence or the OECD Principles on Artificial Intelligence, is arguably a result of this fundamental agreement on shared core values. Contrary to initial expectations that the differences in perception and governance of digital opportunities and risks explored by AGIDE might result from discernible differences in emphasis on particular values, the data collected did not support such distinctions.

AGIDE's research showed that major differences lie in something else: in the narratives of digital ethics. Narratives are stories that are told repeatedly, consisting of a series of events that are selected and arranged in a particular order, often including central characters (protagonists, antagonists), a conflict and a plot. When narratives become dominant in a particular social setting, because they are shared by a wider group and/or promoted by influential actors, they can become powerful drivers of collective behaviour, and they shape how core values are operationalised and put into practice. Findings from the AGIDE project suggest that, although core values were widely shared, digital ethics discourses in different world regions were dominated by different narratives about how these values are challenged, or how they can be protected, and why and how these values matter in the first place.

The data collected in the AGIDE framework was fitted into a matrix according to criteria that emerged from the interviews and workshops. This matrix attempts to categorise the various views and concerns expressed in relation to eight salient aspects, including underpinning ethical

approaches, the primary point of reference (for example, the individual or the community) and its position vis-à-vis technology (for example, whether it is primarily perceived as a victim, a beneficiary or an actor), the primary focus of ethical concern, and appropriate tools of governance. Taken together, the very particular approaches to these eight aspects result in characteristic patterns of storytelling that reflect distinct digital ethics narratives.

From a macro perspective, several characteristic narratives emerged, including what the report calls the „Coloniality" type narrative, the „Beneficiary" type narrative, the „Harmony-Opportunity" type narrative, the „Silicon Valley" type narrative, and the „GDPR" type narrative (with GDPR standing for the EU's General Data Protection Regulation). At the level of practice, these patterns of storytelling may shade into one another. Exploring the narratives from a micro perspective – looking at specific regions or populations – revealed a much greater diversity of patterns, highlighting the nuanced nature of these perspectives. There is no single answer as to what produces the differences in narratives. The data suggest that there is a range of different factors, and that intersectionality of these factors influences the choice of approaches.

It was particularly insightful to observe that many patterns of storytelling are fluid to some extent, and that there are remarkable shifts as far as dominant narratives in a country or region are concerned. However, some narratives seemed to be very deeply entrenched. The European Union (EU) was presented as an illustrative example where the „GDPR" type narrative seems embedded to an extent that might potentially prevent EU policymakers from effectively bringing about change.

Given the scope and timeframe of AGIDE, the findings highlight the need for further research to explore whether the narratives are causes or symptoms of the differences we perceive, or both. Further research is also needed on the factors that contribute to the development of specific narratives, both at the macro and micro level. Finally, we need to better understand the conditions that are conducive to the transformation of established narratives or that cause established narratives to resist even major shifts in the policies pursued, potentially hindering important policy changes.

AGIDE therefore hopes to contribute to a new global discourse and to policymaking that is attentive to differences across and within regions regarding digital ethics narratives, and, importantly, is alert to the significance of resources and power. Accepting and embracing differences means allowing for different interpretations and implementations of shared values, and being open to the possibility of the development of situated ethical codes, without forcing uniformity. Respecting equity, fairness and empowerment also means giving priority to enabling the conditions and possibilities for local implementation, including the development of genuinely local technologies, structures and solutions. Finally, understanding the deep connection between narratives, ethical principles and power negotiations also requires that ethical principles can be translated meaningfully into practice, backed up where necessary with policies and legislation.