

**City University of Hong Kong
Research Centre for Sustainable Hong Kong¹
Policy Paper No. 28**

**Ethical examination of the development of artificial intelligence
Chun-kit Cheuk, Ho-mun Chan, Victor Tin-yau Hung, Linda Chelan Li, Fen Lin,
Viktor Tuzov²**

March 2024

1. Introduction

The discussion on policy regulations for artificial intelligence technology has only started a few years ago. Due to the rapid development, various institutions and technology companies around the world have formulated more than 100 different versions of ethical principles and guidelines, ensuring that researchers and users can act in accordance with widely recognized ethical standards. However, the roadmap to implement these abstract principles into specific legislative and regulatory mechanisms still remains at the preliminary stage.

Artificial intelligence is an indispensable part of Hong Kong's future development. Starting in 2019, we have been conducting textual research, in-depth interviews questionnaire surveys³, and ultimately designed a research project aimed at identifying the ethical principles and value considerations of Hong Kong citizens regarding the application of artificial intelligence. Based on this, we propose a potential AI governance framework.

¹ Established in June 2017 by a cross-disciplinary research team, the Research Centre for Sustainable Hong Kong (CSHK) is an Applied Strategic Development Centre of City University of Hong Kong (CityU). The CSHK conducts impactful applied research with the mission to facilitate and enhance collaborations among the academic, industrial and professional service sectors, the community and the government. The Centre nurtures the collaboration between Hong Kong and other regions and is engaged in applied research in the fields of Hong Kong professional services, the Belt and Road, the Guangdong-Hong Kong-Macao Greater Bay Area, the green economy, COVID-19, etc. Research projects have been repeatedly funded, with a number of research reports, papers and books being published. For more information, please browse the CSHK webpage <http://www.cityu.edu.hk/cshk>. If you have any comments on this policy proposal, please feel free to email: sushkhub@cityu.edu.hk.

² Chun-kit Cheuk is a doctoral student in the Department of Public and International Affairs at City University of Hong Kong; Ho-mun Chan is a Professor in the Department of Public and International Affairs at City University of Hong Kong; Victor Tin-yau Hung is a special Associate Professor at the School of Economics and Management of the University of Hong Kong and an international and professional consultant at the Hong Kong Centre for Sustainable Development; Linda Chelan Li is a Professor in the Department of Public and International Affairs at City University of Hong Kong and the Director of the Centre for Sustainable Development in Hong Kong; Lin Fen is Associate Vice President (Global Strategy) and Associate Professor of Media and Communication at City University of Hong Kong; Viktor Tuzov is a doctoral student in the Department of Media and Communication at City University of Hong Kong; Ray Chan is a member of the Hong Kong Sustainable Development Research Hub.

³ Chun-kit Cheuk, Ho Mun Chan, Victor Hung, Linda Li, Fen Lin & Viktor Tuzov (2023) : "*Ethics and Governance of Artificial Interlligence In Hong Kong – Report*" , Research Centre for Sustainable Hong Kong, City University of Hong Kong.

We found from our 2022 questionnaire survey that under a decontextualized condition, Hong Kong citizens highlighted five ethical values important for the development of artificial intelligence, including: a.) transparency; b.) unbiasedness; c.) robustness; d.) personal privacy; and e.) personal freedom. All of them were comparably important for the survey respondents. In order to identify the potential differences in public perception of these values, we added several scenarios presented by different AI innovations. As a result, the ethical values of "personal privacy" and "personal freedom" have been found to be the most important ones among the proposed scenarios.

In this article, we will share our observations on the results of the research project and, based on this, propose several suggestions suitable for the governance of artificial intelligence development in Hong Kong.

2. Decontextualized ethical considerations

According to the definition of the Organization for Economic Cooperation and Development (OECD), artificial intelligence (Artificial Intelligence) refers to a machine-based system that can make predictions on human-defined goals, provide recommendations, or make decisions that affect real or virtual environments. Artificial Intelligence systems are designed to operate at different levels of autonomy.

In the process of the development of artificial intelligence, previous studies have examined different ethical principles and guidelines for artificial intelligence issued by the public sector, private enterprises and research institutions through text analysis. Among them, the research by ETH Zurich pointed out that transparency, justice, fairness, and privacy are the most widely adopted and, therefore, have greater importance than other values or principles⁴. Based on the conclusions of these texts and our in-depth interviews with 22 industry professionals, experts and scholars, we selected five ethical values as research objects. These five values can be explained as follows:

- 1) **System transparency:** Whether the operating principles and limitations of the intelligent system are clear and whether the system has sufficient disclosure of the source and information used.
- 2) **System's operation without bias:** Whether there is any form of bias or discrimination in the operation of the intelligent system, ensuring the fairness of its operation.
- 3) **System robustness:** The effectiveness and security of the system, including its ability to operate normally in the face of external instability, interference, and malicious attacks.
- 4) **Personal privacy:** How the system protects users' personal privacy data.
- 5) **Personal freedom:** The impact on users' freedom to make behavioral or lifestyle choices when using intelligent systems. The wider the range of choices, the greater the degree of freedom.

⁴ Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature machine intelligence*, 1(9), 389-399.

In March 2022, we successfully interviewed 510 Hong Kong citizens aged 18 or above through random telephone sampling. We first decontextualized the questionnaire, which means that respondents were asked to rate based on their own understanding without any pre-set contexts and explanations of the five ethical values mentioned above.

As a result, the respondents rated the five ethical principles in a similar way: ranked by mode, the highest score for all options was "7"; ranked by a median, all of the scores were "6"; and finally, ranked by mean, the difference between the lowest score "system's operation without bias" (5.47) and the highest score "system robustness" (5.75) was only 0.18. (See Table 1)

Table 1. Scores, calculated in different ways for various ethical values or principles
n=510

	System transparency	System's operation without bias	Protection of personal privacy	Protection of personal freedom	System robustness
Mode	7	7	7	7	7
Median	6	6	6	6	6
Mean	5.61	5.47	5.71	5.56	5.75

*1 point represents "very unimportant"; 7 points represent "completely important"

The above results reflect that in the absence of specific application scenarios, the respondents generally agreed on the importance of the above five ethical principles in the application of artificial intelligence and showed a tendency for maximization.

3. Ethical considerations in different contexts

We then added two different scenarios to identify potential users' trade-offs to be made while using the "Health Code System" and "Fraud Detection System". Each of the scenarios presented a situation unique to the use of a particular technology. After adding the "Health Code System" as an application example in Scenario 1, the proportion of respondents who think "personal privacy" is more important reached 58.8%, which is 35.5% higher than those who think "system's effectiveness" (23.3%) is more important. Switching to Scenario 2 with the "Fraud Detection System," the proportion of respondents who selected "personal privacy" as a more important value reached 60.4%, which was also about 35% higher than those who believed that "system's effectiveness" was more important. (See Table 2)

Table 2. Trade-offs between "personal privacy" and "system's effectiveness" in different application scenarios

n=510

	Scenario 1 Health Code System		Scenario 2 Fraud Detection System	
	Number of respondents	Proportion	Number of respondents	Proportion
(1) Personal privacy is more important	300	58.8%	307	60.4%
(2) Difficult to choose	91	17.8%	71	14.0%
(3) System's effectiveness is more important	119	23.3%	130	25.6%
The difference between (1) and (3)	--	35.5%	--	34.8%

In both situations, "personal privacy" plays a key role. In other words, even if the development of artificial intelligence could help prevent the spread of Covid-19 or even assist in the detection of fraud, in most people's eyes, these cannot be exchanged for "personal privacy."

Later, we used another value, "personal freedom," for a further comparison. In this situation, the results were not as straightforward. In Scenario 1 with the "Health Code System," the difference between the respondents who believe "personal freedom" is more important and those who believe "system's effectiveness" is more important narrows to 6.1%. In Scenario 2 presented by "Fraud Detection System", the gap between "personal freedom" and "system's effectiveness" increased to 27.3%. As for the newly added Scenario 3, represented by "Unmanned Cars," the respondents prioritized "personal freedom" over "system's effectiveness," with the gap ranging up to 15.2%. (See Table 3)

Table 3. The trade-offs between "personal freedom" and "system's effectiveness" in different application scenarios

n=510

	Scenario 1 Health Code System		Scenario 2 Fraud Detection System		Scenario 3 Self-driving Cars	
	Number of respondents	Proportion	Number of respondents	Proportion	Number of respondents	Proportion
(1) Personal freedom is more important	227	44.6%	284	55.8%	246	48.7%
(2) Difficult to choose	86	16.9%	80	15.7%	90	17.8%
(3) System's effectiveness is more important	196	38.5%	145	28.5%	169	33.5%
The difference between (1) and (3)	--	6.1%	--	27.3%	--	15.2%

Based on the results, we can propose an alternative interpretation here, that is, in some cases, "personal freedom" is not as important as "system's effectiveness," which ranges differently in the following scenarios: "Health Code System" (38.5%), "Self-driving Cars" (33.5%) and "Fraud Detection System" (28.5%). The potential harm to the safety of the individuals across these situations could be seen as a factor that impacts the result.

The survey was conducted during the pandemic period when the "Health Code System" was considered an important tool to prevent the spread of the virus. Therefore, a considerable proportion of people may be willing to give up their "personal freedom" (taking time to scan the barcodes when entering or leaving designated places and being denied entry if certain conditions are not met, such as vaccination) to make the health code system effective. In the context of "Self-driving Cars," we also provided a specific scenario. If faced with only two options: "hit pedestrians" or "turn to hit the tunnel wall and hit yourself," 51.4% of respondents would rather hit themselves, while only 30.7% said they would hit pedestrians.

Another point worth noting is that even though 38.5% of the respondents in Scenario 1 accepted that "system's effectiveness" is more important than "personal freedom," fewer people (23.3%) are willing to accept that "system's effectiveness" outweighs "personal privacy." This demonstrates the importance of personal privacy and reflects the public's belief that authorities can protect privacy by creating appropriate system settings.

4. Moral dilemmas and solutions

Behind different value trade-offs, respondents have varying opinions on whether to adopt artificial intelligence technology in individual scenarios. In the "Health Code" and "Fraud Detection System" scenarios, the difference between support and opposition to the technologies only ranges from 2.2% to 3.3%. As for the "Unmanned Cars" scenario, the majority of respondents opposed it. (See Table 4)

Table 4. *Artificial Intelligence adoption in different scenarios*
n=510

	Should it be adopted?	
	It should be	It shouldn't be
Scenario 1. Health Code System	49.9%	47.7%
Scenario 2. Fraud Detection System	45.6%	48.9%
Scenario 3. Unmanned Cars	29.4%	65.4%

The survey results in Table 4 are consistent with the social environment prevalent at that time. When the "Health Code" was launched in August 2020, supporters argued that it was an effective way to control the spread of the virus, arguing that it could identify people at different levels of risk. These measures were meant to lift other excessive social distance restrictions and help social and economic recovery. However, the opposing side was worried that the "Health Code" would be attached to whereabouts records and have access to personal information, which would potentially infringe on personal privacy and freedom. With the further development of artificial intelligence and improved computing accuracy, data access,

real-time monitoring of activities, and even software and hardware design will become increasingly refined. Hence, the above disputes will only become more acute.

We asked questions about possible moral dilemmas in different scenarios. Regardless of the application scenario, once moral dilemmas arise, respondents hope that the opinions of "affected citizens" will be prioritized. (See Table 5)

Table 5. *Whose opinion should be given priority when an ethical dilemma arises?*
n=510

	Scenario 1 Health Code System	Scenario 2 Fraud Detection System	Scenario 3 Self-driving Cars
1. Public institutions	19.4%	24.6%	8.8%
2. Private enterprises	2.0%	1.5%	1.7%
3. Third-party professionals	12.0%	14.5%	24.5%
4. Affected citizens	59.9%	53.5%	58.4%

It is worth noting that respondents also have a high degree of recognition for "third-party professionals." The percentage of respondents who think their opinions should be considered first ranges from 12% to 25%. Among them, confidence in "Self-driving Cars," a relatively new technology, is particularly high. On the contrary, respondents show lower confidence in the largest provider of services, "private enterprises." Only 1.5% to 2.0% of the respondents believe that the opinions of private enterprises should be considered first.

5. Summary

When we discuss ethical principles and policy norms for the development of artificial intelligence with different stakeholders in society, many people indicate these concepts as "abstract" and "vague", making it difficult to formulate a specific and operational framework. However, based on the above analysis, we believe that the main reason for such opinions may be that people often do not link ethical principles to particular practical situations in real life.

Our survey results show that, under decontextualized conditions, all ethical values are considered important, from the system's effectiveness to individual freedom and privacy. But the irony is that when all the options are equally important, society and policymakers will rather be more perplexed. Because, in most cases, every social policy requires different levels of trade-offs.

Therefore, ethical principles should be applied in specific contexts to play a normative role in policy formulation. This is why we included different scenarios in the questionnaire. The result also reflects that the ethical values previously considered equally important acquired necessary variation after adding different application scenarios. This opened the possibility of making judgments and trade-offs regarding a particular application situation.

In addition, when facing ethical dilemmas in policy-making, we were not surprised that respondents treated the affected entities as the most important objects of consultation. However, the survey also reflects that a considerable proportion of people believe that third-party professionals are an important communication channel that can help clarify the difficulties faced in developing artificial intelligence.

We believe that the above survey results and observations will help society formulate regulatory and normative measures suitable for the development of local artificial intelligence. We will leave further suggestions in the next article.