< >

Montréal Declaration Responsible AI_
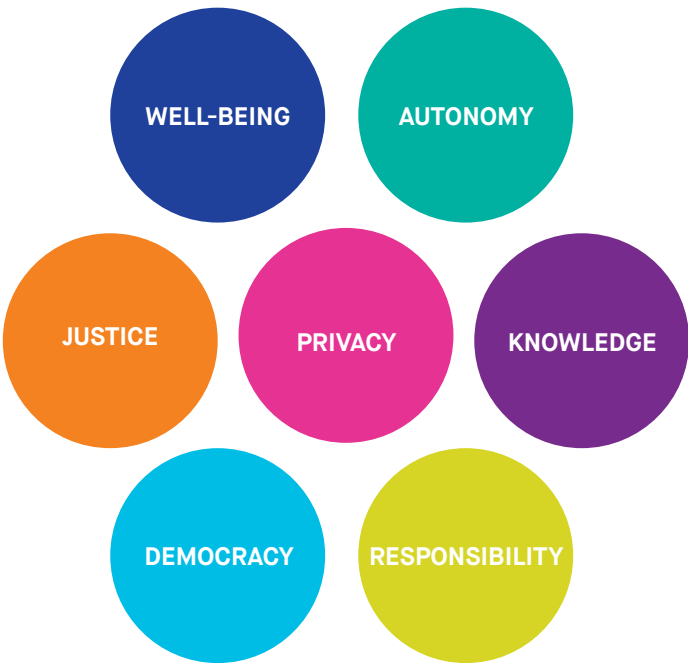
</ >

THE DELIBERATIONS AT A GLANCE

Excerpts from the results of the Montreal Responsible AI Declaration citizen deliberations

On November 3, 2017, Université de Montréal launched the co-construction process for the Montreal Declaration for a Responsible Development of Artificial Intelligence (*Montreal Declaration*). Eight months later, we present the first results of this citizen deliberation process that is now at the halfway point. It's a very favourable assessment: dozens of events were organized to spark discussions about the social issues raised by artificial intelligence (AI), and fifteen deliberation workshops were held over three months, involving over 500 citizens, experts and stakeholders from all horizons.

The *Montreal Declaration* is a collective work that aims to put AI development at the service of the individual and common good, and guide social change by making recommendations with a strong democratic legitimacy.

The selected citizen co-construction method relies on a preliminary declaration of general ethical principles articulated around fundamental values:

WELL-BEING

AUTONOMY

JUSTICE

PRIVACY

KNOWLEDGE

DEMOCRACY

RESPONSIBILITY

If one of the goals of the co-construction process is to fine-tune the ethical principles suggested in the preliminary version of the *Montreal Declaration*, an equally important goal consists of making recommendations to provide a framework for AI research, as well as its technological and industrial development.

FIRST, WHAT IS AI?

Very briefly, AI consists of simulating certain learning processes of the human intelligence, to learn from it and replicate it. For example, discovering complex patterns among a large quantity of data, or reasoning in a probabilistic fashion, in order to sort information into categories, predict quantitative data, or aggregate data. These cognitive skills are the basis for other skills such as choosing among several possible actions to reach a goal, interpret an image or a sound, predict a behaviour, anticipate an event, diagnose a pathology, etc. These AI realizations rest on two elements: data and algorithms, series of instructions that perform a complex action.

TO CONCRETELY DISCUSS THE ETHICAL ISSUES OF AI, THE CO-CONSTRUCTION METHOD WORKSHOP relies on the preliminary version of the *Montreal Declaration*. Schematically, after deciding on the "why?" (which desirable ethical principles should be included in a declaration on the ethics of AI?), it's a matter of prospectively anticipating, along with the participants, how ethical issues around could arise in the coming years, in the fields of health, judicial system and predictive police, smart cities, education and culture, the workplace and public services. Then, we imagine how we could respond to these issues. For example, through a measure such as a sectorial certification, a new actor mediator, a form or a standard, through a public policy or research program.

Citizens and stakeholders therefore took part in the citizen café or entire co-construction days where they had the chance to debate prospective scenarios.

Other citizens choose to contribute to the reflection by filling out a questionnaire online or tabling a brief. The results of these specific initiatives will be discussed in the global report on the activities tied to the *Montreal Declaration*, which should be published in the fall of 2018.

Discussion table around the theme of the workplace, Musée de la civilisation, Quebec City, April 6, 2018

CO-CONSTRUCTION WORKSHOP RESULTS – THE GENERAL TRENDS

Generally speaking, the participants recognized that the arrival of AI came with important potential benefits. Namely, in their field of work, participants recognized the time savings that AI devices could bring. However, it was also mentioned that AI development had to be done with caution and right now to prevent abuse, although some consider the possibilities brought on by AI to still be limited.

The citizens highlighted the need to implement different mechanisms to ensure the quality, intelligibility, transparency and relevance of the information being communicated. They also discussed the difficulty of guaranteeing truly enlightened consent.

The great majority of the participants recognized the necessity to align public interests with private ones and prevent the apparition of monopolies, or limit the influence of corporations through more cohesive and legal measures.

The participants also recommended putting mechanisms in place that would come from and involve independent, trained people to favour the diversity and integration of those who are most vulnerable, and protect the mixed aspect of the lifestyles.

Whatever the use, the majority of the participants insisted on the fact that AI must remain a tool, and that the final decision must come from a human being.

Three members of the citizen deliberation and co-construction team during an activity at Mordecai-Richler Library, March 10, 2018

PRIORITIES ACCORDING TO THE *MONTREAL DECLARATION* PRINCIPLES
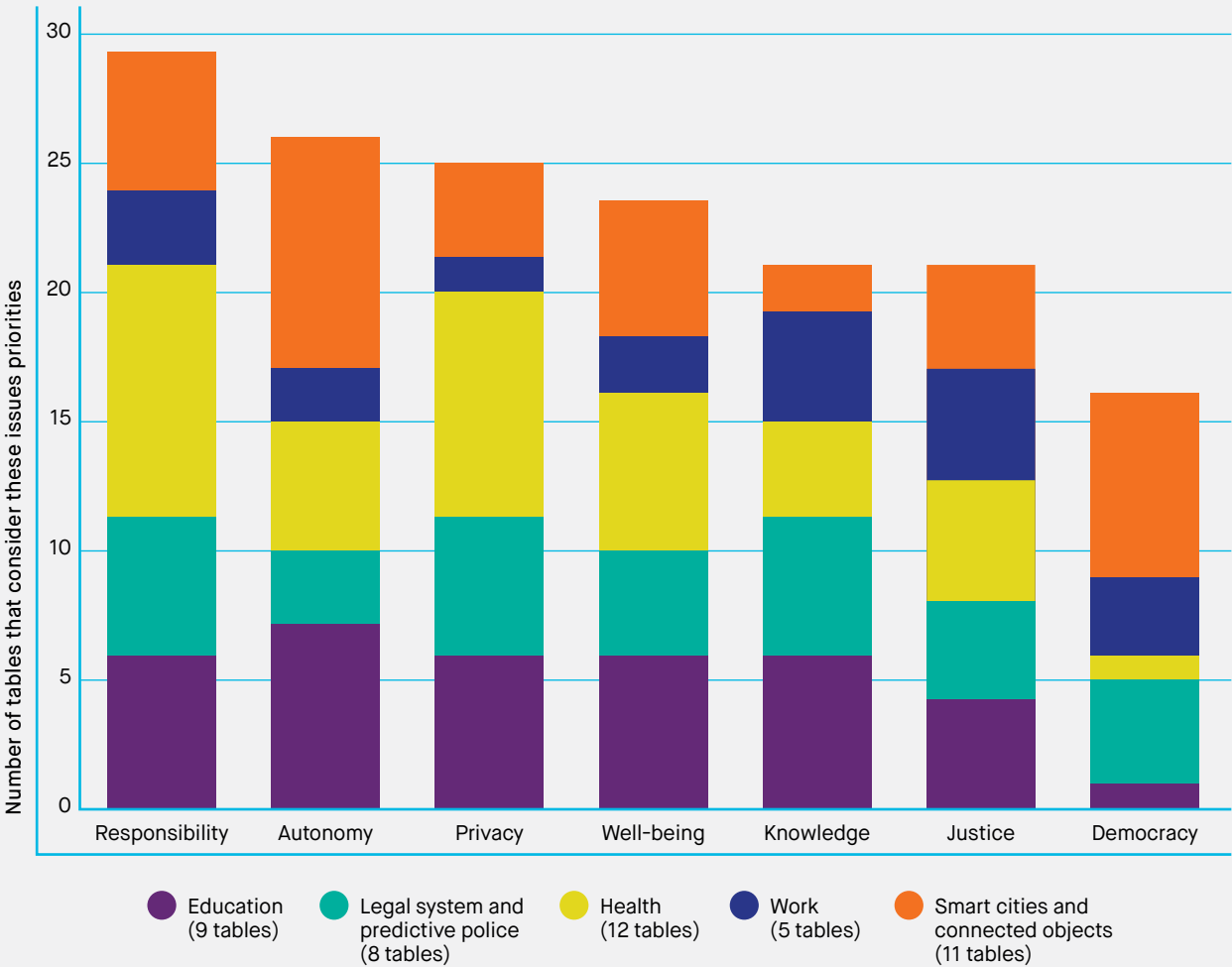
The responsibility principle has often been deemed the most pressing issue, followed by autonomy, privacy, then well-being (individual and collective), knowledge and justice. It's worth noting, however, that they are all closely linked.

As for the autonomy principle, which is often selected as a priority, it has to do with preservation, or even encouraging individual autonomy when faced with the risks of technological determinism and dependency on tools. It also raises the issue of a double liberty of choice: being able to follow your own choice when faced with an AI-guided decision, but also the choice not to use these tools without risking social exclusion.

The well-being principle is also an important one for participants. It is implicit at every table, illustrating a collective wish to move towards a just and equitable society that fosters the development of all.

In a general sense, the well-being principle has also taken on the form of a call to maintain a genuine human and emotional relationship between experts and users in every field.

Table 1: Principles referred to in the priorities identified by citizens.

Education (9 tables), Legal system and predictive police (8 tables), Health (12 tables), Work (5 tables), Smart cities and connected objects (11 tables)

ISSUES THAT CAN LEAD TO THE CREATION OF NEW PRINCIPLES, OR NEW THEMES TO EXPLORE AND DELIBERATE

The impact of the use and development of AI on the **environment** raises issues, namely on the way to guarantee the responsible and equitable use of material and natural resources.
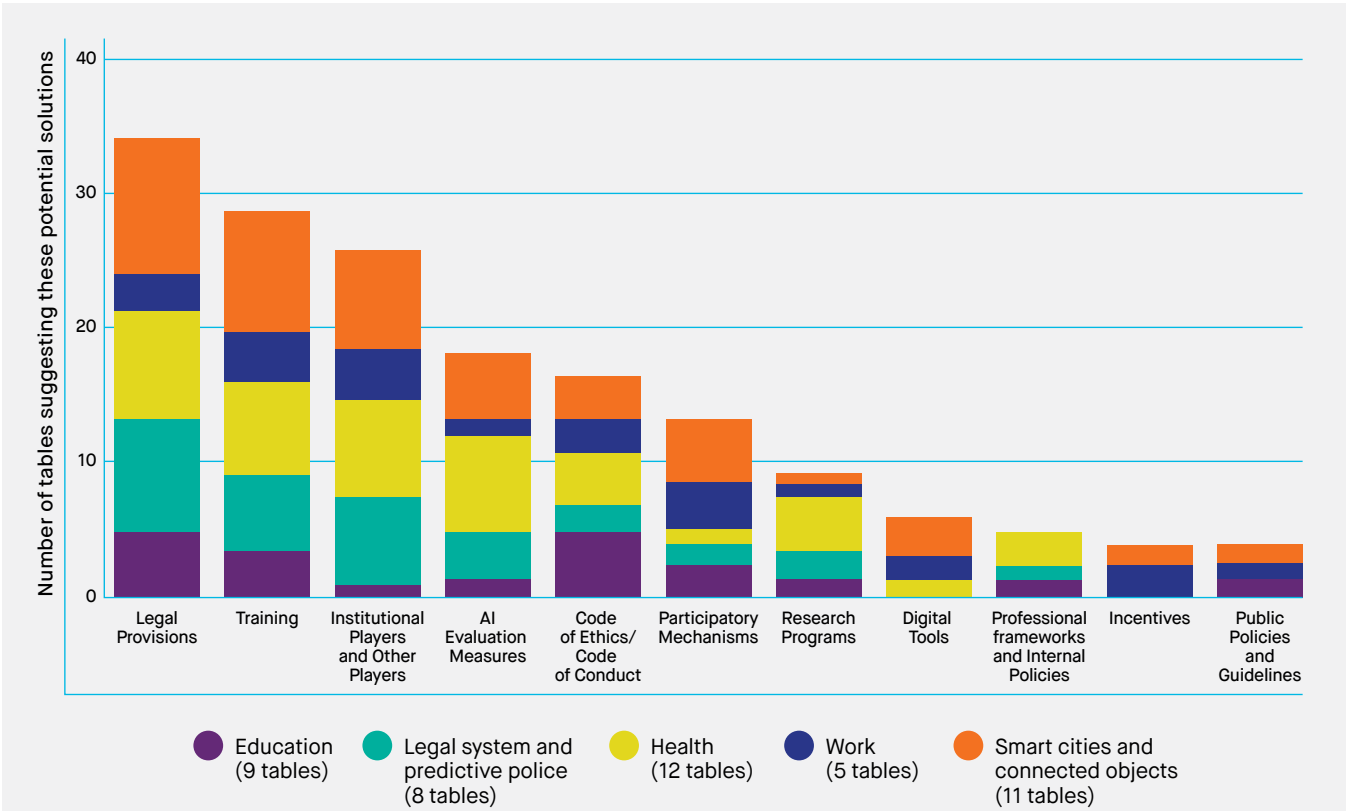
The justice principle was discussed on the basis of two types of issues, which could lead to two new principles: **a diversity principle** looking to avoid discrimination by finding bias-free mechanisms and **an equity or social justice principle**, which would require AI benefits to be accessible to all, and that the development of AI not contribute to the growing economic and social inequalities, but rather help bridge the gap.

A principle of caution: the issues related to the trust towards the development of AI technologies were regularly raised. This trust issue is also closely tied to the question of the reliability of AI systems.

A transparency principle: this principle implies being able to understand an algorithmic decision and react to it. For this, citizens think it's important that the algorithmic procedures be explainable so they can see and understand which criteria were considered in the decision.

Whatever the field, the citizens identified many issues regarding the relationship between human beings and AI.

Table 2: Potential solutions suggested to respond to the identified issues

POTENTIAL SOLUTIONS

All the participants of the co-construction tables agreed on **3 potential solutions** to guarantee socially responsible AI development, regardless of the field:

1. Legal provisions;
2. Training offered to all and
3. The identification of key independent players for AI management.

CONTINUE THE DELIBERATION

The *Montreal Declaration* project concentrated its first phase on five key sectors: education, health, work, smart city and judicial system and predictive police. An entire year of co-construction wouldn't even cover all the reflection themes. The co-construction initiative will therefore continue in September 2018, allowing for discussions about new themes that had barely been touched upon in the scenarios used in the co-construction phase. For example: environment, democracy, media and propaganda, as well as security and integrity.

We will present public policy recommendations around priority fields of action. To date, we can say that three fields of action have established themselves: digital literacy, diversity and inclusion, and transition and social mutations. The final results will be presented in December 2018.

All these transformations brought by the development of AI in different social spheres make us question ourselves, as citizens, on the society to build. At the heart of the tension between hopes and fears, it's the arrangements between humans and technologies that it is essential to observe and analyze in critical and prospective fashion. If one demand can truly be said to have been unanimous in the co-construction debates, it is indeed that of dedicating the central role to humans in a world that is more and more artificially intelligent.

"We must develop technologies that favour human autonomy and freedom of choice."

– A participant.

The full report on the first results of the Montreal Declaration Responsible AI deliberations is available on our web site:
www.montrealdeclaration-responsibleai.com/results-deliberations-june-2018

Montréal Declaration
Responsible AI_

www.montrealdeclaration-responsibleai.com