Commission of the Bishops' Conferences
of the European Union

COMECE

February 2019

Contribution to the Stakeholders Consultation on Draft AI Ethics Guidelines

A. Introduction

Rationale and Foresight of the Guidelines

The High-Level Expert Group explains in the introduction the character of these drafted guidelines. They underline that the guidelines will offer guidance to all relevant stakeholders concerning ethical challenges of AI, they will be not legally binding and they do not intend to be a substitute to any form of policy-making or regulation.

The authors define that the scope of the guidelines covers AI applications in general – although they are aware that a tailored approach is needed for the different challenges of AI context-specificity.

But we have to **criticize the undifferentiated use of the term "stakeholder".**

The drafted guidelines are mentioning the problem, that different situations raise different challenges. But as a result of this recognition it is **not convincing just to mention the necessity of differentiating** business-to-consumer or business-to-business or public to citizen in general, but not **differentiating the necessity of balancing out the different interests** in the wider context of the common good of these different stakeholders.

An ethical assessment has to deal with different factors and tools whether an AI developer, a user or the public sector is affected.

Chapter I

Respecting Fundamental Rights, Principles and Values – Ethical Purpose

We very much **welcome the underlining of the human-centric approach of AI in the EU in the drafted guidelines.** AI has to serve the common good**. AI has to serve the lives of all human beings**. It has to be considered that human life not only has a personal dimension but also a community dimension - community in its human, universal dimension.
The structure of Chapter I is misleading. The used **terminology is not coherent with EU law.**
The fundamental rights are seen as the bedrock for the formulation of ethical principles.
From our point of view the **argumentation should start the other way around:**
A society's ethics is **based on general values** (depending of cultural, societal, anthropological and religious convictions). From these convictions **more concrete norms for acting are build up (principles)** - they are more concrete and are influencing a society, **only some of them are implemented into fundamental rights.**
The different character of fundamental rights and principles always has to be clear:
A fundamental right gives a concrete subject the possibility to defend its position against the state or against other actors with legal means.
A principle is a given regularity which gives an ethically order for acting.
Legal obligations are not in competition with not legally binding ethical obligations: but of course, there is a clear interaction between law and ethics.

In **Chapter I, part 3 the list of fundamental rights** is not correct and coherent to the use of technical terms in the EU treaties.
Part 3.3 is mentioning "Respect for democracy, justice and the rule of law" as a fundamental right. This is not corresponding with the legal terminology of the EU treaties. In Art. 2 TEU "Democracy, justice and rule of law" are mentioned as values.
In Chapter I, part 3.4 "solidarity" is mentioned as a fundamental right. This is also not coherent with the terminology of the EU law and with other International binding instruments.
The use of a coherent terminology is essential to allow these guidelines to become a useful and concrete ethical tool for business, public sector or citizens.

We are skeptical concerning the concept of the principle of "**explicability" (Chapter I, part 4)**
In this term the two components of "**transparency" and "accountability" are mixed up.** We are concerned that the very important aspect of "accountability", specially under the important aspect that it has to be user – centered, can be neglected. Our concerns are underlined by the fact, that Chapter II is mentioning "accountability" as a requirement, but not differentiates clearly enough between public and private sector. "**Accountability" has to be connected with an "understandability" in order**

to be human centered. Only if the use of AI is understandable for the user, the user knows which level of transparency and accountability he can demand.

Chapter I

Part 5.4 Lethal Autonomous Weapons (LAWS)

With regarding to a possible **development and use of artificial intelligence technology in the security & defence** domain, it is to be noted that respective EU funding instruments (notably the European Defence Fund) should **fully comply with international legal obligations** of both the EU and its Member States. Technologies and weapons that are not compatible with the **legal standards of international human rights law, international humanitarian law** as well as of **arms control, disarmament and non-proliferation** provisions must not be supported under EU funding.

Reaching an international ban on **fully autonomous weapons** currently seems to be out of scope despite various calls (https://bit.ly/2QCdjVE) and warnings by the scientific community (https://bit.ly/2syu4Hd). Nevertheless, in line with a recent resolution of the European Parliament of 12 September 2018 (https://bitly.com/), the development of **ethically problematic technologies**, including lethal autonomous weapon systems, should be **excluded** from EU defence funding.

Increase of technological sophistication of weapons tends to **disproportionately affect the civilian population** (https://bit.ly/2FlFeHA). Fully autonomous weapons enabling **lethal actions without meaningful human control** pose **major legal as well as security concerns.** Moreover**,** the **de-humanisation** and **de-responsibilisation** in performing lethal actions raises grave **ethical questions** (for a deeper reflection on ethical implications of lethal autonomous weapon systems cf. Caritas in Veritate Foundation: "*The Humanization of Robots and the Robotization of the Human Person*", https://bit.ly/2QD7UgW).

B. Chapter II

Realizing Trustworthy AI

Requirements of Trustworthy AI

Chapter II, part 1 names a list of requirements.
Most of these terms are named in other documents as principles. (See Statement of the EGE on Artificial Intelligence from March 2018).

We would recommend to add to this list the term of the "**primacy of the human being"**. The draft mentions by itself already this expression in application of the Oviedo Convention.

By this requirement the human centered approach of the draft could be underlined and it would be clarified that it is a premise for the personal and community dimension of the existence of the human being.

C. <u>General remarks</u>

We have to criticize the **procedure of publishing the draft of the guidelines and of the consultation.**

The time for contributions to this consultation is much too short. Publishing the paper just before Christmas with a first deadline for the 18.1.2019 and a postponed deadline for the 1.2.2019 for the contributions is not convincing. The purpose of the drafted guidelines is to invite all different stakeholders to share their opinions and to support the final version. The given deadline is predestined to exclude stakeholders which do not have the capacity for an immediate reaction.