

John Hall, Head of Strategy & Portfolio, Atos Fellow and Editor in Chief of the Scientific Community, Atos

Ethical principles for business in a digital world

Ethics encompasses the moral principles that govern a person's behavior and the way that activities are conducted. It helps set the dividing line between what we consider to be right and wrong.

Broadly speaking, there are three schools of ethical thought - virtue ethics ("am I a good person?"), consequentialism ("do my actions have good consequences") and duty-based ethics ("am I following an ethical code?"). These questions have helped shape ethical principles and codes of conduct for individuals and organizations in all walks of life. They help establish frameworks around areas such as the fairness and integrity of business transactions or the way that medical practitioners operate. Although most large enterprises work to a defined code of business ethics, these are not typically directly translatable to the application of digital technologies such as Artificial Intelligence. However, by drawing parallels with ethical frameworks in the healthcare profession, as they relate to the care of patients, we can gain some helpful insights as to how they might be applied.

Registered Nurses in the US are required to adhere to seven principles: Justice, Beneficence, Nonmaleficence, Accountability, Fidelity, Autonomy, and Veracity.

We can paraphrase the intent of these principles as:

Justice: Care must be fair, just and equitable

Beneficence: Care must do the good and right thing for the patient

Nonmaleficence: Care must do no harm

Accountability: Nurses must accept responsibility for their own actions

Fidelity: Nurses must keep their promises

Autonomy: Patients' rights, opinions and beliefs must be respected

Veracity: The truth must not be withheld from patients.

These all seem perfectly sensible and appropriate in terms of human-to-human interaction. But what about when we translate our thinking to the digital world? How can we make robots and AI algorithms share our anthropocentric values?

In 2019, the European Commission published its 'Ethics guidelines for trustworthy AI', proposing a set of seven key requirements that AI systems should meet in order to be deemed trustworthy. Looking at these recommendations we see some very clear parallels with the nursing principles described above. The EC guidelines can be paraphrased as follows (together with their nursing parallels shown in **blue**):

Diversity, non-discrimination and fairness: Unfair bias must be avoided, AI systems should be accessible to all. (**Justice**)

Societal and environmental well-being: AI systems should benefit all human beings (including future generations) and should be sustainable and environmentally friendly. (**Beneficence**)

Technical robustness and safety: AI systems need to be resilient and secure to ensure that unintentional harm is minimized and prevented. (**Nonmaleficence**)

Accountability: Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes. (**Accountability**)

Privacy and data governance: Full respect for privacy, protection, quality, legality and integrity of data must be guaranteed. (**Fidelity**)

Human agency and oversight: AI systems should empower human beings, and proper human oversight mechanisms must be ensured. (**Autonomy**)

Transparency: Data, systems and AI business models should be transparent, traceable and explainable. (**Veracity**)



With this in mind, perhaps the world of digital ethics is not so different in principle from the world of ethical human care. However, enforcing these principles in an environment that is constantly evolving and where digital actions are not always visible and traceable is likely to be easier said than done.

Business directions

Until relatively recently, the operating boundaries and control of digital technologies were, for the most part, well-defined and understood. However, particularly with developments in areas such as Machine Learning, Human Machine Interfaces, Data Analytics and even Digital Business Models, the scope and impact of what is possible has expanded massively. We are now faced not only with the question of “could we?” do something, but “should we?” do it, striking at the heart of the ethical theory questions of “am I doing good?” and “am I doing

right?” Even in human terms, these can be quite subjective questions. But translate them into a digital world in which the boundaries of man and machine, and physical and virtual, are becoming increasingly blurred, and the scope for subjectivity is enlarged. Making the need for guidelines such as those from the EC even more pressing.

Businesses of all sizes and across all industries and sectors need to understand the ethical implications of adopting digital technologies. These go far beyond merely adhering to the EU’s regulation on data protection (GDPR), to a clear understanding of what it means to conduct business in a way that is acceptable to customers, peers and society at large. Organizations that are able to demonstrate their adherence to an appropriate digital ethics framework will be those that build market trust and as a consequence enjoy sustainable operations.

Digital ethics is not so different in principle from the world of ethical human care. However, enforcing these principles in an environment that is constantly evolving and where digital actions are not always visible and traceable is likely to be easier said than done.