# The Challenge of Human Dignity in the Era of Autonomous Systems

**Paola Inverardi**

**Abstract** Autonomous systems make decisions independently or on behalf of the user. This will happen more and more in the future, with the widespread use of AI technologies in the fabric of the society that impacts on the social, economic, and political sphere. Automating services and processes inevitably impacts on the users' prerogatives and puts at danger their autonomy and privacy. From a societal point of view, it is crucial to understand which is the space of autonomy that a system can exercise without compromising laws and human rights. Following the European Group on Ethics in Science and New Technologies 2018 recommendation, the chapter addresses the problem of preserving the value of human dignity in the context of the digital society, understood as the recognition that a person is worthy of respect in her interaction with autonomous technologies. A person must be able to exercise control on information about herself and on the decisions that autonomous systems make on her behalf.

Nowadays, citizens continuously interact with software systems, e.g., by using a mobile device, in their smart homes, or from on board of a (autonomous) car. This will happen more and more in the future, with the widespread use of artificial intelligence (AI) technologies in the fabric of society that impacts on the social, economic, and political spheres. Effectively described by Floridi's metaphor of the mangrove society (Floridi 2018), the digital world will be increasingly dominated by autonomous systems (AS) that make decisions independently or on behalf of the users. Automating services and processes inevitably impacts on the users' prerogatives and puts at danger their autonomy and privacy.

Besides the known risks represented by, e.g., unauthorized disclosure and mining of personal data or access to restricted resources, which are receiving a huge amount of attention, there is a less evident but more serious risk which attains the core of the fundamental rights of the citizens. Worries about the growing of the data economy

P. Inverardi (✉)

Dipartimento di Scienze ed Ingegneria dell'Informazione e Matematica, Università dell'Aquila, L'Aquila, Italy

e-mail: paola.inverardi@univaq.it

and the increasing presence of AI-fuelled autonomous systems have shown that privacy concerns are insufficient: ethics and the human dignity are at stake. "Accept/ not accept" options do not satisfy our freedom of choice, and what about our individual preferences and moral views?

Autonomous machines tend to occupy the free space in a democratic society in which a human being can exercise her freedom of choice. That is the space of decisions that are left to any individuals when such decisions do not break fundamental rights and laws but are the expression of personal ethics. From the case of privacy preferences in the app domain to the more complex case of autonomous driving cars, the potential user is left unprotected and inadequate in her interaction with the digital world.

A simple system that manages a queue of users to access a service by following a by design fair ordering, e.g., first in first out, may prevent users to exchange their positions in the queue by personal choice, thus depriving users to exercise a free choice driven by her moral disposition, e.g., leave her position to an older lady.

What is considered fair by the system's developer may not match users' ethics.

The above example may seem artificial and of little importance, but it is not. In the years of digital transformation we have witnessed the side effect of increasing the rigidity of the processes implemented by digital systems beyond what the law indicated. How many times have we heard answers like "yes this could be possible, but the system does not allow it"? The above queue managing system may have associated to the ordering position a personal identifier and already made available all the personal information to the service provider. Although it may appear more complex to exchange positions, it would be not a problem for a digital system to manage the exchange. It only requires that the system is properly designed to take into consideration the user's right of choice. Overlooking this attitude in the era of autonomous technology may put at high danger our personal ethical values.

More complex interactions between systems and users shall be made possible in order to allow users' ethics to freely manifest.

However, even when such interaction is made possible, think, e.g., of the (by GDPR law mandatory in Europe) possibility to express users' consent to cookies profiling, the ways systems are presenting such interaction to the user is extremely complex and time expensive even for an expert user and often turns out in a accept/ not accept choice.

In a digital society where the relationship between citizens and machines is uneven, moral values like individuality and responsibility are at risk.

From a societal point of view, it is therefore crucial to understand which is the space of autonomy that a system can exercise without compromising laws and human rights.

Indeed, autonomous systems interact within a society, characterized by collective ethical values, with multiple and diverse users, each of them characterized by her individual moral preferences.

The European Group on Ethics in Science and New Technologies (EGE) recommends an overall rethinking of the values around which the digital society is to be structured (EGE 2018), the most important being the value of human dignity in the

context of the digital society, understood as the recognition that a person is worthy of respect in her interaction with autonomous technologies. A person must be able to exercise control on information about herself and on the decisions that autonomous systems make on her behalf.

There is a general consensus about this, but legislation follows and does not prevent the problem, and it is debatable whether regulatory approaches like GDPR or others are effectively protecting the human dignity of users. Besides regulation, active approaches have been proposed in the research on AI, where systems/software developers and companies should apply ethical codes and follow guidelines for the development of trustworthy systems in order to achieve transparency and account-ability of decisions (AI HLEG 2019; EU 2020). However, despite the ideal of a human-centric AI and the recommendations to empower the users, the power and the burden to preserve the users' rights still remain in the hands of the (autonomous) systems producers.

The above-described active approaches do not guarantee our freedom of choice that is manifested by our individual preferences and moral views. Design principles for meaningful human control over AI-enabled AS are needed. Users need (digital) empowerment in order to move from passive to active actors in governing their interactions with autonomous systems, and it is necessary to define the border in the space of decisions between what the system can decide on its own and what may be controlled and possibly overridden by the user. This also means that the system shall be designed to be open to more complex interactions with its users as far as users' moral decisions are concerned.

But how to draw the border between systems' decisions and users' ones?

Reflections on digital ethics can help in this respect. Digital ethics, as introduced in Floridi (2018), is the branch of ethics that aims at formulating and supporting morally good solutions through the study of moral problems relating to personal data, (AI) algorithms, and corresponding practices and infrastructures. It identifies two separate components, hard and soft ethics. Hard ethics is the base to define and enforce values by legislation and institutional bodies, i.e., hard ethics is what makes or shapes the law and represents the values collectively accepted, e.g., GDPR in Europe.

It is insufficient, since it cannot and shall not cover all the space of ethical decisions. Soft ethics complements it by considering what ought and ought not to be done over and above the existing regulation, not against it, or despite its scope, or to change it, or to by-pass it (e.g., in terms of self-regulation).

Personal preferences fall in the scope defined by soft ethics, e.g., the varieties of privacy profiles that characterize different users. A system will implement decisions to choices that correspond to both hard and soft ethics. The producer will guarantee compliance with the hard ethics rules but who does take care, and how it can care of, the values and preferences of each person?

We claim that soft ethics can express users' moral preferences and should mold their interaction with the digital world. Empowering a person with a software technology that supports her soft ethics is the means to make her an independent and active user in/of the digital society.

Depending on the system's stakeholders, the term user includes individuals, groups, and the society as a whole.

Thus, the capability of AS to make decisions does not only need to comply with legislation but also with any users' moral preferences (including privacy) when they manifest. This leads to the challenge of dealing with moral agreements between the system's hard and soft ethics (e.g., implemented by the system producer) and the user's soft ethics in making her decisions. It is worth noticing that if the user's soft ethics does not manifest, the system will make decisions according to hard ethics and its default soft ethics, e.g., the fair ordering algorithm of our queue example.

Let us now discuss a more elaborated example that is set in the automotive domain.

1. Setting: A parking lot in a big mall.
2. Resource contention: Two autonomous connected vehicles (named A and B hereafter), with one passenger each, are competing for the same parking lot. Passenger of vehicle B has a bad health status.
3. Context: A and B are rented vehicles, therefore they are multi-user and have a default ethics that determines their decisions. The default ethics of A and B (provided by the cars' producers) are utilitarian. Thus, the cars will look for the free parking lot that is closer to the point of interest; in case of contention, the closest car gets in.
4. Action: A and B are approaching the parking lot. A is closer, therefore it would take the parking lot. However, by communicating with B, it receives the information that the passenger in B is in bad health condition. Indeed, the passenger in B, who has a tradeoff privacy disposition, has disclosed such piece of personal information. The soft ethics of the passenger in A has a generosity disposition that manifests in the presence of bad health people, and, consequently, actions are taken to leave the parking lot to B. This use case shows how personal privacy is strictly connected to ethics: by disclosing a personal piece of information like this, the bad health passenger's tradeoff privacy disposition manifests the utilitarian expectation that surrounding drivers might have generosity disposition.

This example shows something that already happens in our ordinary reality when, e.g., owners of cars display signs, e.g., baby on board, beginners at drive, or disabled on board, about the persons in the car.

If one could imagine a sort of soft ethical initial configuration step of the vehicle for a single owner, what would happen if the car is multi-owner or the business model in the automotive domain changes from proprietary to rented due to the increased autonomy? How will the passenger disclose her piece of information and inform the surrounding vehicles? And how would a passenger be able to set the soft ethical part of the autonomous vehicle decisions according to her own ethical preferences?

From a system design perspective, there is the need of a software architectural view of the digital world which decouples autonomous systems from users as independent au pair actors. Users need to be digitally empowered in order to be

able to make, possibly complex, interactions with the surrounding AS through reliable (with respect to users' ethical preferences) protocols.

In the above direction, the separation between hard and soft ethics (Floridi 2018), initial results on design principles to empower the user (Autili et al. 2019) and on achieving moral agreements among autonomous stakeholders (Liao et al. 2019) can be exploited to concur at realizing the principle of human dignity as stated by the EGE.

# References

AI HLEG (2019) The High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Autili M., et al. (2019) A Software Exoskeleton to Protect and Support Citizen's Ethics and Privacy in the Digital World. IEEE Access 7, 2019.

EGE (2018) European Group on Ethics in Science and New Technologies. Statement on artificial intelligence, robotics and autonomous systems. https://ec.europa.eu/research/ ege/pdf/ ege_ai_statement_2018.pdf.

EU (2020) European Commission. White paper on artificial intelligence, 2020. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

Floridi L. (2018) "Soft ethics and the governance of the digital." *Philosophy & Technology* 31(1), pp. 1-8.

Liao B., Slavkovik M., and van der Torre L. (2019) Building Jiminy Cricket: An Architecture for Moral Agreements Among Stakeholders. In Proc. of AIES19, 2019.