Information Polity 27 (2022) 163–174
DOI 10.3233/IP-211529
IOS Press

Emotional AI: Legal and ethical challenges[1]

Thomas Gremsl[a,*] and Elisabeth Hödl[b]
[a]*Institute of Ethics and Social Teaching, University of Graz, Graz, Austria*
[b]*Institute of the Foundations of Law, University of Graz, Graz, Austria*

Abstract. The European Commission has presented a draft for an Artificial Intelligence Act (AIA). This article deals with legal and ethical questions of the datafication of human emotions. In particular, it raises the question of how emotions are to be legally classified. In particular, the concept of "emotion recognition systems" in the sense of the draft Artificial Intelligence Act (AIA) published by the European Commission is addressed. As it turns out, the fundamental right to freedom of thought as well as the question of the common good and human dignity become relevant in this context, especially when such systems are combined with others, such as scoring models.

Keywords: Emotional data, social scoring, social credit systems, artificial intelligence, data protection, freedom of thought, digital ethics, social ethics

Key points for practitioners:

- Opens up new perspectives.
- Contributes to awareness raising.
- Combines ethical and legal perspectives.
- Aims to put the individual and common good of people more at the centre of the digitalisation discourse.

1. Introduction

1.1. Human emotions

Emotions such as fear, anger, sadness, joy, pleasure form the core of human experience, shape our lives in profound ways and structure the basis of what captures and deserves our attention (McStay, 2018a). Data on human emotions (hereafter 'emotional data') is now collected using a variety of devices and in a wide range of life contexts. Animojis, for example, are an evolved version of the familiar emoji symbols in the universe of the technology company Apple. More precisely, they are 3-D emojis that users can create in the messaging app. There is a choice of animals or objects such as kittens, dragons, unicorns, big-eared bears or pale-faced aliens that can reflect one's facial expression. If the user laughs, the bear laughs; if the user sticks out her tongue, the unicorn does too. Apple uses its Face ID facial recognition system for this entertainment application. The technology behind it is the TrueDepth camera system, which is made up of a variety of technical components (floodlight, infrared camera, front camera, dot projector, proximity sensor, ambient light sensor, speaker, microphone). By combining these technologies, more than 50 different muscle movements of the face can be recorded and analysed (Tillman, 2021).

[1] This article received a correction notice (Erratum) post publication with DOI 10.3233/IP-229012, available at http://doi.org/10.3233/IP-229012.

[*] Corresponding author: Thomas Gremsl, Institute of Ethics and Social Teaching, University of Graz, Heinrichstraße 78b/2, 8010 Graz, Austria. E-mail: thomas.gremsl@uni-graz.at.

1.2. *Emotional AI*

In the emotionalisation of modern media life, global technology companies are working on broader techniques towards more widely applicable emotional AI. Emotional AI is a term that cannot simply be defined universally, especially in an interdisciplinary context. We subsequently understand emotional AI as AI systems that access and process emotional data. As far as we know, all the big players such as Amazon, Apple, Facebook (Metaverse), Google, IBM and Microsoft are working on tools for emotional AI to develop "empathic products" for different applications and settings. The areas of such applications are diverse and are likely to ultimately extend to all areas of human life: safety in autonomous vehicles by recognising emotions in drivers, facilitating communication in call centres using voice analytics, use for market analysis, for all industries such as wellness, health, fitness, dating, security (McStay, 2018b, 3). Thus, a "datafication" of human emotional life is emerging. Research is also increasingly addressing these issues (Barrett et al., 2019; Fraunhofer-Institut für Integrierte Schaltungen IIS).

With a view to these developments and the associated data technologies, the subject of this paper will be a contribution to the risk and danger analysis of the systemic collection of data on human emotions.

2. Planned standards of the European Union

With the proposal for an Artificial Intelligence Act (AIA), the European Union has presented a concept for a set of standards based on the insight that artificial intelligence systems (AI) are becoming increasingly important in public life and that the growing importance of these systems requires sensible, efficient and sustainable regulations (21.4.2021 COM(2021) 206 final 2021/0106 (COD)). With the use of AI systems, not only the opportunities but also the dangers and risks for human life and coexistence will increase, which is why the proposal is based on a risk-based approach to AI systems. A uniform definition of the term "AI systems" does not exist due to the complexity of the technologies on "artificial intelligence", but initially intelligence emanating from machines is assumed, which contrasts with the natural intelligence of living beings. The AIA proposal defines AI system as "software developed using one or more of the techniques and concepts listed in Annex I that is capable of producing results such as content, predictions, recommendations, or decisions that influence the environment with which it interacts, with respect to a set of goals specified by humans" (Art 3 Z 1).

2.1. *Reality of life*

The proposal takes into account structural, social, political and economic aspects in the use of AI. This essentially corresponds to the reality of many decision-makers, who are increasingly required to use the opportunities of AI systems for their areas of responsibility, but at the same time try to minimise increasing risks. Legislators are therefore called upon to create appropriate framework conditions that allow both to use the advantages of the technologies and at the same time try to minimise possible risks. The social significance of this tension is also reflected in the fact that non-legislative initiatives and self-regulatory measures by private actors are increasing on national as well as international levels. The Partnership on AI to Benefit People and Society, for example, brings together technology companies, academia and research, and civil society organisations calling for an appropriate approach to AI in society. A recently published Civil Society Statement of 30.11.2021 (signed by 115 organisations from different European countries) references the Commission's proposal and draws particular attention to structural power inequalities that arise from the use of AI systems. The appeal is made to institutions of the European Union to undertake the sharpening of fundamental rights objectives in the envisaged regulatory framework. What is apparent, in any case, is the interest of civic participation in discourse (An EU Artificial Intelligence Act for Fundamental Rights, 2021).

2.2. Emotion recognition systems

In this way, society generates more information about human emotions and thus derives intentions and attitudes in a systemic way. This tendency is taken into account in the AIA proposal in Art 3, Z 34, which defines "emotion recognition systems" as AI systems that serve the purpose of determining or inferring emotions or intentions of natural persons on the basis of their biometric data. As can be seen from the definition, motion detection, for example, is intended to make emotions interpretable by analysing human movements. Think of a person wandering around a deserted underground car park in the early hours of the morning, peering through car windows into the interior of vehicles. From the security industry's point of view, a behavioural prediction could arise here, such as the likelihood of an imminent offence.

How this systemic process is to be understood will be explained using Facial Emotion Recognition (FER). This is the analysis of sentiments by means of technologies known under the umbrella term "affective computing". We are talking about a multidisciplinary field of research that deals with the possibilities of interpreting human emotions and states of mind by means of computer technology. AI-applications are often used in this context. Facial expressions, as forms of non-verbal communication and their interpretation by means of technologies, are the subject of research in psychology, specifically in the field of human computer interaction. Roughly speaking, an FER analysis is carried out in three steps: (1) Face Detection, (2) Facial Expression Detection, (3) Expression Classification to Emotional State. Depending on the respective algorithm, these facial expressions can be classified into categories. These are, for example, anger, disgust, fear, joy, sadness, surprise. It can also be classified into compound emotions such as happily sad, happily surprised, happily disgusted, sadly fearful, sadly angry, sadly surprised. Or it can be assigned to physiological or mental states, such as tiredness or boredom. In addition, combinations with biometric identifications are possible, i.e. with similar analyses of voice, text or health data (EDPS, 2021). What we have to recognise and classify from a legal and ethical point of view is the systemic process in which AI technology is used to turn biometric data into data about people's emotions, which we will call "emotional data" here.

2.3. Dangers and risks

If we reflect on the legal and ethical challenges of information and communication technologies and do so with a focus on "emotional data", the question of what a legal definition of the term "emotions" and "emotional data" can be seems central. Especially in the context of a set of norms – which lays down rules for rights, duties and sanctions – it must also be clarified whether emotions and values are related. People can have irrational emotions or emotions can persist even if the associated value has long disappeared (Berninger & Döring, 2009; Mulligan, 2009; Tappolet, 2009). This presuppositional question of how emotions and value judgements are connected is relevant for the definition of "emotional data" and its legal significance. Furthermore, ethical considerations must be made as to what follows for a society from a technology-based linking of emotion recognition systems and value judgements. Both problem contexts – the legal and the ethical – are relevant for reflections on modern surveillance technologies.

3. Legal problem context

3.1. Data protection issues

The General Data Protection Regulation (GDPR) also contains provisions for dealing with AI systems, in particular by standardising obligations in data processing and in shaping the rights of the data subject when using personal data. The GDPR applies when AI systems are fed with personal data, when theyuse them or when they serve as the basis for decisions. Thus, the principles enshrined in the GDPR, such as the prohibition of discrimination or the principle of purpose limitation, also have relevance for AI systems. As with all large and comprehensive data collections, there is a risk of systematic misuse by third parties.

The GDPR also sees itself as a regulatory system for a European data space in which the possibility of data use in increasingly digitalised environments is possible and necessary. Personal data is data that identifies or makes identifiable a person. Recital 26 of the GDPR states that "in order to determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used by the controller or by any other person to identify the natural person, directly or indirectly, such as segregation". In determining whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the cost of identification and the time required for it, taking into account the technology and technological developments available at the time of the processing.

How can the term "emotional data" be defined in such a way that it is suitable for the legal context – in particular in the context of "emotion recognition systems" mentioned in the draft AIA? As we have seen, biometric data become "emotional data" through systemic processing in the emotion recognition system.

In a first step, the treatment of biometric data under data protection law will therefore be considered. In a second step, the question will be posed which insights can be found for dealing with "emotional data".

The definition of "biometric data" can be found in Article 4 (14) of the GDPR. These are "personal data obtained by means of special technical procedures relating to the physical, physiological or behavioural characteristics of a natural person which enable or confirm the unique identification of that natural person, such as facial images or dactyloscopic data". Biometrics is the science of recording and measuring living beings, the numerical measurement in the field of biology. Biometrics involves the automated identification of living persons based on physical characteristics or typical behavioural patterns. Biometric data is suitable for the unique identification of people if the measured characteristics are unique. Since biometric data have a direct link to the human body, they are difficult to change or falsify. Nevertheless, certain biometric characteristics can change in the course of a life. The definition ties in with the concept of "special technical procedures" by which personal data are obtained. Thus, only data collected by means of technical procedures are covered by the definition. Hence, simple photographs of persons do not constitute biometric material. Only the further technical processing of the image data should lead to the existence of biometric data (Hödl, 2021). According to its wording, the regulation also includes acoustic identification features of voice recognition as well as biotechnological or chemical measuring methods. Biometrics can be found in the identification of a person by means of fingerprints, automated recognition of the retina and iris, complex voice recognition procedures, signatures, faces, body movements, statistical or dynamic imaging procedures, by means of X-rays or infrared, acoustic voice profiles, typing behaviour of a person at a keyboard or life scans in real time. Biometrics thus allows for unique identification and the linking of different data sets that would otherwise not be directly related. Biometric data are raw data, such as features captured directly with a sensor, but also biometric templates, which are feature vectors extracted from the raw data and typified. By means of filtering and conversion, user-specific features are extracted from the biometric data to create a biometric template.

Let us return to the procedure of the FER analysis, which, as we have seen, is carried out in three steps: (1) Face Detection, (2) Facial Expression Detection, (3) Expression Classification to Emotional State. Emotional data is useful for uniquely identifying people if the features measured are unique. The features of the face are unique, but still fall into the category of biometric data. The features of Facial Expression Detection and Expression Classification to Emotional State can be unique if certain patterns are repeated. Lastly, during the third step of the process, biometric data become emotional data and, analogous to biometric templates, emotional templates should also be considered.

From our view it would therefore be conceivable to define emotional data in terms of data protection law as personal data on emotional characteristics or states of a natural person obtained by means of special technical procedures, which enable or confirm the unique identification of this natural person.

In any case, the data of a face detection are already personal data due to their biometric quality and thus fall into the category of sensitive data according to Article 9 (1) of the GDPR for which a special level of protection is provided. Recital 51 of the GDPR states: “Personal data which by their nature are particularly sensitive in relation to fundamental rights and freedoms deserve specific protection because the context of their processing could give rise to substantial risks to fundamental rights and freedoms”. In this sense, data protection law takes into account sensitive data, such as racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, data concerning health, sexual orientation, as mentioned in Article 9 (1). For these categories of data, special consent requirements are provided under the GDPR regime. When looking for the common characteristics of sensitive data, the “significant risk to fundamental rights and freedoms” must be taken into account, as mentioned in recital 51 of the GDPR. If one wants to classify emotional data in the category of sensitive data, it must be asked whether its processing is associated with a significant risk to fundamental rights and freedoms.

3.2. Fundamental legal issues

As articulated in the AIA proposal, due to “opacity, complexity, data dependence and autonomous behaviour [. . .] the use of AI may lead to the violation of some of the fundamental rights enshrined in the EU Charter of Fundamental Rights”. The AIA aims to “protect these fundamental rights to a high degree and addresses different sources of risk through a clearly defined risk-based approach”. It emphasises that all participants in value chains are subject to a set of trustworthy AI requirements and proportionate obligations to ensure that the rights protected by the Charter are protected. Central among these are human dignity (Article 1), respect for private life and protection of personal data (Articles 7 and 8), non-discrimination (Article 21) and equality between women and men (Article 23), the right to freedom of expression (Article 11) and freedom of assembly and association (Article 12), the right to an effective remedy and to a fair trial and the presumption of innocence and rights of defence (Articles 47 and 48), and the general principle of good administration. It is also noteworthy that particular reference is made to group-specific rights, such as the right of workers to fair and just working conditions (Article 31), consumer protection (Article 28), the rights of the child (Article 24) and the integration of persons with disabilities (Article 26). In addition, the right to a high level of protection and improvement of the quality of the environment (Article 37), especially in relation to the health and safety of persons, is at stake.

With regard to the fundamental rights assessment of the risk and dangers of emotional data, the factual scope of protection of Art 9 ECHR; Art 10 CFR which includes freedom of thought, conscience and religion, also seems relevant. For the problem area of “emotional data”, the aspect that seems important is that which encompasses “freedom of thought”. Freedom of thought protects the internal human processes that lead to the formation of opinions and convictions, including religion and conscience. Freedom of thought aims to protect the innermost core of human self-determination and thus respect for the individual personality. “This is not only limited to inner thought processes, but also encompasses an expression of thoughts formulated in – unconscious or conscious, involuntarily or arbitrarily conducted – self-talk, in which the person feels ‘alone with himself or herself’ and, depending on the circumstances, may also include other conversations and files stored in an information technology system” (Graupner, 2021). This fundamental right therefore includes freedom of thought but also the freedom to conceal and keep one’s thoughts secret (Grabenwarter, 2003; Grabenwarter, 2008).

What risk minimisation provisions are envisaged for AI systems with regard to fundamental rights? Recital 70 of the AIA addresses information obligations for "emotion recognition systems" and Article 52 standardises transparency obligations for the use and deployment of "emotion recognition systems". But are these information and transparency obligations really sufficient safeguards with regard to the protection of fundamental rights in the use of "emotional data"? Annex III of the draft AIA mentions high-risk AI systems (to be classified and assessed as high-risk according to Art 6(2)). According to Annex III No. 6 b (law enforcement) and Annex III No. 7a (migration, asylum and border control), these are AI systems that are intended to be used by law enforcement authorities as lie detectors and similar instruments or to determine the emotional state of a natural person.

4. Ethical problem context

4.1. Emotional data and scoring

For the problem area of "emotional data" raised in the context of AI systems, a problem area should be highlighted, such as that addressed in the context of data protection law. Profiling (Art 22 GDPR) enables the analysis and prediction of certain personal areas of people's lives. By means of systematic procedures – mathematical-statistical analysis of empirical values – the future behaviour of groups of people and individuals with certain characteristics is to be predicted. Scoring is based on the consideration that if certain comparable characteristics of an individual are present, similar future behaviour is likely (Keller, 2012). It can also be understood as a procedure that is intended to map certain aspects of an individual that result from interaction with his or her social environment into a numerical value, the social score (Hoffrage & Marewski, 2020). An example of scoring is credit scoring of individuals: Credit scoring is used by credit institutions as a statistical procedure to assess risk for private standardised instalment loans and small loans. Such credit scoring is already used in several countries, such as the USA or Germany. However, their use is in a tense field between the access to credit, only made possible by such a system and the danger of systematic discrimination against individual population groups (Lischka & Klingel, 2017).

A question of particular ethical relevance is: How does society deal with the AI-based classification of human feelings into a scoring system on the basis of "emotional data"? Let's take the example of a fast-food chain that wants its employees to smile and spread a positive mood when serving burgers. Those who fulfil this requirement particularly well receive points within the framework of an internal credit system and consequently benefits. The analysis of "emotional data" to predict probabilities is used in the USA, for example, to classify suitability for police service, as the concept of the "connected cop" makes clear: Is the person in question too scared and too quick to draw the gun? (McStay, 2018a). The use of "emotional data" also becomes explosive in areas dealing with mental health. Current fields for emotional data applications are provisons of personal services, customer behaviour analyses and advertising, healthcare, employment, education, public safety, crime detection, gaming and sex industries. Emotional data will therefore play a central role in marketing and in the development of new business models and services in almost all areas of life. From a data protection perspective, one of the specific risks of profiling and automated decision-making is that it can cause changes in behaviour (EDPS, 2021). When a person knows that he or she will be subjected to a profiling process, self-censorship, chilling effects in society and feelings of insecurity between citizens can be induced, especially when it comes to non-democratic governments, so that people's political behaviour is influenced.

Ethics must therefore focus on the well-being of people – all of them – who are affected by these technologies, and this in turn draws attention to two key issues: the common good debate (4.2) and human dignity (4.3).

4.2. Common good

A look at history shows that people use technology. Regardless of how risky or cruel it may have been. There is no doubt that this way of thinking has produced achievements. Prosperity-oriented societies like those of the industrialised countries thrive on the ductus of innovation. In order to minimise the dangers and risks of new technologies as far as possible, ethical implications of technological innovations must therefore be pointed out today more than ever. The challenges associated with the increasing technological achievements of the digital transformation thus represent an imperative for action for ethics. Against this backdrop, ethics should open up perspectives for a humane design of such technologies and systems, keeping the well-being of all people in mind. An important point of reference in this regard is the common good. As Catholic social teaching (CST) states, "[...] the common good calls for social peace, the stability and security provided by a certain order which cannot be achieved without particular concern for distributive justice; whenever this is violated, violence always ensues. Society as a whole, and the state in particular, are obliged to defend and promote the common good." (Pope Francis, 2015, para. 157). From a socio-ethical perspective, an essential aspect of the idea of the common good is thus to open up perspectives that promote the shaping of structural and institutional conditions that enable individuals to realise self-development and perfection in the best possible way (Veith, 2004). It is the state, and thus the legislature, that constitutes a primary space for action for the idea of the common good (Remele, 2021). Understood in this way, the state has an instrumental character with the purpose of being conducive to the self-realisation of the person, indeed of all persons (Anzenbacher, 1998). In this context, it is important to emphasise that the common good is to be realised subsidiarily by individuals in the sense of personal responsibility, personal competence and personal empowerment (Zsifkovits, 2012). According to Veith, the common good, in which the well-being of all individuals is a constitutive element, is thus a social value that calls for the constant improvement of social structures with a view to promoting the personhood of people in society (Veith, 2004). The individual and the common good are thus constructively interrelated.

In China for example, scoring systems, in the sense of a social credit system encompassing almost all areas of society, are already being piloted in several cities and municipalities. The declared goal of those in charge is to create a "harmonious society" (change Magazin – Das Magazin der Bertelsmann Stiftung, 2018). Thus, for the purpose of achieving the common good idea of a group, the latest surveillance technologies are coupled to the social credit system. This means continuously monitoring and evaluating each individual – 24 hours a day, 7 days a week.

Through the use of such systems, in particular the privacy of the individual is torpedoed, undermined and eroded. In our (Western) societies, privacy is linked to the well-being of the individual and is considered a high good to be protected. For example, Art 8 ECHR, Art 12 UDHR and Article 7 of the Charter of Fundamental Rights of the European Union (CFR) enshrine the protection of private life. And yet it is precisely this area that is of enormous, essential importance to us humans. As an undisturbed space for personal development, as a space in which we can simply be ourselves, away from social customs and guidelines, and have the power to decide how and with whom we share this sphere, this sphere is of immense importance for our lives. As the report of the German Data Ethics Commission states, this space is closely connected to the dignity of every human being. Self-determination in choosing and pursuing our life goals and lifestyles is an expression of human freedom. This self-determination also includes the right of the individual to determine who may collect and use which personal data, when and for what purpose (Datenethikkommission der Bundesregierung, 2019).

When scoring systems are coupled with emotion recognition systems, the result may not be increased objectivity or efficiency, but may also pose dangers to people concerned, such as limiting freedom of thought, becoming screenable subjects for the state. Furthermore it might become increasingly interesting for powerful corporations to use this technology. From an ethical perspective, we are experiencing a new quality here: social profiling is experiencing a yet unseen dangerous horizon. Emotional AI scans and measures bodies, analyses vital signs, conscious or unconscious looks or expressions. The system analyses the entire body with all its functions and subsequently tries to turn our inner life inside out – it tries to look into our soul life and into our world of thoughts. It not only attempts to make our exterior but also our interior objectifiable and usable for certain purposes. From an ethical point of view, a massive warning must be issued here. Furthermore, these outlined potential dangers raise questions about the regulation of such AI-systems.

At this point, we would refer to the regulatory concept of Katharina Zweig and her colleague Tobias D. Krafft for AI and digital risk technology. In this framework, the extent to which a system should be regulated depends on the overall damage potential of the system and the degree of individual dependence on the system in question (Klingel et al., 2020). The European Union has also been working intensively on the topic of AI and risk technologies and their regulation for some time. The ethical approach chosen here, which aims at trustworthy AI, is based on the fundamental rights enshrined in the EU Treaties, the EU Charter and human rights. The ethical principles are derived from the understanding of fundamental rights, which include respect for human autonomy, prevention of harm, fairness and the explainability of systems. In order to be able to realise these prerequisites and create trustworthy AI, the primacy of human action over computer-based decisions is emphasised. The common ground is seen in respect for human dignity, reflecting the "human-centred approach" in which people are accorded a unique and inalienable moral status with primacy in the civil, political, economic and social spheres (High-Level Expert Group on Artificial Intelligence, 2019).

The AIA's concept for regulations and measures in dealing with artificial intelligence focuses on the graduated risk-based approach. Social scoring systems are classified as an "unacceptable risk". According to the Commission, these systems must be banned because they allow social behaviour to be assessed, thereby posing an unacceptable risk to the livelihood and rights of people in society (COM(2021) 206 final 2021/0106(COD)).

This truly personal space, which is an important dimension for the processes of our personality development, must therefore continue to be safeguarded. People must be able to live in public space without fear of constant surveillance and evaluation of their social life. This includes that their feelings, emotions or moods in public space are not analysed and interpreted by AI systems based on economic, political or other interests. Technical means must not be used unilaterally to exert pressure on members of society. The focus should be on supporting members of society who are affected by negative impacts of technologies. The needs of the socially weaker and vulnerable must be given special consideration within the framework of the common good-oriented use, the design of digital systems and in the context of digitally supported social processes.

4.3. *Human dignity*

"Human dignity is inviolable. It must be respected and protected" (Art 1 CFR).

The central message of our current fundamental rights catalogues is almost preamble-like the inviolability of human dignity (Art 1 CFR, Art 1 UDHR)). In this context, however, it must also be asked what is to be understood by the philosophically meaningful concept of human dignity. The German philosopher Immanuel Kant is often referred to with the concept of human dignity in this regard. For him, dignity is removed from any calculability; it has no price and knows no equivalents. Kant emphasises the self-purposefulness of the human being and in this respect decisively excludes a devaluation of the human being to the level of an object and an associated evaluative comparability (Kant, 1785). Every human being is endowed with this dignity, which is an absolute incomparable value, and one cannot lose it at any time (Schaber, 2012).

In its report from 2019, the Data Ethics Commission of the German Federal Government also refers to the fundamental character of human dignity and implicitly refers in its argumentation to philosophical aspects articulated by Immanuel Kant. In doing so, it also decidedly refers to the fact that every human being is an individual and not a pattern of data points, who may not be classified in a classifying system across all their areas of life, as would be the case with a super-scoring system, for example. The protection of human dignity must therefore be taken into account, especially in the use of AI systems (Datenethikkommission der Bundesregierung, 2019).

Human dignity gives rise to the claim that every human being must be treated as a subject and not as an object in all state proceedings. This claim to social value and respect gives rise, for example, to the prohibition of humiliating or degrading punishments in criminal proceedings (see BVerfGe 72, 105; 115f with further references), or the prohibition of the use of lie detectors or a truth serum (certain drugs) against an individual's will. Under the title of human dignity, the fundamental legal conditions of individual and social existence of human beings must be preserved.

This means that no person may be degraded to the status of a mere object of state power, insofar as their subject quality is thereby called into question. The philosopher Martha Nussbaum has done systemic work on this. According to this, a person is objectified when one of the following qualities is applied: (1) Instrumentality, (2) Denial of Autonomy, (3) Inertness, (4) Fungibility, (5) Volabilty, (6) Ownership and (7) Denail of subjectivity – here meaning that the person is treated as if there were no reason to consider their inner experiences or feelings (Nussbaum, 1995, p. 257).

In this context, reference should be made once again to Annex III of the draft AIA, in which high-risk AI systems are mentioned in accordance with Article 6(2). In item 6 b (law enforcement) and in item 7a (migration, asylum and border control), AI systems are mentioned that are intended to be used by law enforcement authorities as lie detectors and similar instruments or to determine the emotional state of a natural person. Objectification works both ways within a relationship between humans. As one person is objectified, there is at the same time a self-objectification of the other, who in turn loses the self-awareness of being a subject (Elias & Scotson, 2002).

But what when one of these two human subjects is exchanged with an object, like a machine, and the machines' access to the individual data of the persons concerned turns them into digital objects for the purposes of companies or institutions? Especially when it occasionally involves the technical-digital objectification attempts of one of the most intimate and personal aspects of human existence – our emotions. While the first example involves an inevitable subject-object interaction in the respective concrete context, the second example presents a new, one might even say a special challenge. Does this mean that human dignity of people impacted by such AI systems is affected? With regard to Kant, this question must be affirmed, because the human being must never be an object, never a means for other purposes; he or she is always an end in himself or herself (Sandel, 2019). This raises the question of when a person becomes an object. In view of the criteria established by Nussbaum, of which only one would have to apply in order to identify such an objectification of a person, we see several criteria fulfilled here. Going into all the criteria individually would go beyond the scope of this paper. Therefore, from a research pragmatic point of view, we would like to briefly illuminate two criteria as examples. In view of the previous discussion in this paper, we have chosen the denial of autonomy and the denial of subjectivity.

The criteria that ultimately apply in a specific case always depend on the AI system concerned and the respective context. Hence, no generally valid statement can be made here. From our point of view, an emotional AI – especially if it is linked to a scoring system – would at least partially restrict the personal autonomy of the person concerned and thus their self-determination. For example when the behaviour displayed creates a score that has an impact on the real reality of life, as happens in some cities and municipalities in the context of the social credit system in China (Kühnreich, 2018; Schlieker, 2019; Sommer, 2020).

However, it is precisely the criterion of "denail of subjectivity" that we see as particularly tangential in the context of emotional data and emotional AI. The technical-digital objectification of emotions in data patterns subsequently makes it possible to classify them. This means that the feelings, inner experiences and emotions of individuals are not – or only to a limited extent – taken into account. This classification into different classes and groups and the associated objectification of the individual for various economic, political or other purposes – regardless of whether it is a matter of exercising control or protecting those affected – disregards human dignity. And with this, absolute boundaries are exceeded which from an ethical perspective must be particularly respected and protected.

5. Outlook and conclusion

Technological development shows us more and more how data is interconnected. Profiling and prediction for various purposes are – and this should further be refined – treated with particular caution from a data protection law perspective. It is becoming apparent that with emotional data, new categories of data are emerging that are not yet sufficiently covered by data protection law in this form. Emotional data and their processing affect the privacy of the individual in a particularly drastic way and their processing therefore poses a particularly high risk. It will therefore be necessary to discuss differentiated regulatory concepts for dealing with emotional data. Our society faces enormous ethical and legal challenges. As has been shown, emotions are increasingly being recorded and evaluated automatically. It is therefore necessary to create awareness of these categories of data in data protection law and, from an ethical point of view, to consider the interests of the individual but also the issue of the common good. All in all, as life becomes increasingly technological, the importance of human-centred design will become apparent.

We have been able to illustrate that emotional data and emotional AI will pose major challenges, especially regarding their regulation. Although regulatory impulses can be derived indirectly from already existing regulations, more explicit regulation of the handling of such data and technologies is needed. The very high probability of coupling such data and AI with social credit systems poses a further threat to individual and common welfare, since several ethical boundaries are being crossed here. Catholic social teaching and social ethics, for example, provide important ethical perspectives for dealing with these challenges. Additionally the reference to the dignity debate using philosophical perspectives from Kant and Nussbaum provides important impulses for dealing with these manifold challenges.

Especially complex topics such as emotional data and emotional AI touch on several fundamental rights (e.g. freedom of expression, privacy). Such technologies can be used under the sign of the good, but also under negative signs. For example, they can be used to foster security, but they can also lead to a new level of unwarranted and unnecessary control and surveillance. It is precisely this referencing of emotional data and emotional AI in the context of modern surveillance technologies that poses complex ethical and legal challenges to our society today. As a result, it is necessary to shape these challenges in a humane way. In the end, it must always be about people and their well-being. Fundamental rights of people must not be restricted due to new technological capabilities in the context of particular economic, political or other interests.

Acknowledgments

The Open Access publication of this paper was supported by the Panelfit project. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039. This document reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.

References

An EU Artificial Intelligence Act for Fundamental Rights: A Civil Society Statement. (2021). https://edri.org/wp-content/uploads/2021/12/Political-statement-on-AI-Act.pdf.

Anzenbacher, A. (1998). Christliche Sozialethik: Einführung und Prinzipien. UTB für Wissenschaft: [Große Reihe]: Theologie.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollatk, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the*, *20*(1), 1-68. doi: 10.1177/1529100619832930.

Berninger, A., & Döring, S. A. (2009). Einleitung: Emotionen und Werte. In S. A. Döring (Ed.), Suhrkamp-Taschenbuch Wissenschaft: Vol. 1907. Philosophie der Gefühle (1st ed., pp. 433-438). Suhrkamp.

Beyond Verbal. (2013). Access Moodies to Analyze Your Emotions and Attitude in Real Time [Youtube]. Access Moodies to Analyze Your Emotions and Attitude in Real Time [accessed 5 October 2021].

Beyond Verbal (2021). https://en.wikipedia.org/wiki/Beyond_Verbal [accessed 5 October 2021].

Brooker, C. (2016). Nosedive [Netflix] (= Black Mirror S3 E1).

change Magazin – Das Magazin der Bertelsmann Stiftung. (2018). Was steckt wirklich hinter Chinas Social Credit System? https://www.change-magazin.de/de/china-social-credit-system-was-steckt-wirklich-dahinter/. [accessed 5 October 2021].

Datenethikkommission der Bundesregierung. (2019). Gutachten der Datenethikkommission. Berlin.

Daum, Jeremy, Untrustworthy: Social Credit Isn't What You Think It Is, VerfBlog, 2019/6/27, https://verfassungsblog.de/untrustworthy-social-credit-isnt-what-you-think-it-is/, doi: 10.17176/20190627-112616-0. [accessed 5 October 2021].

Dorloff, A. (2018). China auf dem Weg in die IT-Diktatur. https://www.deutschlandfunk.de/sozialkredit-system-china-auf-dem-weg-in-die-it-diktatur.724.de.html?dram:article_id=421115 [accessed 5 October 2021].

EDPS – European Data Protection Supervisor (2021). TechDispatch – Facial Emotion Recognition. #1/2021, https://edps.europa.eu/data-protection/our-work/publications/techdispatch/techdispatch-12021-facial-emotion-recognition_en. [accessed 5 October 2021].

EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1.

European Commission. (2021). Europe fit for the Digital Age:Commission proposes new rules and actions for ex-cellence and trust in Artificial Intelligence. https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682. [accessed 5 October 2021].

European Union, Charter of Fundamental Rights of the European Union, 26 October 2012, 2012/C 326/02, available at: https://www.refworld.org/docid/3ae6b3b70.html [accessed 5 October 2021].

Elias, N., & Scotson, J. L. (2002). Etablierte und Außenseiter (1st ed.). Suhrkamp.

Fraunhofer-Institut für Integrierte Schaltungen IIS. Gesichtserkennung: Schnell, robust und in Echtzeit. https://www.iis.fraunhofer.de/de/ff/sse/imaging-and-analysis/ils/tech/shore-facedetection.html. [accessed 5 October 2021].

Grabenwarter, Art 9 ECHR, in Korinek/Holoubek et al (eds), Bundesverfassungsrecht (6th ed, 2003) Rz 8 and 10; Grabenwarter, Art 9 ECHR, in Pabel/Schmal (eds), IntKommEMRK (8th ed, 2008) Rz 32 and 35 f.

Graupner, Eheaufhebungsgrund Bisexualität? EF-Z 2021/50, 3/2021, 113

Grunwald, A. (2019a). Digitalisierung als Prozess: Ethische Herausforderungen inmitten allmählicher Verschiebungen zwischen Mensch, Technik und Gesellschaft. *Zeitschrift Für Wirtschafts-und Unternehmensethik*, *20*(2), 121-145.

Grunwald, A. (2019b). Der unterlegene Mensch: Die Zukunft der Menschheit im Angesicht von Algorithmen, künstlicher Intelligenz und Robotern. riva Premium.

High-Level Expert Group on Artificial Intelligence. (2019). Ethics Guidelines for Trustworthy AI.

Hödl, E. (2021). Art 4, RZ 146-153. In R. Knyrim (Ed.), Großkommentar. Der DatKomm: Praxiskommentar zum Datenschutzrecht in 2 Mappen inkl. 53. Lfg (p. 148). MANZ Verlag Wien.

Hoffrage, U., & Marewski, J. N. (2020). Social Scoring als Mensch-System-Interaktion. In O. Everling (Ed.), Social Credit Rating: Reputation und Vertrauen beurteilen (1st ed., pp. 305-329). Springer Fachmedien Wiesbaden.

Kant, I.(1785). Grundlegung der Metaphysik der Sitten. https://www.projekt-gutenberg.org/kant/sitte/sitte.html accessed [5 October 2021].

Keller, H. (2012). Scoring. https://wirtschaftslexikon.gabler.de/definition/scoring-53269. [accessed 5 October 2021].

Klingel, A., Krafft, T. D., & Zweig, K. A. (2020). Mögliche Best-Practice-Ansätze beim Einsatz eines algorithmischen Entscheidungsunterstützungssystems am Beispiel des AMS-Algorithmus. In M. Hengstschläger (Ed.), Digitaler Wandel und Ethik (1st ed., pp. 190-215).

Kühnreich, K. (2018). Soziale Kontrolle 4.0? Chinas Social Credit System. *Blätter Für Deutsche Und Internationale Politik(7)*, 63-70.

Lischka, K., & Klingel, A. (2017). Wenn Maschinen Menschen bewerten: Internationale Fallbeispiele für Prozesse algoritmischer Entscheidungsfindung. Bertelsmann Stiftung.

McStay, A. (2018a), Emotional AI, The Rise of Empathic Media.

McStay, A. (2018b), The Right to Privacy in the Age of Emotional AI, https://www.ohchr.org/Documents/Issues/DigitalAge/ReportPrivacyinDigital-Age/AndrewMcStayProfessor%20of%20Digital%20Life,%20BangorUniversityWalesUK.pdf [accessed 5 October 2021].

Mulligan, K. (2009). Von angemessenen Emotionen zu Werten. In S. A. Döring (Ed.), Suhrkamp-Taschenbuch Wissenschaft: Vol. 1907. Philosophie der Gefühle (1st ed., pp. 462-495). Suhrkamp.

Nell-Breuning, O. V. (1954). Gemeingut, Gemeinwohl. In O. V. Nell-Breuning & H. Sacher (Eds.), Wörterbuch der Politik (2nd ed., pp. 51-58). Herder.

Nida-Rümelin, J., & Weidenfeld, N. (2018). Digitaler Humanismus: Eine Ethik für das Zeitalter der künstlichen Intelligenz (4th ed.). Piper.

Nussbaum, M. C. (1995). Objectification. *Philosophy and Public Affairs*, *24*(4), 249-291.

Pope Francis. (2015). Encyclical letter Laudato si'. On care for our common home. https://www.vatican.va/content/francesco/en/encyclicals/documents/papa-francesco_20150524_enciclica-laudato-si.html. [accessed 5 October 2021].

Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts. COM(2021) 206 final 2021/0106(COD) 2021.

Re:publica 2019. Beyond Black Mirror – China's Social Credit System. https://www.youtube.com/watch?v=jm8TKf0eTEs. [accessed 5 October 2021]

Remele, K. (2021). "Es geht uns allen besser, wenn es allen besser geht": Die ethische Wiederentdeckung des Gemeinwohls. Matthias Grünewald Verlag.

Sandel, M. J. (2019). Gerechtigkeit: Wie wir das Richtige tun (6th ed.). Ullstein: Vol. 37537. Ullstein.

Schaber, P. (2012). Menschenwürde und Selbstverfügung. Jahrbuch Für Recht Und Ethik, 319-329.

Schlieker, A. (2019). Digitale Überwachung in China: Diktatur 2.0 oder nur effizienteres Regieren? In S. Rietmann (Ed.), Soziale Arbeit als Wohlfahrtsproduktion: Vol. 15. Beratung und Digitalisierung: Zwischen Euphorie und Skepsis (1st ed., pp. 109-128). Springer Fachmedien Wiesbaden GmbH; Springer VS.

Shabani/Borry (2018). Rules for processing genetic data for research purposes in view of the new EU General Data Protection Regulation, volume 26, pp. 149-156.

Spiekermann, S. (2019). Digitale Ethik: Ein Wertesystem für das 21. Jahrhundert.

Sommer, T. (2020). Überwachungsstatt China. In O. Everling (Ed.), Social Credit Rating: Reputation und Vertrauen beurteilen (1st ed., pp. 203-207). Springer Fachmedien Wiesbaden.

Stalder, F. (2019). Kultur der Digitalität (4th ed.). Edition Suhrkamp.

Tappolet, C. (2009). Emotionen und die Wahrnehmung von Werten. In S. A. Döring (Ed.), Suhrkamp-Taschenbuch Wissenschaft: Vol. 1907. Philosophie der Gefühle (1st ed., pp. 439-461). Suhrkamp.

Tillman, M. (2021). Was sind Animoji? So erstellen und verwenden Sie Apples animiertes Emoji. https://www.pocket-lint.com/de-de/software/news/apple/142230-was-sind-animoji-wie-benutzt-man-das-animierte-emoji-von-apple. [accessed 5 October 2021].

Universal Declaration of Human Rights (UDHR) 1948, (resolution 217 A), adopted 10 December 1948.

Veith, W. (2004). Gemeinwohl. In M. Heimbach-Steins (Ed.), Christliche Sozialethik 1 (pp. 270-282). Pustet.

Wikipedia. (2021). Human evolution. https://en.wikipedia.org/wiki/Human_evolution [accessed 5 October 2021].

Zsifkovits, V. (1980). Gemeinwohl. In A. Klose, W. Mantl, & V. Zsifkovits (Eds.), Katholisches Soziallexikon (2nd ed., pp. 854-861). Verlag Tyrolia; Verlag Styria.

Zsifkovits, V. (1990). Grundprinzipien der katholischen Soziallehre. *Theologisch-Praktische Quartalsschrift*, *138*(1), 16-25.

Zsifkovits, V. (2012). Orientierungen für eine humane Welt. Zeitdiagnosen: Bd. 27. LIT.

Zweig, K. A. (2019). Ein Algorithmus hat kein Taktgefühl: Wo künstliche Intelligenz sich irrt, warum uns das betrifft und was wir dagegen tun können (Originalausgabe).