



AI Ethics and the Capability Approach

Michael Litschka, Digital Media Technologies, Univ. of Applied Sciences St. Pölten, AT, michael.litschka@fhstp.ac.at

This contribution deals with recent theoretical approaches in AI (artificial intelligence) ethics and their (possible) connection with Amartya Sen's capability approach as well as my application of it to media and AI capabilities. While such an integration is considered on a general level in some AI ethics publications, a more detailed analysis of media and AI capabilities is still an ongoing project in the field of applied ethics. When looking for solutions to ethical dilemma situations arising from the development and use of new AI technologies, further cooperation between the social sciences, economic-ethical concepts such as the capability approach, and more traditional philosophical approaches considering deontological principles, virtues, and justice is needed.



1. Introduction

Artificial intelligence (AI) will affect our social coexistence to an extent that is still underestimated. On the one hand, hopes are high: Robots that take over dangerous tasks for us (e.g., demining) or facilitate the care of sick people; computer-aided medical applications (e.g., the analysis of millions of medical images); forecasting models (e.g., climate change predictions); more efficient value chains in content production and distribution; etc. Many other applications are conceivable and are being discussed in computer science, economics, media and communication studies, sociology, political science, and related fields. On the other hand, questions about the social impact of AI-supported applications and the various ethical dilemmas that we face as developers, users, regulatory institutions, and policymakers as well as ordinary citizens need to be asked. Such dilemmas are reinforced by the use of "large language based models" such as Chat GPT, algorithm-based business models of digital platforms, unfair applications of forecasting models, privacy and data protection issues, lack of social feedback on the models, or the fear that an AI system may become too autonomous. In the case of so-called social media, the dangers of fake news dissemination, hate speech communication, and filter bubble production are becoming ever more prevalent. Challenges must be met by multiple stakeholders: Users of software solutions need to understand (and agree with) the quality, transparency, and underlying values of the used algorithms; companies and authorities need to address the distribution of responsibilities, the level of risk they will accept, and the kind of regulation the respective problem demands (self-, co-, or external regulation); the general public will have to acquire AI competences ("AI literacy") to understand the opportunities and risks of AI tools and use cases. All these stakeholders can be addressed by thinking through the implications of the "capability approach," as envisioned¹ by Amartya Sen (e.g., Sen 1987, 1992), and by further developing this approach to a concept of "media capabilities" (Litschka 2019).

The following section (2) presents an overview of Sen's approach in its original economic-ethical understanding. Section 3 describes my application of the concept to the field of media reception and AI competencies. Section 4 discusses actual and possible fields of integration of either approach in current AI ethics literature. While literature in business ethics has been more focused on the possibilities of Sen's theory to overcome weaknesses of utilitarian ethics and to inform economic policy how we may be better able to measure economic inequality and justice, literature on media and AI ethics has been inclined to develop ethical design principles, virtues of programmers,

¹ I will focus on Sen's take on capabilities while of course acknowledging that the concept has been co- and further developed by others including Martha Nussbaum (e.g., Nussbaum 2006).

managers, and users, as well as deontological categories like discourse and process orientation (with justice being a part of this stream). Section 5 concludes by looking at possible common grounds of the described approaches.

2. The Capability Approach: Basic Features and Fields of Application

Some fields in the social sciences, e.g., communication science or economics, look at the problems mentioned above primarily from the perspective of changes in the use of AI, the influence of these changes on society (and the human condition), socio-technical phenomena that are based on digitization, and the choice of technologies by human beings. Examples for such approaches would be mediatization theory (e.g., Krotz 2001) or the uses and gratifications approach (e.g., Katz et al. 1974; see also Karmasin and Litschka 2013 for a critique). But the influence of AI as a socio-technical phenomenon affects our freedom of choice, life chances, and democratic cohesion—developments which can be well comprehended with the capability approach. Can we always consciously and actively influence the way we (fairly and transparently) interact with new AI technologies? The basic idea behind that question is the importance of the possibilities and chances that people have in order to understand these developments and, subsequently, make an informed choice (of technologies and uses).

The capability approach is an economic and philosophical theory countering some perceived problems of mainstream (neoclassical) economics and a possible solution to the limited information base of utilitarianism by incorporating the basic rights, abilities, and choices of individuals into its framework. The approach's focus is on people's opportunities for self-realization in economy and society. Sen criticizes "revealed preference theory" and "rational choice theory" as important parts of neoclassical economics for their belief that a complete ordering of preferences and internal consistency of choice would represent the real utility of a person.

[T]his approach presumes both too little and too much: too little because there are non-choice sources of information on preference and welfare as these terms are usually understood, and too much because choice may reflect a compromise among a variety of considerations of which personal welfare may be just one. (Sen 1977, 92f.)

In contrast to utilitarianism (which adds up utility sums and, in this way, considers people to be equal), Sen (1987, 1992) expands the "information basis" of his evaluative theory. Some information is included when making a judgment and some is (often implicitly) excluded; for example, utilitarianism excludes information bases other than "utility." In addition to the "well-being" of a person, i.e., the personal benefit they gain from an action, it would be equally important to analyze the "agency" aspect of

the same person, i.e., their ability to form goals and values, possibly without being able to derive a benefit from them. In addition, social contingencies can completely distort the notion of "utility" (e.g., due to low social status). According to Sen, the neoclassical economic approach also disregards the fact that freedom itself can be a deontological category. Freedom and the alternative courses of action it makes possible can certainly have an intrinsic value and therefore always has two components (Sen 1999, 198f.):

- the opportunity aspect: freedom helps us to achieve the goals we choose (Sen calls this "well-being");
- the process aspect: freedom gives us control over our choices, regardless of which choice of goals we ultimately make (Sen calls this "agency").

According to Sen, it is not the consequences of actions that should be the most important point of reference for ethics, but the freedoms and possibilities of the individual to pursue their goals. Sen calls these possibilities "capabilities," the freedom of choice that individuals have and the ability to make use of them. Even if we do not choose an alternative, it is important to have it (such as in the example of starvation and fasting: the first is forced, lacking an alternative, the second voluntary, a conscious choice). Our ability to convert resources into goals varies greatly, depending on age, gender, genetic dispositions, disabilities, etc. (Sen 2003, 96). It is unlikely that an equal distribution of basic goods, as Rawls (1999) envisioned, will also result in equal opportunities for individuals to realize their goals. Freedom is linked to means and ends, and neither equality of ends nor equality of means will guarantee equal freedoms (Sen 1992, 85ff.). If we want to judge the quality of life in a mediatized world, it is therefore important to take

note not only of the ownership of primary goods and resources, but also of interpersonal differences in converting them into the capability to live well [...]. This approach focuses on the substantive freedoms that people have, rather than on the particular outcomes they come up with (Sen 1999, 192).

The capability approach has been criticized for its paternalistic stance (e.g., Sugden 2008), as it may call for policies which restrict individual liberties and further an ideal of "democratic control," similar to the "nudging" approach as promoted by, e.g., Thaler and Sunstein (2008). I would argue that the difference between "positive" (freedom to act) and "negative" (absence of coercion) freedom should be considered here: Sen wants to strengthen positive freedoms, not weaken negative ones. Another critique concerns the comparative view within the capability approach, i.e., whether we should stick to a perfect principle of justice (as, e.g., Rawlsian ones) or compare

states of justice on an empirical basis (which Sen suggests). Valentini (2011) argues that "perfect" justice should not be neglected based on Sen's arguments as they still allow for comparative judgements when using Rawls' concept of a "reflective equilibrium." However, it is beyond the scope of this paper, which focuses on concrete applications of the capability approach in media and AI ethics and pertinent discussions in literature, to delve into the finesses of such critiques.

3. Media Capabilities: Concept and Reach

According to the capability approach, we are not only benefit-orientated but also interested in agency aspects when using AI technologies—in other words, the number of "functionings" (or combinations of functionings) that we can (but do not have to) achieve matters. This cannot be left to the individual alone, but requires socialization by parents, schools, universities and the like, as well as conceived policy measures, e.g., economic, education, or technology policies. Taking the capability approach seriously, these agents must not rely solely on resources and skills (i.e., the opportunity aspect), but empower people for these tasks (e.g., by means of specific capabilities). In a paper from 2019, I applied this logic to the concept of "media capabilities" (Litschka 2019). In this section, I will restate my line of thought and suggest developing an analogous understanding of "AI capabilities" to be developed in further research.

Let us start with a conceptual difference between the traditional concept of media competencies and the idea of capabilities. Seen as ability for media critique, knowledge about media, their system and use, knowledge of media production, intermediation of communication, and other elements in the history of media pedagogy, media competencies (see, e.g., Moser 2010, 241ff., for an overview and critique) are an important part of our individual dealings with technology. However, I argue that media competencies are analytically bound (too much) to the individual and do not convey enough information on this specific individual's possibilities to handle, for instance, AI technologies. While, for instance, critical theory clearly states that media literacy is not only conceived and taught individually but also in a socio-critical way (e.g., Kellner and Share 2019), in a capabilities perspective we need to give all levels of responsibility (micro, meso, and macro) equal regard. It would not be enough to assign people individual responsibility for media competency and not to conceive corresponding organizational and political incentive systems to make such responsibility feasible. So, while the concept of media competencies and media capabilities share some of the same societal goals (such as enabling citizens to make informed election decisions and participating in sensible discourses about politics and society), they stress different responsibility patterns and decision processes.

Sen also suggests further analyses of communicative rationality (similar to Habermas 1991), choice, and freedoms to understand the complete normative picture of media reception and different responsibilities distributed on several levels in the media economy. Even if the analytical focus stayed on the individual, economically rational arguments in utilitarian theories like the uses and gratifications approach would not be able to grasp important normative concepts like obligations of individuals, the embedding of individuals in a responsive media society, or mass media as social institutions that cannot be changed by isolated individual decisions (Christians 2007). Only deontological (and discursive) theories and, in my opinion, capability-orientated frameworks can give us such a complete normative view.

Following this view, capabilities at the individual level can be interpreted as media and AI literacy and competence in the sense of an ability to choose and consume media and AI offerings that satisfy our needs. We therefore have the opportunity (and the consumption capital) to deal with media and AI goods and services in a self-determined way in order to increase our well-being. This ability would then feed directly into our utility function, which encompasses the opportunity aspect of media capabilities. The goals ("functionings" according to Sen 1987) that we can achieve with this competence could be, for example, the status of a well-informed and educated person, a higher social status, or simply a higher salary in the respective job environment.

A complete understanding of media capabilities, as introduced above, must also include the "agency" or process aspect of reaching one's goals. This encompasses all possibilities of choice and the multitude of functionings (or combinations of functionings) we might want to reach (but need not actually choose to reach). As far as media reception is concerned, a recipient might be interested in participating in democratic election processes and political discourses, without a bystander being able to attach such behavior to economic rationality, as the utility function of the recipient is not involved. One of the possible concepts which grasp such behavior could be "commitment" as the adherence to (group) rules, in this example the contribution to democratic cooperation and values in one's home country. While media pedagogic approaches stress the abilities and competences of people, media capabilities are about "being enabled to do something." What would be needed in a mediatized economy to make this approach realistic (see Litschka 2019)?

First of all, there is the need for a critical mass of media diversity representing the most diverse points of view and values. Moreover, a basic media education for the consumption and reception of media is necessary in order to be able to build up the necessary consumption capital (Kiefer and Steininger 2014). This task cannot be taken over by the individual alone but lies in the socialization that parents, schools

universities, and the like provide. In addition, the responsibility of media enterprises must be addressed as they often decide upon our chances to get access to media products, influence our world views through public relations and advertising, and concentrate media power in a few platforms. Only the enabling of people to actually use their possibilities of choice in the media economy—that is, real media capabilities as realization chances—will convert basic rights and freedoms of media usage into functionings. Compared to the concepts of media competencies (e.g., Moser 2010), media literacy (e.g., Buckingham 2017), and cultural competencies within media literacy (e.g., Jenkins 2009)—though all trying to go beyond the analysis level of individual behavior—the media capabilities approach additionally includes justice and publicity deliberations.

Considerations of justice have always played a part in the debate on the normative problems of AI technologies. Concepts to be mentioned here are the equal and open access to AI applications (no digital divide between wealthy/educated and less wealthy/educated people) or that there be no inequalities in the treatment of people through bias or biased input data. Rawls (1999, 2001) and his "theory of justice" or related and further developed positions such as Amartya Sen's (2010) "comparative justice" are repeatedly used for analyzing these issues. The former theory emphasizes equality of opportunity through a principle of the "veil of ignorance" in decision-making and develops concrete principles of justice, i.e., the greatest possible freedom for individuals as long as this does not interfere with the freedoms of others, and restriction of inequality through the fair equality of opportunities for all and the relative betterment of the worst-off in the population. The latter theory deals with the way certain institutions work and the actual behavior of people in order to achieve a gradual improvement in their living conditions without the need for fundamental principles. In both approaches, communicative reason, already emphasized by Habermas (1991), plays a central role as the required form of rationality (in contrast to, e.g., economic rationality). Media capabilities can enhance justice in a media society by demanding a pivotal role for global mass media to make public discourse possible, strengthen the role for disadvantaged communities, build values through open discourses, use all information available also for interpersonal comparisons of states of well-being, and last but not least stress meso- and macro-level responsibilities like the corporate social responsibility of media and AI companies or regulatory responsibilities of media and technology policies.

The following section reviews literature on AI ethics with regard to possible and potential inclusion of the capabilities approach. While the media economy is in the focus of my media capabilities concept, we now place AI technologies and businesses

in the center of attention. As such we could probably speak of "AI capabilities" without principally changing the nature and theoretical underpinning of the above-described line of argument.

4. Recent Literature on Capabilities: Halfway to AI Capabilities?

Recent literature on AI ethics incorporates many philosophic-ethical approaches to deal with a wide range of problem areas, including interdisciplinary legal, sociological, economic, political, and technological issues (see Litschka and Krainer 2019, Rath et al. 2019, Dubber et al. 2020, and Veliz 2024 for current overviews). To begin with, social media communication can endanger democracy per se, because it is performed on platforms which potentially spread fake news, deep fakes, or hate speech (see Litschka et al. 2024b) and may generate echo chambers and filter bubbles (Pariser 2011). There have also been discussions of the loss of trust in professional journalism, legacy media, and political institutions as well as possible copyright infringements (e.g., with learning data) and election manipulations. Increased attention has recently been on the addictive risks of media and AI usage, privacy problems with data-based systems, and the general lack of transparency of algorithms and decisional freedoms in the sense of autonomy (see, e.g., Spiekermann 2019 and Vallor 2024). Related philosophical problems are the questionable moral status of AI, the issue of the responsibility of AI systems and robots, the adaption of user behavior to expected behavior (as self-fulfilling prophecy), and missing concepts of AI justice (see, e.g., Coeckelbergh 2022 and Litschka et al. 2024a).

In the following, I will summarize some major ethical stances in recent AI literature and whether they are connected to the basic ideas of the above-described capability approach in general and specifically to media (or rather AI) capabilities.

Virtue ethics approaches (e.g., Cohen 2012, Spiekerman 2019, Ess 2020, Vallor 2024) deal with the question of what character traits one must have and exercise in order to lead a good life and how artificial intelligence can either promote these traits or at least be programmed in such a way that these virtues are not made impossible. We need to ask ourselves what kind of person we should be or become in order to be satisfied in the constant pursuit of our technological interactions, i.e., to be able to achieve so-called "eudaimonia." This is about a new awareness of values in dealing with digital technologies and the question of the "why" of new technological developments. To what extent, according to one of the central questions, can AI incorporate values or enable us to pursue our self-determined values? Moreover, we need to understand that AI technology "mirrors" our preconceived (and often falsely balanced) values and by the same token—i.e., by changing ourselves—could be moved in a better direction (Vallor

2024). An example would be the constant distraction caused by social media and the lack of focus or even the excess of stress it can create. An important value here would be the ability to set more focus time and less distraction by default. Such approaches also concern the question of how our values can be prioritized, balanced, and pursued in a self-determined way.

Following virtue-ethical thinking in AI development, authors like Dignum (2019) or Spiekermann (2019) argue for design principles for software and hardware that include certain value-based principles such as accountability, responsibility, and transparency. They maintain that ethical considerations should be reflected in the approval and development processes (ethics in design), in the decision-making processes of AI systems (ethics by design), and in considering the impact of the systems on their integrity (e.g., codes of conduct; ethics for design). Taking Spiekermann's (2019) virtue ethics solution, which implies a new awareness of values in our dealings with (digital) technology, as an example, we see a connection of virtues and capabilities, because according to virtue ethics the correct balancing of virtues by a person helps achieving the "telos" of their life, thereby making them capable of choosing the right technology offering. She cites as an example the constant bombardment of communicators active on social networks and the lack of focus, depth, and completeness that can be experienced on these platforms. This needs to be improved through more careful (value-conscious) programming, different settings, and more conscious use of technology. In a similar vein, Floridi et al. (2019) suggest incorporating the principles of beneficence, non-maleficence, autonomy, justice, and explicability in intergovernmental guidelines for AI technology development.

While virtue ethics as direct descendent of Aristotle's and Plato's early philosophical concepts is not about the enhancement of people's capabilities to prosper in a just society and also does not conceptionally use the social choice logic inherent in the capability approach, we might still argue that virtues connected with technologies (e.g., in ethics by design approaches, see Dignum 2019) enable us to use AI to strive for "eudaimonia" (happiness) through the development of specific character traits. However, the concept is rooted in individual ethics while the capability approach places more importance on organizational and political endeavors to raise capabilities.

Deontological approaches (e.g., Floridi et al. 2018, Thimm and Bächle 2019) emphasize the centrality of human dignity and the unconditional opportunity to realize ourselves as persons even in a digitalized world. This claim must be universalizable, i.e., it must be possible to formulate it as a law valid for all participants. For example, profiling travelers at airports according to their appearance is banned for good reason in the EU's AI Act (EU 2024), as it is considered too risky (because of bias, distortion,

prejudice) and cannot be applied universally. The debate about the autonomy of our decisions and those of AI also plays a major role in deontological approaches. Autonomy is seen as a condition for the possibility of assuming ethical responsibility and is (at least at present) by no means guaranteed by the fact that machines can (or could) make independent decisions in various areas. Subjects and objects of responsibility in AI issues are distributed in a network manner and are complex. In any case, Kantian autonomy requires far more than can currently be achieved by machines.

The capability approach itself stresses the importance of deontological categories while acknowledging the necessity of judging consequences of policy measures in order to compare different social states. By doing so, it emphasizes "processes" on the way to reaching specific goals, something that utilitarianism deems unnecessary. Sen (1985, 4), for example, gives a procedural understanding of the importance of having markets in a society:

If this rights-based, "procedural" view is accepted, then the traditional assessment of the merits and demerits of the market, in terms of the goodness of outcomes, would be quite misplaced. The moral necessity of having markets would follow from the status of rights and not from the efficiency or optimality of market outcomes.

In accentuating "rights" of people as an unalienable aspect of a person, Sen imports a deontological tenet into his theory. As far as the market for AI is concerned, in much of economic literature, but also embodied in statements of big tech managers in the industry, we find utilitarian thinking and a strong belief in either the functioning of markets or marketplaces of ideas (Karmasin and Litschka 2013). While from a perspective of stressing the innovative potential of functioning technology markets and deterrence against too much state influence these theories are reasonable, they do not seem to be able to tackle the multifaceted dilemmas encountered by current uses of AI by organizations and citizens.² We have, for instance, already witnessed the problems that market concentration can cause on platform markets and in the social media industry (see, e.g., Litschka et al. 2024a). Utilitarian ethics does not seem to help alleviate these problems.

An important discussion within virtue and deontological ethics is the questionable moral status of AI (Coeckelbergh 2020, 50–60). As long as we cannot foresee the actual possibility of normative machine thinking (e.g., Rath et al. 2019), we might not want to give AI "full moral agency," and as long as machines do not have mental states, emotions, or free will to make decisions, it would also be inappropriate to assign them such a status. While they may be much faster in applying principles to specific

² For a thorough analysis of weaknesses in utilitarian theory, see, e.g., Sen and Williams 1982.

cases, the kind of moral reasoning we expect humans to have—at least in theory—is of a different nature, especially when weighing conflicting interests. "Many people think that moral agency is and should be connected to humanness and personhood," states Mark Coeckelbergh (2020, 54): "They are not willing to endorse posthumanist or transhumanist notions." In a similar vein, Powers and Ganascia (2020, 40f.) are doubtful of real moral AI agency as attempts to model ethical reasoning into machines (e.g., deontic logic and incorporating all possible consequences of an action, respectively also embracing conflicting ethical frameworks) have not been convincing so far. Applying a capability approach to these reservations, AI capabilities may only be found in humans, and strengthening these capabilities in the above-described sense of agency is the foremost task of AI and media education in schools, universities, and related companies.

Coeckelbergh (2020, 163) also addresses the issue of our pluralist views on AI ethics. Culture, power, and political processes influence our stance on specific ethical issues. This problem (of value pluralism and value aggregation) has been, among others, taken up by social choice theorists like Arrow (1997) and Sen (1999), referring to the diverse value positions that exist in our society due to different cultures, educational backgrounds, traditions, etc. This problem could be observed, for example, in the freedom vs. security or health debate during the COVID pandemic: both sides had arguments for or against tough policy measures, and both values are important in our liberal democracies. As we must expect this problem to arise during our discourse on AI ethics as well, the capability approach suggests an "impartial spectator" weighing pros and cons of diverse viewpoints, but only if the necessary (interpersonal) comparisons of social states based on an enlarged informational basis (beyond utility) can be secured.

5. Common Grounds: AI Ethics and Capabilities

While the rising amount of literature on AI ethics renders it almost impossible to review the whole field in relation to the appearance of any form of capabilities, I would still suggest the following appraisal: concrete applications of the capability approach or at least its offsprings like media or AI capabilities are rare.³ However, as I have tried to show, a number of possible connections of established AI ethics and capabilities approaches are conceivable as the following thoughts will exemplify.

Virtue ethics and capabilities share the goal of making a person and its character more "complete" by revealing the complex motivational basis of our decisions. The ability to reach goals beyond utility maximization (e.g., because commitment is a

³ See, however, this very recent capability approach orientated non-principlist appraisal of medical AI tools by Ratti and Graves (2025) based on Martha Nussbaum's thought.

driving factor, not utility), can be developed by adopting the right character traits (virtues) and by having possibilities to choose from (capabilities). The major difference is the individualistic stance of virtue ethics compared to the important enabling tasks of organizations and policies in the capability approach.

Deontological approaches stress the importance of unalienable rights and the role of autonomous, rational, and universalizable decision-making—comparable to the capability approach's focus on a rights-based view of the economy where procedures, processes, and the right kind of rationality need to be taken into consideration in addition to outcomes of interactions. AI ethics is often concerned with issues of justice, which Sen (2010) understands as a universalized concept. He demands that all possible viewpoints (and not only regional or "near" ones) are included so that publicly deliberated judgments (and interpersonal comparisons of social states) can be ameliorated. Public reasoning—made possible, among others, by a functioning system of global media—is not only the keystone of democracy, but also of universal social contracts. It is important for the future development of AI ethics to embrace this kind of reason: while rationality demands arguing our reasons in front of ourselves, reason demands that our arguments hold in front of all others. If we want to include different cultures and traditions in AI development and usage (see also Ess 2020) in order to reach at least partly universalizable agreements on AI regulations, this discursive-ethical principle is still valid.

A further connection (and possible future research area) between capabilities and AI-ethical approaches is the concept of value pluralism. For Sen (e.g., 2010, 37), justice arises from applying communicative rationality to the comparison of social states (contrary to developing a principle of "perfect" justice as in Rawls' theory). This way, we can (1) compare different situations according to their status of being just or unjust, (2) judge actual social improvements, (3) go beyond focusing on one's home country, making the viewpoints of other nations matter ("open impartiality"), and (4) may choose very diverse principles of justice in an original state, because even well-reasoned norms and values are pluralist.

Regarding regulation, the pluralistic and divergent principles of ethics that are recognized by and acceptable for society can, in my view, only be reconciled by public and impartial use of reason. This use of reason, e.g., by applying the Smithian notion of the "impartial spectator," can generate principles of justice, virtues, capabilities, autonomy, and other philosophically founded ethical norms. In practice, this implies cooperation of AI companies, AI users, and regulatory authorities, e.g., by organizing public stakeholder dialogues. The role of independent and globally active media in arguing for and distributing accepted values has also been stressed by Sen (2010, 201),

and authorities or governments must ensure that this exchange of information is not hindered by concentrated media and platform markets. Recent regulation activities like the DMA (Digital Markets Act), DSA (Digital Services Act), or the AI Act on European level are first steps into this direction. Kirchschräger (2021), for example, has also called for international regulations of AI development and would probably agree with Sen that such policies should overcome regional and culturally parochial points of view.

Summarizing these connections of the capability approach to recent AI-ethical approaches and thereby concluding this article, I suggest the following common grounds:

1. "Publicity" must apply to AI developments, understood here as the creation of an (unlimited) public sphere for the exchange of arguments between equal and free citizens. Citizens must understand the basic structure of the digitized society and its influence on their life chances and (be able to) agree to it. Simply knowing how algorithms work, for example, is not enough, because the principles must be understood and accepted.
 2. This creates an obligation for companies and developers to make AI models accessible to public deliberation. Regarding the fundamental freedoms and capabilities for citizens as demanded by Sen, it can be assumed that algorithmic discrimination is contrary to the right of equal citizenship for all.
 3. Sometimes the active elimination of inequalities might be required, that is, substantive rather than procedural equality achieved. If private (tech) companies fear that too much equality and state intervention might inhibit their innovative strength, our answer should be that autonomy and freedom are not undermined by justification processes and regulations, but rather by the possible unjust influence of a technology such as AI on society at large (e.g., Gabriel 2022, 12).
-

References

- Arrow, Kenneth. 1997. "Social Responsibility and Economic Efficiency." In *Ethics in Business and Economics*, edited by T. Donaldson and T. W. Dunfee. Ashgate.
- Buckingham, David. 2017. *Media Education: Literacy, Learning and Contemporary Culture*. Polity Press.
- Christians, Clifford G. 2007. "Utilitarianism in Media Ethics and Its Discontents." *Journal of Mass Media Ethics* 22 (2–3): 113–31.
- Coeckelbergh, Mark. 2020. *AI Ethics*. MIT Press.
- Cohen, Julie E. 2012. *Configuring the Networked Self Law, Code, and the Play of Everyday Practice*. Yale University Press.
- Dignum, Virginia. 2019. *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer Nature.
- Dubber, Markus D., Frank Pasquale, Sunit Das. 2020. *The Oxford Handbook of Ethics of AI*. Oxford University Press.
- Ess, Charles. 2020. *Digital Media Ethics*. 3rd ed. Polity Press.
- EU. 2024. *Regulation 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations*.
- Floridi, Luciano, Josh Cowls, Monica Beltrametti, et al. 2018. "An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds and Machines* 28 (4): 689–707.
- Floridi, Luciano and Josh Cowls. 2019. "A Unified Framework of Five Principles for AI in Society." *Harvard Data Science Review* 1 (1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Gabriel, Iason. 2022. "Towards a Theory of Justice for Artificial Intelligence." *Daedalus* 151 (2): 1–12.
- Habermas, Jürgen. 1991. *Erläuterungen zur Diskursethik*. Suhrkamp.
- Jenkins, Henry. 2009. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. MIT Press.
- Karmasin, Matthias and Michael Litschka. 2013. "Normativität in der Medienökonomie." In *Normativität in der Kommunikationswissenschaft*, edited by Matthias Karmasin, Matthias Rath, Barbara Thomaß. Springer.
- Katz, Elihu, Jay G. Blumler, Michael Gurevitch. 1974. "Utilization of Mass Communication by the Individual." In *The Uses of Mass Communication: Current Perspectives on Gratifications Research*, edited by Jay G. Blumler and Elihu Katz. Sage.
- Kiefer, Marie Luise and Christian Steining. 2014. *Medienökonomik*. Oldenbourg Wissenschaftsverlag.
- Kirchschläger, Peter G. 2021. *Digital Transformation and Ethics: Ethical Considerations on the Robotization and Automatization of Society and Economy and the Use of Artificial Intelligence*. Nomos Verlag.

- Krotz, Friedrich. 2001. *Die Mediatisierung kommunikativen Handelns: Der Wandel von Alltag und sozialen Beziehungen, Kultur und Gesellschaft durch die Medien*. Westdeutscher Verlag.
- Litschka, Michael. 2019. "The Political Economy of Media Capabilities: The Capability Approach in Media Policy." *Journal of Information Policy* 9: 63–94.
- Litschka, Michael, Florian Saurwein, Tassilo Pellegrini. 2024a. *Open Data Governance und digitale Plattformen: Ethische, ökonomische und regulatorische Herausforderungen und Perspektiven*. Springer.
- Litschka, Michael, Florian Saurwein, Tassilo Pellegrini. 2024b. "Digitale Plattformen und offene Daten: Theoretische Ideale vs. praktische Kontextbedingungen." In *Digitalisierte Massenkommunikation und Verantwortung: Politik, Ökonomik und Ethik von Plattformen*, edited by Michael Litschka, Claudia Paganini, Lars Rademacher. Nomos.
- Litschka, Michael and Larissa Krainer, eds. 2019. *Der Mensch im digitalen Zeitalter: Zum Zusammenhang von Ökonomisierung, Digitalisierung und Mediatisierung*. Springer.
- Moser, Heinz. 2010. *Einführung in die Medienpädagogik: Aufwachsen im Medienzeitalter*. 5th ed. VS Verlag.
- Nussbaum, Martha C. 2006. *Women and Human Development: The Capabilities Approach*. Cambridge University Press.
- Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.
- Powers, Thomas M. and Jean-Gabriel Ganascia. 2020. "The Ethics of the Ethics of AI." In *The Oxford Handbook of Ethics of AI*, edited by Markus D. Dubber, Frank Pasquale, Sunit Das. Oxford University Press.
- Rath, Matthias, Friedrich Krotz, Matthias Karmasin, eds. 2019. *Maschinenethik: Normative Grenzen autonomer Systeme*. Springer VS.
- Ratti, Emmanuele and Mark Graves. 2025. "A Capability Approach to AI Ethics." *American Philosophical Quarterly* 62 (1): 1–16.
- Rawls, John. 1999. *A Theory of Justice: Revised Edition*. Harvard University Press.
- Rawls, John. 2001. *Justice as Fairness. A Restatement*. Harvard University Press.
- Sen, Amartya. 1977. "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory." *Philosophy and Public Affairs* 6: 317–44.
- Sen, Amartya. 1985. "The Moral Standing of the Market." *Social Philosophy & Policy* 2 (2): 1–19.
- Sen, Amartya. 1987. *On Ethics and Economics*. Oxford University Press.
- Sen, Amartya. 1992. *Inequality Reexamined*. Oxford University Press.
- Sen, Amartya. 1999. "The Possibility of Social Choice." *American Economic Review* 89 (3): 349–378.
- Sen, Amartya. 2003. *Ökonomie für den Menschen: Wege zu Gerechtigkeit und Solidarität in der Marktwirtschaft*. 2nd ed. DTV.
- Sen, Amartya. 2010. *The Idea of Justice*. Penguin.
- Sen, Amartya and Bernard Williams, eds. 1982. *Utilitarianism and Beyond*. Cambridge University Press.

- Spiekermann, Sarah. 2019. *Digitale Ethik: Ein Wertesystem für das 21. Jahrhundert*. Droemer.
- Sugden, Robert. 2008. "Why Incoherent Preferences Do Not Justify Paternalism." *Constitutional Political Economy* 19: 226–48.
- Thimm, Caja and Thomas C. Bächle. 2019. "Autonomie der Technologie und autonome Systeme als ethische Herausforderung." In *Maschinenethik: Normative Grenzen autonomer Systeme*, edited by Matthias Rath, Friedrich Krotz, Matthias Karmasin. Springer.
- Valentini, Laura. 2011. "A Paradigm Shift in Theorizing About Justice? A Critique of Sen." *Economics & Philosophy* 27: 297–315.
- Vallor, Shannon. 2024. *The AI Mirror*. Oxford University Press.
- Veliz, Clarissa, ed. 2024. *The Oxford Handbook of Digital Ethics*. Oxford University Press.

