Legal and Ethical Challenges in Multimedia Research

Vivek K. Singh
Rutgers University

Elisabeth André
University of Augsburg

Susanne Boll
University of Oldenburg

Mireille Hildebrandt
Radboud University, Netherlands and Vrije Universiteit Brussels

David A. Shamma
FX Palo Alto Laboratory

***Abstract*—Multimedia research has long moved beyond laboratory experiments and is being rapidly deployed in real-life applications including advertisements, search, security, automated driving, and healthcare. Hence, the developed algorithms now have a direct impact on the individuals using the abovementioned services and the society as a whole. While there is a huge potential to benefit the society using such technologies, there is also an urgent need to identify the checks and balances to ensure that the impact of such technologies is ethical and positive. For instance, if the multimedia technologies are being used to detect and protect pedestrians from accidents by autonomous vehicles, then the pedestrian detection performance needs to be equitable across demographic descriptors, such as gender and race of the pedestrians. Similarly, while logs of driving behaviors are important in many applications, making such information available to corporate entities and third parties could raise important privacy challenges. This position article aims to: first, increase the awareness of such concepts and existing legal constraints in the multimedia research community, second, initiate a discussion on community guidelines on how to conduct multimedia research in a lawful and ethical**

Digital Object Identifier 10.1109/MMUL.2020.2994823
Date of current version 12 June 2020.

Published by the IEEE Computer Society

manner, and third, identify some important research directions to support a vision of lawful and ethical multimedia research.

Recent growth spurt in multimedia research has led to some exciting developments in terms of multimedia content understanding and search, self-driving cars, and medical analysis. At the same time, there have been reports questioning both the processes and the outcomes for the developed technologies. For instance, Metz questions the ethics of using public image data in YFCC100M and IBM's Diversity in Faces datasets for training face recognition algorithms.[1] Under EU data protection law, any use must have a specific purpose and be limited to that purpose, while also requiring a valid legal basis. This clearly also goes for publicly available data, including images. Similar concerns have been raised about the outcomes of the developed algorithms. For instance, a study by Buolamwini & Gebru has reported that face detection algorithms work much more accurately for white men than dark-skinned women, raising the question whether dark-skinned women should require that their images become part of the training set or resist such inclusion (as it may be used for unwarranted surveillance purposes that have disparate effects for black women).[2] Further, multiple authors have criticized the use of video analysis software for automatic tracking of people in both civilian and military settings.[3] In fact, San Francisco has recently banned the use of face recognition technology for government applications.[4] Meanwhile, the Data Protection Authority in Hamburg, Germany has ordered Google to ban its employees "from listening to and reviewing EU data subjects" voice recordings for three months, to investigate potentially unlawful processing under the GDPR.[6]

Each of these issues raises important ethical concerns and many times the opinions of the experts in the multimedia community might not match with those in the popular media. There is an important need for the multimedia research community as a whole to have free and frank discussions on this topic and be cognizant of the myriad research and activism literature that is available regarding the potential benefits and harms of such technologies.[28–30] In doing so, the multimedia research community should not confuse ethical discussion with legal obligations as many of the so-called ethical concerns have clear answers (and obligations) as per law. In fact, a proper understanding of these obligations could lead the way to actionable respect for fundamental rights and freedoms at the level of research design. This will allow the multimedia research community to identify a set of community norms and guidelines on the processes and the outcomes of the technologies being developed. Developing such an understanding and a set of best practices would allow the multimedia research community to lead the conversation around these technologies rather than reacting to news stories about them (see Figure 1).

In this discussion, we must be careful not to confuse ethics with law. Many of the ethical challenges to be discussed below are part of the legal framework that applies to the multimedia applications based on machine learning. Since May 2018, the General Data Protection Regulation (GDPR) applies to the processing of personal data of people in the EU, whether or not the processing is done in the EU or by a company established in the EU. When relevant, we will discuss the requirements of the GDPR, taking into account that applications meant to be deployed within the EU will have comply. The persistent confusion over what legal frameworks apply and what they mean for developers calls for dedicated attention to law for computer scientists; this viewpoint cannot do more than appetize the reader to take a deep dive into why and how law matters for their work.[31] The GDPR is one of the most advanced legislations, and though its scope is significant, many legal frameworks in other jurisdictions may have very different implications. The principles we highlight in this viewpoint are not necessarily anchored and enforceable at the global level. This means that whether and how they are legal or ethical principles is an empirical question. The fact that human rights courts have been weighing the corresponding fundamental rights against economic

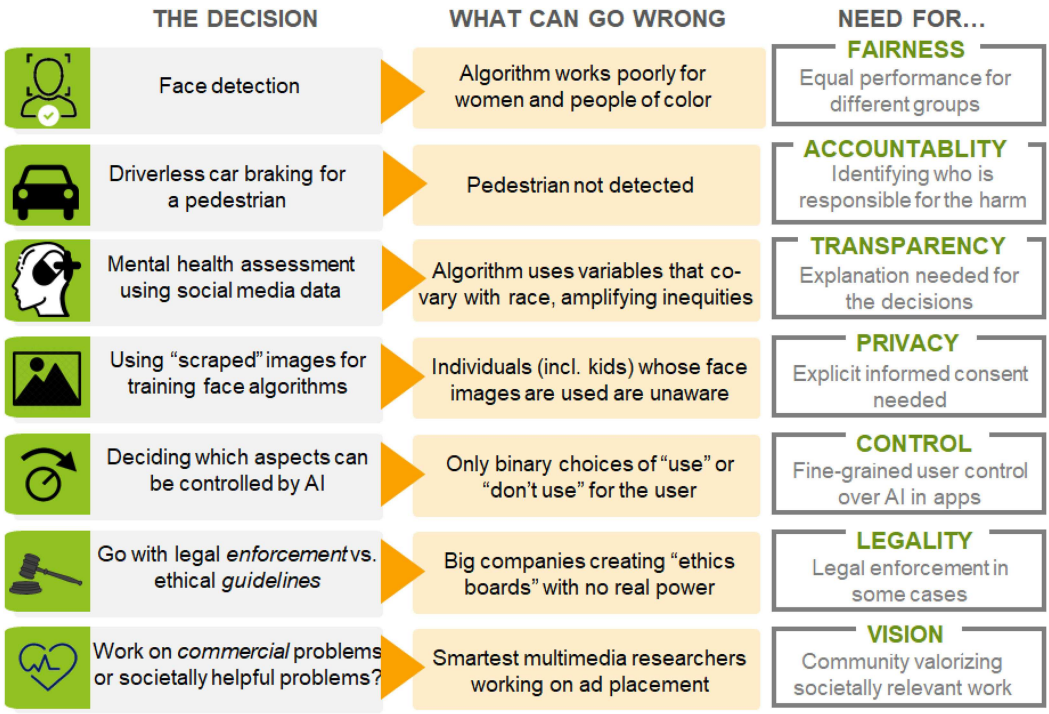

Figure 1. Summary of ethical challenges in multimedia research.

and public security interests for decades means that both ethics and computer science have lots to learn from the judicial scrutiny this has involved. Just like legislatures and courts have lots to learn from technical, scientific and ethical experts.

Fairness

Fairness means that the models developed do not systematically favor or disfavor a particular set of people. Angwin *et al.*[13] showed that parole decision algorithms being used in New York state were much more likely to assign positive outcomes to white defendants than black defendants. Buolamwini & Gebru found facial image based gender recognition algorithms to be much more accurate for white and male individuals than others.[2] Similarly, the dependence of multimedia algorithms for pedestrian detection on the age or race of pedestrians could result in unequivocally unfair outcomes.[25] Multimedia algorithms being used for parole decisions, driving decisions, and security applications can have important life-altering effects on people and it is important to ensure that the outcomes of the algorithms do not systematically favor or disfavor a specific set of people. Obviously, when algorithms detect that a particular characteristic of people correlates with higher risk (of recidivism, defaulting on loans, or causing road accidents), justice authorities, banks, or insurance companies will argue that treating these people differently is fair. An active research community at the intersection of machine learning, law and ethics is involved in this domain.[31–33] It is important that the multimedia community takes note of this and integrates the state-of-the-art design solutions to prevent violations of human rights, such as the right to nondiscrimination. In the research community dedicated to Fairness Accountability and Transparency (ACM-FAccT conferences), this has resulted in raising the more fundamental question of whether and when machine learning should be deployed, warning against computational solutionism.[33]

Accountability

"Algorithmic accountability ultimately refers to the assignment of responsibility for how an algorithm is created and its impact on society; if harm occurs, accountable systems include a mechanism for redress."[26] This is especially important in scenarios where there are multiple humans, companies, algorithms, and algorithm designers involved in the process.[7, 22] For instance, when a

pedestrian is injured due to the decisions taken by a self-driving car, it is important to have accountability in place.[23] Going forward, if people are denied bail or organ transplants unfairly, it would be important to identify accountability in the process. However, liability obscurity associated with the use of modern ML algorithms is considered a major issue.[41] Attempts are being made towards certifying algorithms in order to enhance the transparency of accountability. It is, however, hard to predict whether multimedia systems that make use of highly dynamic ML algorithms will always behave according to a specification. Thus, certification is not a trivial task. Furthermore, it is unclear whether and how existing certification procedures can be adapted to modern multimedia systems. Part of a certification might include the use of state-of-the-art algorithms that check whether a multimedia system is suffering from bias. Finally, it is important to note that responsibility has to be taken over by humans and not by the machine, and usually there is not just one stakeholder in charge of it. The European Commission is currently considering adaptations of the liability regime for AI, making sure that accountability is not dependent on whatever a company may deem ethical, but on the need to compensate harm and damage.[34] This may result in more foresight and help developers propose the state-of-the-art applications that incorporate safety by design and data protection by design. Here again, such accountability may result in the choice not to develop or use certain applications at all, as this may be the only responsible approach, considering the consequences.

Transparency

Algorithmic transparency is the principle that the factors that influence the decisions made by algorithms should be visible, or available, to the people who use, regulate, and are affected by systems that employ those algorithms.[17] Note that transparency and fairness are two different things. It is possible for a decision to be very transparent but not fair (e.g., admit only males) and vice-versa (e.g., a hypothetical neural network that ensures fairness but no one can understand why and how). The GDPR includes a "right to explanation" of decisions made by algorithms, whenever the decision is automated and has a significant effect on individual persons whose data are being processed.[35] In its rudimentary sense, many credit scoring applications in the US, Germany, and other countries identify the factors affecting a person's credit score, though given factors will often operate as a proxy for hidden variables that result in discrimination. Especially when deep learning algorithms have been used, a much higher level of transparency is required to figure out potential discrimination. This concerns behavioral targeting, where the EPIC (Electronic Privacy Information Center) has called for regulations that require advertisers to disclose the demographic factors behind targeted political ads, as well as the source and payment.[24]

Privacy and Data Governance

Multimedia research needs to ensure high quality of results in a way that also ensures human dignity in the process and in the results. This is often presented as a zero-sum game, but that is not necessarily the case.[36] Human subject research in medicine and the social sciences has a long-standing history of "informed consent" from the participants. While the web-based data collection is great for scaling up the studies, there is rarely a notion of "informed consent," i.e., explicitly informing the individuals about all the actions that will be undertaken using such data. Multiple individuals have expressed regret and raised concern upon realizing that their data has been used by machine learning algorithms for training tasks such as face detection.[1] Though the GDPR does not make processing dependent on consent, if consent is used as a processing ground, the GDPR requires that consent is both informed and freely given, and can be withdrawn as easily as given. Moreover, processing that is not necessary for a given purpose is unlawful, whether based on consent or one of the other legal basis. This implies that under the GDPR, repurposing of data processing may be illegal. This implies that an image posted on Facebook or elsewhere on the world wide web cannot be processed for a purpose that was not communicated to the data subject and for which no legal basis applies.

Control

As the AI elements in multimedia are entering myriad applications, an important question is

whether an individual can actually decide and control how much AI is being used. For instance, should the users know that AI is used in a service, be it a natural language based chat bot or an image analysis software, which creates automatic captions for each photograph? Similarly, it would be important for users to have choices beyond the binary "install/don't install" and be able to control the degree to which automatic AI processing is part of the workflow. For instance, one user may want to allow AI in a video-conferencing application for facial identity analysis but not for face touch-ups and vice-versa. We need frameworks that support such a process workflow and make it easy for users to control their choices at different points of time. These are pivotal questions, directly related to human autonomy and dignity, especially in the light of attempts to "nudge" people into compliant behavior "behind their back"—for instance, based on emotion detection in facial images.[40]

Legal Compliance

With the growth of research impact outside the lab environment, there is a need for legal compliance at the level of the design of an application. Laws like the GDPR have made many of the above aspects a critical legal requirement rather than being "good things to do," notably by requiring a data protection impact assessment in case of likely high risk to fundamental rights and freedoms, and by requiring data protection by design to mitigate such risks. These are legal obligations that level the playing field. For instance, imposing these duties on all companies that want to operate in the market, there can be strong economic incentives created for companies to pay keen attention to the consequences of deployment of AI. This should help research communities, such as in multimedia to come up with different types of research design that incorporate the consequences of design choices.

Problems Targeted

Multimedia research has paid significant attention to commercially viable applications such as ad placement and product recommendation but relatively much less effort to societally relevant but less directly marketable applications such as long-term support for education, healthcare or tackling climate change.[27] In fact, "social good" has been identified as a key focus area of multimedia research in a recent NSF workshop on Multimedia Challenges and a column by the Associate Editor-in-Chief in IEEE Multimedia urges researchers to be mindful of social impact of the applications being created.[27,42] Hence, it might be a good time for us as a community to introspect, and prioritize research on socially relevant themes in the coming decade.

EMERGING SOLUTION PATHS AND OPEN RESEARCH QUESTIONS

There are multiple approaches that are painting an optimistic picture regarding multimedia research in addressing each of the abovementioned issues. However, much more work is needed. Here, we identify multiple research questions and research areas that are ripe for exploration and development.

Use of Web-sourced Data for Large Dataset Creation in Relation to Data Protection

One of the biggest drivers for deep learning-based multimedia research is the recent availability of large-scale web-sourced image and video datasets. Clearly, not every image or video that is available on the web should be downloaded and used as part of a dataset. Notably, the GDPR always requires a valid legal ground and an explicit, specified and legitimate purpose for the processing of facial images as they concern identifiable data. Even consent is only valid when given for an explicit, specific, and legitimate purpose. The issue does not end there and there are ethical ramifications of using one's facial data, for say, profiling applications in the future. Although the GDPR has a broad research exception, this mostly applies to research in the public interest, rather than commercial interests. Multiple scholars have argued that even where the GDPR does not apply, just because the data is "public" does not mean it is acceptable for researchers or corporate agencies to reuse it for their purposes.[5] The default prohibition of automated targeting as codified in the GDPR squarely addresses this issue. Therefore an important question for the multimedia research community is to identify the guidelines for

legal compliance and, within the space left open by the law, for ethically creating such large-scale datasets.

Multiple emerging efforts in multimedia research are now focusing on less data-hungry approaches for artificial intelligence. These approaches include the creation of domain-aware (e.g., physics inspired) approaches, zero and one-shot learning approaches, and transfer learning. (e.g.,[19, 20]) While domain (e.g., physics) inspired approaches clearly do not need lots of data to get started, other machine learning approaches are also trying to reduce the amount of new data needed to tackle each emerging problem. However, it is still early days in this space, and the legal and ethical requirements mentioned above clearly call for an important research direction—one on approaches that do not require large datasets.

Informed Consent and Control in Relation to Copyrighted Content and Portrait Rights.

Creative Commons provides an approach for identifying the permissions on what can be done with the images.[37] These licenses concern the copyright of the "author" of the image, not the person depicted in the image. However, Creative Commons licenses were defined before the deep learning and the corresponding opportunities for identifying individuals became commonplace, meaning that we can now assume that *insofar as facial images are concerned, those depicted have a so-called portrait right in the picture, which concern their privacy right rather than the photographers copyright.*

One of the possible settings considered during a panel discussion[43] on this topic at the 2019 ACM Multimedia conference was a "No AI" permission setting, which would make it illegal for algorithms to use the image for training of sophisticated face matching algorithms. However, the solution is not as simple as it appears. For instance, does the above "No AI" tag also include image cropping, touch-up, lighting, or other filters? When does the processing become "AI" is not one with a clear definition, and understanding the user's perception/understanding of what they are signing up for remains an important research problem. Understanding this would require work by those who not only understand the underlying technology but also understand the human perspective on these topics. Under the GDPR the issue plays out differently, as consent can only be provided for a specific purpose. Providing consent for "whatever" processing as long as it does not involve AI would be invalid.

Another point that generated large agreement in the panel discussion was providing users the ability to withdraw consent at any later point of time, as is now required under the GDPR. Note that under the GDPR, the mere fact that one has made public one's image does not imply consent for processing by whoever for whatever purpose. Some online systems have started designing web repositories (see OpenPDS, an open source personal data store[8]) that allow for users to remove their data at any point of time. However, in image and multimedia research the issue is more complicated. If there is a model that has learnt using millions of images, does the model also need to be discarded if the consent for (even one of) the supporting images is retracted? This question has informed the work on differential privacy,[38] which solves the problem to the extent that the model will not allow for reidentification. Under the GDPR, this would mean that the model does not qualify as personal data, and therefore the GDPR does not apply to the model. Again, there is a need for more research and identifying community norms in this space and the research findings could inform the legal viewpoint in this space.

Algorithmic Bias

Multiple studies have now accumulated evidence that computer vision and multimedia algorithms can be biased in terms of their performance across demographic groups. The reasons for these biases include the imbalance in training data sets, lack of positive training samples for historically marginalized communities, lack of training data to allow for convergence, and the lack of awareness regarding the leakage of demographic information (e.g., a "moustache detector" hidden in the layers of convolutional neural networks) in the developed algorithms. The default legal prohibition of indirect discrimination on grounds, such as gender and ethnicity may have unexpected repercussions when proxies are used that result in effective discrimination of women or

ethnic minorities. Hence, an important question for the multimedia research community is how to develop multimedia algorithms that support both high accuracy and low bias?

Some of the possible approaches to counter this include those suggesting the use of datasets with balanced representation of people with different demographic characteristics—some of which may be artificially generated, creating adversarial approaches that penalize algorithms for any perceptible bias,[10] and those that propose posthoc adjustment of results for countering bias.[11] As an illustration of this kind of work in multimedia research, a recent paper by Alasadi *et al.*, describes a GAN (generative adversarial network) approach for face matching where one network optimizes for face matching, whereas another network tries to reduce bias. Specifically, the second network tries to infer demographic properties from the hidden layers of the first network and evidence of gender encoding (even when not directly required for the assigned task) is considered evidence of bias. The competition between the two networks yields models that balance accuracy and fairness.[10] We note, however, that since facial recognition systems are sometimes used for surveillance purposes that disadvantage specific groups, it may or may not be in the interest of those groups to become more identifiable. [39]

Explainability and Control of Algorithms

One of the side effects of the development of deep learning approaches is the complexity of the developed algorithms, which comes with the side effect of no human being able to explain the details of the algorithms developed in terms of the features being implemented or the decision rationale. This has costs in terms of interpretability of the models and the lack of transparent causal reasoning for the decisions being made by the system. If such a system needs to make important decisions (e.g., in life and death scenarios in autonomous vehicles) then an explanation of the underlying processes is important. The GDPR requires that automated decisions that seriously affect people are accompanied by meaningful information about the "logic of processing," which implies that such decisions are prohibited if no meaningful information can be provided. Multiple research efforts in machine learning have started focusing on explainability in AI (See[12] for a review). One limitation is that current approaches might help experts to understand the inner workings of ML approaches, but they are of lesser usefulness to domain experts without any ML background. We need a user-centered perspective on the use of AI in applications and systems in which individuals can understand which information and decision is AI supported and if and how they can opt in or opt out.

Community Norms on Research

Given the wide variety (geographic, disciplines, political) of viewpoints represented within the ACM multimedia community, can there be a common set of guidelines that make sense to all researchers? *While inherently difficult, multiple disciplines ranging from nuclear physics to drug testing have come up with globally accepted guidelines for research*. Also, what should be the mechanism for supporting the development of such an ethical framework and how can such guidelines be implemented in the review process? Finally, how do we prevent discussing legal obligations as if they were ethical principles (often framed as "ethics washing"). There is a need for fundamental research as well as organized consensus-building within the multimedia research community to agree on a common set of norms that would be applicable across the globe. Some of these norms could be made part of the paper review/acceptance process in the community going forward. For instance, some communities require access to data to allow for replication of results before accepting research papers. Others insist that the paper cannot be accepted without a formal review by an ethics board. The multimedia research community has been pioneering some efforts on replication of results and perhaps the scope can be broadened to allow for dedicated benchmarks that allow authors and the wider community to reserve their research efforts for work that aligns with the most basic ethical norms (in the case that these norms are not already part of the applicable legal framework).

In summary, there is an urgent need to raise awareness about ethical and legal challenges in

multimedia research. While there are multiple challenges, there are also opportunities to undertake meaningful research, which is technically robust and societally beneficial.

REFERENCES

1. R. Metz, If your image is online, it might be training facial-recognition AI, 2019. [Online]. Available: https://www.cnn.com/2019/04/19/tech/ai-facial-recognition/index.html
2. J. Buolamwini and T. Gebru, "Gender shades: Intersectional accuracy disparities in commercial gender classification," in *Proc. Conf. Fairness, Accountability Transparency*, 2018, pp. 77–91.
3. Facial recognition poses serious risks. Congress should do something about it, 2018. [Online]. Available: https://www.washingtonpost.com/opinions/facial-recognition-poses-serious-risks-congress-should-do-something-about-it/2018/07/18/c4c8973c-89c1-11e8-a345-a1bf7847b375_story.html?utm_term = .8a991491dd14
4. San Francisco just banned facial-recognition technology, 2019. [Online]. Available: https://www.cnn.com/2019/05/14/tech/san-francisco-facial-recognition-ban/index.html
5. K. Albury, "Just because it's public doesn't mean it's any of your business: Adults' and children's sexual rights in digitally mediated spaces," *New Media Soc.*, vol. 19, no. 5, pp. 713–725, 2017.
6. Hamburg regulator bans Google from listening to smart speaker audio, 2019. [Online]. Available: https://globaldatareview.com/article/1195881/hamburg-regulator-bans-google-from-listening-to-smart-speaker-audio
7. Data for Good: FATES, Elaborated, 2018. [Online]. Available: https://datascience.columbia.edu/FATES-Elaborated
8. Y. A. De Montjoye, E. Shmueli, S. S. Wang, and A. S. Pentland, "openPDS: Protecting the privacy of metadata through SafeAnswers," *PloS One*, vol. 9, no. 7, 2014, Art. no. e98790.
9. M. Merler, N. Ratha, R. S. Feris, and J. R. Smith, "Diversity in faces," 2019, arXiv:1901.10436.
10. J. Alasadi, A. Al Hilli, and V. K. Singh, "Toward Fairness in Face Matching Algorithms," in *Proc. 1st Int. Workshop Fairness, Accountability, Transparency MultiMedia*, 2019, pp. 19–25.
11. P. K. Lohia, K. N. Ramamurthy, M. Bhide, D. Saha, K. R. Varshney, and R. Puri, "Bias mitigation post-processing for individual and group fairness," in *Proc. ICASSP IEEE Int . Conf . Acoustics, Speech Signal Process .*, 2019, pp. 2847–2851.
12. M. Du, N. Liu, and X. Hu, "Techniques for interpretable machine learning," *Commun. ACM*, vol. 63, no. 1, 68–77, 2019.
13. J. Angwin, J. Larson, S. Mattu, and L. Kirchner, May 23, 2016, "Machine bias," ProPublica.
14. S. Barocas and A. D. Selbst, "Big data's disparate impact," *Calif. L. Rev.*, vol. 104, 2016, Art. no. 671.
15. C. Dwork, M. Hardt, T. Pitassi, O. Reingold, and R. Zemel, "Fairness through awareness," in *Proc. 3rd Innov. Theoretical Comput. Sci. Conf.*, 2012, pp. 214–226.
16. A. Datta, M. C. Tschantz, and A. Datta, "Automated experiments on ad privacy settings," in *Proc. Privacy Enhancing Technol.*, vol. 2015, no. 1, 92–112, 2015.
17. N. Diakopoulos and M. Koliska, "Algorithmic transparency in the news media," *Digital Journalism*, vol. 5, no. 7, pp. 809–828, 2017.
18. "False Testimony," *Nature*, vol. 557, no. 7707, May 2018, Art. no. 612.
19. Y. Yang, Y. Luo, W. Chen, F. Shen, J. Shao, and H. T. Shen, "Zero-shot hashing via transferring supervised knowledge," in *Proc. 24th ACM Int. Conf. Multimedia*, 2016, pp. 1286–1295.
20. R. Stewart and S. Ermon, "Label-free supervision of neural networks with physics and domain knowledge," in *Proc. 31st AAAI Conf. Artif. Intell.*, 2017, pp. 2576–2582.
21. W. Schreurs, M. Hildebrandt, E. Kindt, and M. Vanfleteren, Cogitas, ergo sum, "The role of data protection law and non-discrimination law in group profiling in the private sector." *Profiling the European. Citizen*. Berlin, Germany: Springer, 2008, pp. 241–270.
22. Ethics guidelines for trustworthy AI, 2019. [Online]. Available: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai
23. Self-driving Uber car that hit and killed woman did not recognize that pedestrians jaywalk [Online]. Available: https://www.nbcnews.com/tech/tech-news/self-driving-uber-car-hit-killed-woman-did-not-recognize-n1079281
24. EPIC Promotes 'Algorithmic Transparency' for Political Ads. [Online]. Available: https://epic.org/2017/11/epic-promotes-algorithmic-tran-1.html

25. A new study finds a potential risk with self-driving cars: Failure to detect dark-skinned pedestrians, 2019. [Online]. Available: https://www.vox.com/future-perfect/2019/3/5/18251924/self-driving-car-racial-bias-study-autonomous-vehicle-dark-skin
26. R. Caplan, J. Donovan, L. Hanson, and J. Matthews, "Algorithmic accountability: A primer," *Data Soc.*, vol. 18, 2018.
27. S. F. Chang *et al.*, "Report of 2017 NSF Workshop on Multimedia Challenges, Opportunities and Research Roadmaps," 2019, arXiv:1908.02308.
28. L. D. Introna, "Disclosive ethics and information technology: Disclosing facial recognition systems," *Ethics Inform. Technol.*, vol. 7, no. 2, pp. 75–86, Jun. 2005.
29. M. Flanagan, D. Howe, and H. Nissenbaum, "Values in design: Theory and practice." *Information Technology and Moral Philosophy*, ed. Jeroen Van den Hoven and John Weckert, Cambridge, U.K.: Cambridge Univ. Press, 2007.
30. I. D. Raji, T. Gebru/ M. Mitchell, J. Buolamwini, J. Lee, and E. Denton, "Saving face: Investigating the ethical concerns of facial recognition auditing," Jan. 2020, *ar*Xiv:2001.00964
31. M. Hildebrandt, *Law for Computer Scientists and Other Folk*. Oxford, U.K.: Oxford Univ. Press, 2020, [Online]. Available: https://lawforcomputerscientists.pubpub.org/
32. S. Barocas, M. Hardt, and A. Narayanan, fairness and machine learning, 2019, [Online]. Available: https://fairmlbook.org/
33. J. Powles, Medium: The seductive diversion of "Solving" Bias in artificial intelligence, Dec. 2018, [Online]. Available: https://medium.com/s/story/the-seductive-diversion-of-solving-bias-in-artificial-intelligence-890df5e5ef53
34. European Commission, COM(2020)64, "Final report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, Feb. 2020 [Online]. Available: https://ec.europa.eu/info/files/commission-report-safety-and-liability-implications-ai-internet-things-and-robotics_en
35. M. E. Kaminski, "The right to explanation, explained," Rochester, NY, USA: Social Science Research Network, SSRN Scholarly Paper, Jun. 2018, [Online]. Available: https://papers.ssrn.com/abstract = 3196985
36. A. Cavoukian, "Privacy by Design and the Emerging Personal Data Ecosystem," 2012.
37. Creative Commons—About The Licenses. [Online]. Available: https://creativecommons.org/licenses/
38. C. Dwork, "Differential privacy," in *Automata, Languages and Programming* ed. M. Bugliesi *et al.*, vol. 4052, Berlin, Germany: Springer , 2006, pp. 1–12, [Online]. Available: http://research.microsoft.com/apps/pubs/default.aspx?id = 64346
39. S. Fussell, The strange politics of facial recognition," The Atlantic, Jun. 2019, [Online]. Available: https://www.theatlantic.com/technology/archive/2019/06/democrats-and-republicans-passing-soft-regulations/592558/
40. A. McStay, "Empathic media and advertising: industry, policy, legal and citizen perspectives (the Case for Intimacy)," *Big Data Soc.*, vol. 3, no. 2, Dec. 2016, Art. no. 2053951716666868.
41. J. Kobielus, "What does it mean to certify an ai product as safe?," 2018, [Online]. Available: https://www.dataversity.net/mean-certify-ai-product-safe/
42. A. Hanjalic, "Multimedia research: what is the right approach?," *IEEE MultiMedia*, vol. 24, no. 2, pp. 4–6, 2017, [Online]. Available: https://www.dataversity.net/mean-certify-ai-product-safe/
43. V. K. Singh, E. André, S. Boll, M. Hildebrandt, D. A. Shamma, and T. -S. Chua, "Legal and ethical challenges in multimedia research," in *Proc. ACM International Conference on Multimedia*, 2019, pp. 2514–2515.