|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2022-2024 | | FG-AI4H-P-039 | |
| **ITU-T Focus Group on AI for Health** | |
| **Original: English** | |
| **WG(s):** | | Plenary | Helsinki, 20-22 September 2022 | |
| **DOCUMENT** | | | | |
| **Source:** | | University of Turku | | |
| **Title:** | | Privacy-preserving AI for synthetic and anonymous health data | | |
| **Purpose:** | | Discussion | | |
| **Contact:** | | Juho Vaiste University of Turku Turku | | Tel: +358405155607 Email: [juho.vaiste@utu.fi](mailto:juho.vaiste@utu.fi) |

|  |  |
| --- | --- |
| **Abstract:** | Data-driven technology solutions offer new opportunities in healthcare, for example in sectors such as disease diagnosis, monitoring and follow-up, and self-care. However, data availability is challenging due in part to privacy restrictions. One proposed solution has been to generate synthetic data via differential privacy techniques that provide a guaranteed level of privacy and advanced privacy protection. New machine learning assisted ways of generating synthetic data have shown promising results to the previous problems with the synthetic data in terms of data accuracy and usability. |

**Ethical assessment of the generation of synthetic data by machine learning in the field of healthcare**

Data-driven technology solutions would offer new opportunities in healthcare, in sectors such as disease diagnosis, monitoring and follow-up, and self-care. However, data availability is challenging due in part to privacy restrictions. One proposed solution has been to generate synthetic data via differential privacy techniques that provide a guaranteed level of privacy and advanced privacy protection. New machine learning assisted ways of generating synthetic data have shown promising results to the previous problems with the synthetic data in terms of data accuracy and usability.

Besides promising outlook for the future, synthetization of data uncovers difficult questions of trade-offs and not the least from the ethical perspective. The trade-offs are addressed from an ethical perspective, with separate questions: What level of privacy can be achieved through synthetization? Is the data in the synthesized data accurate and fair? Who determines the level of privacy parameters used in synthetization, and under what conditions?

The presentation introduces the generation of medical and healthcare synthetic data based on an authentic data source. The study applies differential privacy to data anonymization and generative adversarial networks (GANs) to produce synthetic data. The data used in biobanking and prognostic data on the risk of prostate cancer and cardiovascular disease have been used as data examples.

**Introduction**

Data availability is an essential issue for producing innovations and technology solutions in the healthcare sector.[[1]](#footnote-1) Data availability challenges have been identified, for example, in the analysis of imaging services[[2]](#footnote-2), health data analytics[[3]](#footnote-3), and the production of infectious disease surveillance solutions[[4]](#footnote-4). Private healthcare and information technology companies are also interested in developing their products and services to meet the needs of the self-care and home monitoring sectors, among others. The issue of data availability is also linked to the development of new artificial intelligence solutions[[5]](#footnote-5)[[6]](#footnote-6) and Internet of Things (IoT) devices.

Data synthetization in the healthcare industry is not an invention. However, recent advances in neural networks in machine learning have enabled more promising developments than previous technologies, which have suffered from poor data quality. Various solutions based on neural networks have achieved levels of data quality that are equal to or close to the usability of the original data.[[7]](#footnote-7)[[8]](#footnote-8) Differential privacy methods have also been widely introduced in the processing of health data.[[9]](#footnote-9)[[10]](#footnote-10)

Health data is subject to strict legislation; in Finland, the use of health data is regulated inter alia by the National Data Protection Act, the EU Data Protection Regulation (GDPR) and the Secondary Act on the Secondary Use of Social and Health Data. The ethical assessment of synthetization has a close linkage with the legislation and legal guidelines. However, there are elements within existing laws that are not defined in detail and require an independent ethical review, such as adapting organizational best practices and defining technical specifics. On the other hand, an ethical approach to the issue may open aspects that are not yet covered by current legislation, particularly related to an uncertain future.

Various ethical issues have been identified regarding the generation of synthetic health data. Defining and assessing a sufficient level of anonymization is an essential question from the privacy perspective. When generating synthetic data, the original data is anonymized using differential privacy techniques. Differential privacy has become commonplace as a technology to protect privacy. It can be used to produce an information theory promise of privacy, a calculation of the probability that the privacy of the elements of the target data material is protected.[[11]](#footnote-11)

The synthesized data can also reproduce or amplify the bias of the original data, producing unbalanced results. Furthermore, synthetic data generation can leave latent properties in the data, which can be complicated and rudimentary to monitor. In addition, although the evaluation of synthesized data relative to the original data is in principle a technical issue, a lack of understanding of the evaluation process can present ethically significant problems.[[12]](#footnote-12)

The presented ethical viewpoints will be investigated more closely in following three subsections: What level of privacy can be achieved through synthetization? Is the synthesized data fair and equitable? Who determines the level of privacy, and under what conditions? These three ethical questions arise in the use of synthetized health data that is generated with generative adversarial networks (GANs) and guaranteed with differential privacy promises. The ethical issues represent critical themes related to synthesized health data and are still unsatisfactory addressed in the research literature. However, it should be noted that synthetization of health data should always be considered ethically on a case-by-case and organizational basis.

**What level of privacy can be achieved through synthetization?**

Differential privacy methods can guarantee a selected level of privacy. Although differential privacy methods generally provide strong privacy protection, the level of privacy chosen is theoretically a trade-off between privacy and the usefulness and accuracy of the data. Complete anonymization of the data material will result in the data being unusable. The choice of level of privacy requires ethical judgment, and, for example, in the United States, U.S. Pat. The Census Bureau has directed an ethical review to the Data Stewardship Executive Policy Committee.[[13]](#footnote-13)

The information theory promises of privacy achieved via differential privacy is many times presented in mathematical form and might be hard for non-technical audience to understand. From an ethical point of view, however, equal attention needs to be paid to how privacy is explained to different target groups: researchers or healthcare professionals using anonymized data, the healthcare decision-making body or the body that administers it, and healthcare clients, i.e., citizens. Some non-technical illustrations are available, see e.g., Wood et al.[[14]](#footnote-14)

The privacy level selection can be opened by introducing the privacy parameter epsilon 𝜖 to be used. Epsilon determines the impact of each individual in the data on the conclusions drawn from the data.[[15]](#footnote-15) A low epsilon reading indicates strong privacy and anonymity, while a high indicates weak protection. On the other hand, a higher epsilon reading in principle indicates better accuracy of synthetic data, i.e., practical usability. The definition of epsilon concretizes the trade-off involved in applying differential privacy. ‘(Another influential parameter in differential privacy should be sensitivity. Roughly speaking, it means the global limit for single data input.[[16]](#footnote-16) The sensitivity of the data affects the selected level of privacy.)

EXAMPLE FROM PRIVASA DATA WILL BE PRESENTED

The choice of epsilon has been also approached in different ways. For example, Hsu et al.[[17]](#footnote-17) presented the idea of ​​an economic approach and Koskela et al.[[18]](#footnote-18) strict approach. However, the different approaches to the choice of privacy level are difficult to assess without in-depth knowledge of the technology. This seems to signal the immaturity of the communicating technology, especially in terms of practical application.

The trade-off between privacy, data accuracy and usability, and law and ethics is a fascinating subject to consider. In general, there seems to be no unambiguous, case-independent answer to this; choosing an epsilon level is a social question.[[19]](#footnote-19) However, ethical assessment of the needed epsilon level can be drawn into ethical calculus where the application’s benefits and risks are weighed. In addition to the prevailing legislation and ethical guidelines, the required level of privacy can be compared to the privacy risks of research and development activities using traditional data methods. Ethical calculus in the field of medical ethics is usually complicated and deserve extensive discussion, e.g., Wolfe[[20]](#footnote-20) and Tcheng[[21]](#footnote-21).

Commented on the theory of ethics, there is a danger that the trade-off calculus associated with synthesized data will be treated with an overemphasized benefit perspective. The more we emphasize the fundamental, perhaps deontological, value of privacy, the greater the level of demand for the benefits of synthesized data. It should be remembered that actions to promote innovation should be based on political compromises, not ethical compromises.

The theoretical trickiness of the privacy-accuracy trade-off could be eased by other ethical goods, like consent. When utilizing differential privacy, a privacy calculus in a comparison to traditional research methods could be an easier task, and with clear consent communication larger data access could follow. In this case, differential privacy would not work only as a privacy-preserving technique but also as a legal and consent design tool.

Another easing limitation to consider is the use of less sensitive health data. Though the categorization of more sensitive and less sensitive is itself an ethical choice, it seems plausible that some borderlines could be drawn. For example, flu statistics and brain imagery have indeed accepted qualitative differences. Also, private companies' medical or semi-medical applications could offer compelling possibilities, especially if the consent process is integrated, for new data to operate further with privacy-preserving methods.

**Is the synthesized data fair and equitable?**

In addition to accuracy, the synthesized data should meet the criteria of sufficient fairness. For example, in addition to making the data set more inaccurate, it may also highlight the properties of some of the subgroups in the data set over other subgroups. However, synthetic material should represent a wide range of subgroups, such as minorities, and should not be overshadowed by majority groups.[[22]](#footnote-22)

EXAMPLE FROM PRIVASA DATA WILL BE PRESENTED

Various solutions have been proposed to minimize bias. Impact assessment processes have become an established approach and various assessment frameworks have been developed to facilitate impact assessment.[[23]](#footnote-23)[[24]](#footnote-24) The impact assessment aims to identify and address the risks that may lead to unequal and unfair outcomes. Various technical approaches and methods can be used to minimize the identified risks, see e.g., Shrestha et al[[25]](#footnote-25).

In the case of health data, unfair outcomes could mean that one group of the population would be able to enjoy the results of research and technological development unfairly compared to other groups of the population. Synthesizing the data may serve as an excellent catalyst to look at these issues when the data properties need to be looked at in any case. However, the issues with fairness make ethical calculus of differential privacy synthetization even more complex, which does not support the expectation of any ready-made solutions.

There is also a trade-off between fairness and privacy. For example, to consider the specific characteristics of a particular population group, these variables need to be included in the data set. Strong anonymization makes it more difficult or even impossible to consider and analyze fairness issues. There is a need for a broad ethical and societal debate on whether there are factors in the healthcare sector that would require more attention to specific sections of the population. Technological solutions have already become so central to our society that their choices significantly impact the distribution of scarce societal resources. It may often be worth considering whether those most in need should be prioritized in their choices[[26]](#footnote-26).

Fairness issues are studied a lot from a technical point of view. The topic was featured already in 2021 by Cynthia Dwork[[27]](#footnote-27), who has long merit in promoting differential privacy, and currently the research material is extensive.[[28]](#footnote-28) From the point of view of ethical assessment, it is hardly too strongly said that techniques and tools can be found to minimize bias but making the right decisions in each application and context remains a tricky part. Combining medical expertise, ethical and social judgment, and technical expertise can be challenging.

The design and implementation of fair technology solutions have been identified as crucial topics of discussion in the ethical debate on new technologies, such as artificial intelligence technologies. The minimum requirement is that technology solutions do not discriminate against anyone or any section of the population. Attention must also be paid to this requirement when synthesizing data. On the other hand, the question of whether there are any distortions in the field of medicine and health care, possibly influenced by history, that research and development work for some sections of the population should be promoted even more than for other sections of the population could be raised as a matter of broader ethical debate. For example, there are openings for discussion about the biasness of genetic data.[[29]](#footnote-29)

**Who determines the level of privacy, and under what conditions?**

When using the synthesized health data, it is essential to consider the chosen level of privacy and ensure the data's fairness. Within both topics, ethical evaluation is required, when choosing how strong the protection of privacy can be guaranteed, how well the subgroups in the data sets can be taken into consideration, and how the compromise between the two, privacy and fairness, can be resolved.

There has been less discussion about who gets to make these choices and under what conditions. The field of medicine and healthcare is accustomed to heavy ethical requirements. The issue of choosing the level of privacy would be appropriate for the to-do list of different ethics committees. A potential problem is the adequacy of technical expertise and the complexity of utility calculations due to complex technical methods. It is unclear how a generalizable ethical perspective and research can be formed when operating with differential privacy methods. There would be a need for clear case studies to calculate the effects of different choices, parameters and outcomes on the root.

The problems caused by the tricky generalizability and the complexity of the methods can be limiting the use of the methods. Indeed, the prevailing practice is not to use differential privacy methods in critical studies that require high and guaranteed data accuracy. The mandatory trade-off between privacy and accuracy directs the use of differential privacy methods, for example, to mapping and screening studies, which can be conducted in addition to traditional method studies.

In the debate on the ethics of artificial intelligence the explainability of different technological solutions has been raised as an essential issue.[[30]](#footnote-30) In the case of differential privacy applications, explainability is a significant burden, with the method being, almost by definition, complex. In turn, applying differential privacy and synthetization could offer an informative effect on the consent process required for the use of health data in research and development.

It would be desirable for the level of accepted residual risk to be determined by either a defined ethics committee or legislation. Although the level of accepted residual risk may vary from case to case, step-by-step limit values would clarify the development of the technology. For example, different stages could consist of categorical applied health technology, medical research, and critically sensitive data (e.g., genetic data).

**Summary**

The generation of synthetic health data with the help of machine learning techniques offers promises to many sectors of medicine and healthcare. Legislation and ethical principles continue to vigorously protect the use of patient data and health data, as planned in the spirit of the law. However, data synthetization and the associated anonymization techniques seem to make it possible to achieve both strong privacy protection and wider data availability for research and development.

A critical ethical guideline for the generation and use of synthetic data is that the ethical implications should be assessed on a case-by-case basis. In addition, the ethical use of synthetic data requires well-established, controlled and continuously evolving data processing and management processes and practices. In the medical and healthcare sectors, ethical research practices are generally solid and deliberate. However, for synthesized data, as in many other technology-intensive areas of medical application, the necessary technical expertise and understanding must be ensured, for example, in ethical committees.

The topic-specific ethical issues described in this article also need to be assessed case-by-case. For example, best practices, ethical calculus structures and recommended limits can be argued concerning the trade-off between privacy and accuracy. Again, however, the actual assessment should be made with the best available medical and case specific knowledge. The key factors to be assessed are the sensitivity of the data used, the benefits of synthetization, and the ability to understandably communicate the data process to stakeholders.

This article focused on three critical ethical issues that arise in the generation of synthetic health data. However, this does not mean that other issues should be closed and refrained from.

A major ethical question is why and when healthcare research should be promoted by generating and using synthetic data. In which situations does synthetic data achieve clear advantages over the original use of the data? Medical research has been carried out for a long time and it has adopted strict privacy practices. The original data also offer the authentic accuracy. The question arises as to whether synthetic data is intended to offer other benefits. The consent process for medical research is often cumbersome and research projects may be tempted to streamline the requirement for consent by easing it. However, a strong borderline must be drawn between the ethical evaluation of consent with synthetic data and easing consent just by means of accelerating technological development.

**References**

Abowd, J. M. (2017). Why the Census Bureau Adopted Differential Privacy for the 2020 Census of Population. *Presentation at Harvard University, December*, *11*.

Adeli, E., Zhao, Q., Pfefferbaum, A., Sullivan, E. V., Fei-Fei, L., Niebles, J. C., & Pohl, K. M. (2021). Representation learning with statistical independence to mitigate bias. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 2513-2523).

Bhanot, K., Qi, M., Erickson, J. S., Guyon, I., & Bennett, K. P. (2021). The problem of fairness in synthetic healthcare data. *Entropy*, *23*(9), 1165.

Calvo, R. A., Peters, D., & Cave, S. (2020). Advancing impact assessment for intelligent systems. *Nature Machine Intelligence*, *2*(2), 89-91.

Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, *5*(6), 493-497.

Corbett-Davies, S., & Goel, S. (2018). The measure and mismeasure of fairness: A critical review of fair machine learning. *arXiv preprint arXiv:1808.00023*.

Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. Found. Trends Theor. Comput. Sci., 9(3-4), 211-407.

Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012, January). Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*(pp. 214-226).

Dwork, C. (2008, April). Differential privacy: A survey of results. In *International conference on theory and applications of models of computation* (pp. 1-19). Springer, Berlin, Heidelberg.

Eibl, G., & Engel, D. (2017). Differential privacy for real smart metering data. *Computer Science-Research and Development*, *32*(1), 173-182.

Fernandez-Quilez, A. (2022). Deep learning in radiology: ethics of data and on the value of algorithm transparency, interpretability and explainability. *AI and Ethics*, 1-9.

Hsu, J., Gaboardi, M., Haeberlen, A., Khanna, S., Narayan, A., Pierce, B. C., & Roth, A. (2014, July). Differential privacy: An economic method for choosing epsilon. In *2014 IEEE 27th Computer Security Foundations Symposium* (pp. 398-410). IEEE.

Islam, K. T., Wijewickrema, S., & O’Leary, S. (2021). A deep learning based framework for the registration of three dimensional multi-modal medical images of the head. *Scientific Reports*, *11*(1), 1-13.

Jordon, J., Yoon, J., & Van Der Schaar, M. (2018, September). PATE-GAN: Generating synthetic data with differential privacy guarantees. In International conference on learning representations.

Jälkö, J., Lagerspetz, E., Haukka, J., Tarkoma, S., Honkela, A., & Kaski, S. (2021). Privacy-preserving data sharing via probabilistic modeling. *Patterns*, *2*(7), 100271.

Kamel Boulos, M. N., Kwan, M. P., El Emam, K., Chung, A. L. L., Gao, S., & Richardson, D. B. (2022). Reconciling public health common good and individual privacy: new methods and issues in geoprivacy. *International Journal of Health Geographics*, *21*(1), 1-9.

Koskela, A., Jälkö, J., Prediger, L., & Honkela, A. (2021, March). Tight differential privacy for discrete-valued mechanisms and for the subsampled gaussian mechanism using fft. In *International Conference on Artificial Intelligence and Statistics* (pp. 3358-3366). PMLR.

Mahmood, F., Borders, D., Chen, R. J., McKay, G. N., Salimian, K. J., Baras, A., & Durr, N. J. (2019). Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE transactions on medical imaging*, *39*(11), 3257-3267.

Morales, J., Welter, D., Bowler, E. H., Cerezo, M., Harris, L. W., McMahon, A. C., ... & MacArthur, J. A. (2018). A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog. *Genome biology*, *19*(1), 1-10.

Nguyen, Q., Das, S., & Garnett, R. (2021, January). Scarce Societal Resource Allocation and the Price of (Local) Justice. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Shrestha, R., Kafle, K., & Kanan, C. (2022). An investigation of critical issues in bias mitigation techniques. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 1943-1954).

Theunissen, M., & Browning, J. (2022). Putting explainable AI in context: institutional explanations for medical AI. *Ethics and Information Technology*, *24*(2), 1-10.

Tayefi, M., Ngo, P., Chomutare, T., Dalianis, H., Salvi, E., Budrionis, A., & Godtliebsen, F. (2021). Challenges and opportunities beyond structured data in analysis of electronic health records. Wiley Interdisciplinary Reviews: Computational Statistics, 13(6), e1549.

Tcheng, J. E., Madan, M., O'SHEA, J. C., Cohen, E. A., Buller, C. E., Lincoff, A. M., ... & Califf, R. M. (2003). Ethics and Equipoise: Rationale for a Placebo‐Controlled Study Design of Platelet Glycoprotein IIb/IIIa Inhibition in Coronary Intervention. *Journal of Interventional Cardiology*, *16*(2), 97-105.

Vu, D., & Slavkovic, A. (2009, December). Differential privacy for clinical trial data: Preliminary evaluations. In *2009 IEEE International Conference on Data Mining Workshops* (pp. 138-143). IEEE.

Wan, C., & Jones, D. T. (2020). Protein function prediction is improved by creating synthetic feature samples with generative adversarial networks. *Nature Machine Intelligence*, *2*(9), 540-550.

Wolfe, D. (2018). Moving the needle: strengthening ethical protections for people who inject drugs in clinical trials. *Journal of Medical Ethics*, *44*(3), 161-162.

Wood, A., Altman, M., Bembenek, A., Bun, M., Gaboardi, M., Honaker, J., ... & Vadhan, S. (2018). Differential privacy: A primer for a non-technical audience. *Vand. J. Ent. & Tech. L.*, *21*, 209.

Ziller, A., Usynin, D., Braren, R., Makowski, M., Rueckert, D., & Kaissis, G. (2021). Medical imaging deep learning with differential privacy. *Scientific Reports*, *11*(1), 1-8.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Chen, R. J., Lu, M. Y., Chen, T. Y., Williamson, D. F., & Mahmood, F. (2021). Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering*, *5*(6), 493-497. [↑](#footnote-ref-1)
2. Islam, K. T., Wijewickrema, S., & O’Leary, S. (2021). A deep learning based framework for the registration of three dimensional multi-modal medical images of the head. *Scientific Reports*, *11*(1), 1-13. [↑](#footnote-ref-2)
3. Tayefi, M., Ngo, P., Chomutare, T., Dalianis, H., Salvi, E., Budrionis, A., & Godtliebsen, F. (2021). Challenges and opportunities beyond structured data in analysis of electronic health records. Wiley Interdisciplinary Reviews: Computational Statistics, 13(6), e1549. [↑](#footnote-ref-3)
4. Kamel Boulos, M. N., Kwan, M. P., El Emam, K., Chung, A. L. L., Gao, S., & Richardson, D. B. (2022). Reconciling public health common good and individual privacy: new methods and issues in geoprivacy. *International Journal of Health Geographics*, *21*(1), 1-9. [↑](#footnote-ref-4)
5. Mahmood, F., Borders, D., Chen, R. J., McKay, G. N., Salimian, K. J., Baras, A., & Durr, N. J. (2019). Deep adversarial training for multi-organ nuclei segmentation in histopathology images. *IEEE transactions on medical imaging*, *39*(11), 3257-3267. [↑](#footnote-ref-5)
6. Wan, C., & Jones, D. T. (2020). Protein function prediction is improved by creating synthetic feature samples with generative adversarial networks. *Nature Machine Intelligence*, *2*(9), 540-550. [↑](#footnote-ref-6)
7. Jordon, J., Yoon, J., & Van Der Schaar, M. (2018, September). PATE-GAN: Generating synthetic data with differential privacy guarantees. In International conference on learning representations. [↑](#footnote-ref-7)
8. Jälkö, J., Lagerspetz, E., Haukka, J., Tarkoma, S., Honkela, A., & Kaski, S. (2021). Privacy-preserving data sharing via probabilistic modeling. *Patterns*, *2*(7), 100271. [↑](#footnote-ref-8)
9. Vu, D., & Slavkovic, A. (2009, December). Differential privacy for clinical trial data: Preliminary evaluations. In *2009 IEEE International Conference on Data Mining Workshops* (pp. 138-143). IEEE. [↑](#footnote-ref-9)
10. Ziller, A., Usynin, D., Braren, R., Makowski, M., Rueckert, D., & Kaissis, G. (2021). Medical imaging deep learning with differential privacy. *Scientific Reports*, *11*(1), 1-8. [↑](#footnote-ref-10)
11. Dwork, C., & Roth, A. (2014). The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, *9*(3-4), 211-407. [↑](#footnote-ref-11)
12. Fernandez-Quilez, A. (2022). Deep learning in radiology: ethics of data and on the value of algorithm transparency, interpretability and explainability. *AI and Ethics*, 1-9. [↑](#footnote-ref-12)
13. Abowd, J. M. (2017). Why the Census Bureau Adopted Differential Privacy for the 2020 Census of Population. *Presentation at Harvard University, December*, *11*. [↑](#footnote-ref-13)
14. Wood, A., Altman, M., Bembenek, A., Bun, M., Gaboardi, M., Honaker, J., ... & Vadhan, S. (2018). Differential privacy: A primer for a non-technical audience. *Vand. J. Ent. & Tech. L.*, *21*, 209. [↑](#footnote-ref-14)
15. ibid. [↑](#footnote-ref-15)
16. Eibl, G., & Engel, D. (2017). Differential privacy for real smart metering data. *Computer Science-Research and Development*, *32*(1), 173-182. [↑](#footnote-ref-16)
17. Hsu, J., Gaboardi, M., Haeberlen, A., Khanna, S., Narayan, A., Pierce, B. C., & Roth, A. (2014, July). Differential privacy: An economic method for choosing epsilon. In *2014 IEEE 27th Computer Security Foundations Symposium* (pp. 398-410). IEEE. [↑](#footnote-ref-17)
18. Koskela, A., Jälkö, J., Prediger, L., & Honkela, A. (2021, March). Tight differential privacy for discrete-valued mechanisms and for the subsampled gaussian mechanism using fft. In *International Conference on Artificial Intelligence and Statistics* (pp. 3358-3366). PMLR. [↑](#footnote-ref-18)
19. Dwork, C. (2008, April). Differential privacy: A survey of results. In *International conference on theory and applications of models of computation* (pp. 1-19). Springer, Berlin, Heidelberg. [↑](#footnote-ref-19)
20. Wolfe, D. (2018). Moving the needle: strengthening ethical protections for people who inject drugs in clinical trials. *Journal of Medical Ethics*, *44*(3), 161-162. [↑](#footnote-ref-20)
21. Tcheng, J. E., Madan, M., O'SHEA, J. C., Cohen, E. A., Buller, C. E., Lincoff, A. M., ... & Califf, R. M. (2003). Ethics and Equipoise: Rationale for a Placebo‐Controlled Study Design of Platelet Glycoprotein IIb/IIIa Inhibition in Coronary Intervention. *Journal of Interventional Cardiology*, *16*(2), 97-105. [↑](#footnote-ref-21)
22. Bhanot, K., Qi, M., Erickson, J. S., Guyon, I., & Bennett, K. P. (2021). The problem of fairness in synthetic healthcare data. *Entropy*, *23*(9), 1165. [↑](#footnote-ref-22)
23. Calvo, R. A., Peters, D., & Cave, S. (2020). Advancing impact assessment for intelligent systems. *Nature Machine Intelligence*, *2*(2), 89-91. [↑](#footnote-ref-23)
24. ECP - Platform for the Information Society Netherlands (2019) AI Impact Assessment <https://airecht.nl/s/Artificial-Intelligence-Impact-Assessment-English.pdf> [↑](#footnote-ref-24)
25. Shrestha, R., Kafle, K., & Kanan, C. (2022). An investigation of critical issues in bias mitigation techniques. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (pp. 1943-1954). [↑](#footnote-ref-25)
26. Nguyen, Q., Das, S., & Garnett, R. (2021, January). Scarce Societal Resource Allocation and the Price of (Local) Justice. In *Proceedings of the AAAI Conference on Artificial Intelligence*. [↑](#footnote-ref-26)
27. Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012, January). Fairness through awareness. In *Proceedings of the 3rd innovations in theoretical computer science conference*(pp. 214-226). [↑](#footnote-ref-27)
28. Corbett-Davies, S., & Goel, S. (2018). The measure and mismeasure of fairness: A critical review of fair machine learning. *arXiv preprint arXiv:1808.00023*. [↑](#footnote-ref-28)
29. Morales, J., Welter, D., Bowler, E. H., Cerezo, M., Harris, L. W., McMahon, A. C., ... & MacArthur, J. A. (2018). A standardized framework for representation of ancestry data in genomics studies, with application to the NHGRI-EBI GWAS Catalog. *Genome biology*, *19*(1), 1-10. [↑](#footnote-ref-29)
30. Theunissen, M., & Browning, J. (2022). Putting explainable AI in context: institutional explanations for medical AI. *Ethics and Information Technology*, *24*(2), 1-10. [↑](#footnote-ref-30)