

A Short Review of Ethical Challenges in Clinical Natural Language Processing

Simon Šuster^{1,2}, Stéphan Tulkens¹ and Walter Daelemans¹

¹CLiPS, University of Antwerp

²Antwerp University Hospital

{simon.suster, stephan.tulkens, walter.daelemans}@uantwerpen.be

Abstract

Clinical NLP has an immense potential in contributing to how clinical practice will be revolutionized by the advent of large scale processing of clinical records. However, this potential has remained largely untapped due to slow progress primarily caused by strict data access policies for researchers. In this paper, we discuss the concern for privacy and the measures it entails. We also suggest sources of less sensitive data. Finally, we draw attention to biases that can compromise the validity of empirical research and lead to socially harmful applications.

1 Introduction

The use of notes written by healthcare providers in the clinical settings has long been recognized to be a source of valuable information for clinical practice and medical research. Access to large quantities of clinical reports may help in identifying causes of diseases, establishing diagnoses, detecting side effects of beneficial treatments, and monitoring clinical outcomes (Agus, 2016; Goldacre, 2014; Murdoch and Detsky, 2013). The goal of clinical natural language processing (NLP) is to develop and apply computational methods for linguistic analysis and extraction of knowledge from free text reports (Demner-Fushman et al., 2009; Hripcsak et al., 1995; Meystre et al., 2008). But while the benefits of clinical NLP and data mining have been universally acknowledged, *progress in the development of clinical NLP techniques has been slow*. Several contributing factors have been identified, most notably difficult access to data, limited collaboration between researchers from different groups, and little sharing of implementations and trained models (Chapman et al., 2011). For

comparison, in biomedical NLP, where the working data consist of biomedical research literature, these conditions have been present to a much lesser degree, and the progress has been more rapid (Cohen and Demner-Fushman, 2014). The main contributing factor to this situation has been the sensitive nature of data, whose processing may in certain situations put patient’s privacy at risk.

The ethics discussion is gaining momentum in general NLP (Hovy and Spruit, 2016). We aim in this paper to gather the ethical challenges that are especially relevant for *clinical NLP*, and to stimulate discussion about those in the broader NLP community. Although enhancing privacy through restricted data access has been the norm, we do not only discuss the right to privacy, but also draw attention to the social impact and biases emanating from clinical notes and their processing. The challenges we describe here are in large part not unique to clinical NLP, and are applicable to general data science as well.

2 Sensitivity of data and privacy

Because of legal and institutional concerns arising from the sensitivity of clinical data, it is difficult for the NLP community to gain access to relevant data (Barzilay, 2016; Friedman et al., 2013). This is especially true for the researchers not connected with a healthcare organization. Corpora with transparent access policies that *are* within reach of NLP researchers exist, but are few. An often used corpus is MIMICII(I) (Johnson et al., 2016; Saeed et al., 2011). Despite its large size (covering over 58,000 hospital admissions), it is only representative of patients from a particular clinical domain (the intensive care in this case) and geographic location (a single hospital in the United States). Assuming that such a specific sample is representative of a larger population is an example of sampling bias

(we discuss further sources of bias in section 3). Increasing the size of a sample without recognizing that this sample is atypical for the general population (e.g. not all patients are critical care patients) could also increase sampling bias (Kaplan et al., 2014).¹ We need more large corpora for various medical specialties, narrative types, as well as languages and geographic areas.

Related to difficult access to raw clinical data is the lack of available *annotated* datasets for model training and benchmarking. The reality is that annotation projects do take place, but are typically constrained to a single healthcare organization. Therefore, much of the effort put into annotation is lost afterwards due to impossibility of sharing with the larger research community (Chapman et al., 2011; Fan et al., 2011). Again, exceptions are either few—e.g. THYME (Styler IV et al., 2014), a corpus annotated with temporal information—or consist of small datasets resulting from shared tasks like the i2b2 and ShARe/CLEF. In addition, stringent access policies hamper reproduction efforts, impede scientific oversight and limit collaboration, not only between institutions but also more broadly between the clinical and NLP communities.

There are known cases of datasets that had been used in published research (including reproduction) in its full form, like MiPACQ², Blulab, EMC Dutch Clinical Corpus and 2010 i2b2/VA (Albright et al., 2013; Kim et al., 2015; Afzal et al., 2014; Uzuner et al., 2011), but were later trimmed down or made unavailable, likely due to legal issues. Even if these datasets were still available in full, their small size is still a concern, and the comments above regarding sampling bias certainly apply. For example, a named entity recognizer trained on 2010 i2b2/VA data, which consists of 841 annotated patient records from three different specialty areas, will due to its size only contain a small portion of possible named entities. Similarly, in linking clinical concepts to an ontology, where the number of output classes is larger (Pradhan et al., 2013), the small amount of training data is a major obstacle to deployment of systems suitable for general use.

¹Sampling bias could also be called selection bias; it is not inherent to the individual documents, but stems from the way these are arranged into a single corpus.

²The access to the MiPACQ corpus will be re-enabled in the future within the Health NLP Center for distributing linguistic annotations of clinical texts (Guergana Savova, personal communication).

2.1 Protecting the individual

Clinical notes contain detailed information about patient-clinician encounters in which patients confide not only their health complaints, but also their lifestyle choices and possibly stigmatizing conditions. This confidential relationship is legally protected in US by the HIPAA privacy rule in the case of individuals' medical data. In EU, the conditions for scientific usage of health data are set out in the General Data Protection Regulation (GDPR). Sanitization of sensitive data categories and individuals' informed consent are in the forefront of those legislative acts and bear immediate consequences for the NLP research.

The GDPR lists general principles relating to processing of personal data, including that processing must be lawful (e.g. by means of consent), fair and transparent; it must be done for explicit and legitimate purposes; and the data should be kept limited to what is necessary and as long as necessary. This is known as data minimization, and it includes sanitization. The scientific usage of health data concerns "special categories of personal data". Their processing is only allowed when the data subject gives explicit consent, or the personal data is made public by the data subject. Scientific usage is defined broadly and includes technological development, fundamental and applied research, as well as privately funded research.

Sanitization Sanitization techniques are often seen as the minimum requirement for protecting individuals' privacy when collecting data (Berman, 2002; Velupillai et al., 2015). The goal is to apply a procedure that produces a new version of the dataset that looks like the original for the purposes of data analysis, but which maintains the privacy of those in the dataset to a certain degree, depending on the technique. Documents can be sanitized by replacing, removing or otherwise manipulating the sensitive mentions such as names and geographic locations. A distinction is normally drawn between anonymization, pseudonymization and de-identification. We refer the reader to Polonetsky et al. (2016) for an excellent overview of these procedures.

Although it is a necessary first step in protecting the privacy of patients, sanitization has been criticized for several reasons. First, it affects the integrity of the data, and as a consequence, their utility (Duquenoy et al., 2008). Second, although sanitization in principle promotes data access and

sharing, it may often not be sufficient to eliminate the need for consent. This is largely due to the well-known fact that original sensitive data can be re-identified through deductive disclosure (Amblard et al., 2014; De Mazancourt et al., 2015; Hardt et al., 2016; Malin et al., 2013; Tene, 2011).³ Finally, sanitization focuses on protecting the individual, whereas ethical harms are still possible on the group level (O'Doherty et al., 2016; Taylor et al., 2017). Instead of working towards increasingly restrictive sanitization and access measures, another course of action could be to work towards heightening the perception of scientific work, emphasizing professionalism and existence of punitive measures for illegal actions (Fairfield and Shtein, 2014; Mittelstadt and Floridi, 2016).

Consent Clinical NLP typically requires a large amount of clinical records describing cases of patients with a particular condition. Although obtaining consent is a necessary first step, obtaining explicit informed consent from each patient can also compromise the research in several ways. First, obtaining consent is time consuming by itself, and it results in financial and bureaucratic burdens. It can also be infeasible due to practical reasons such as a patient's death. Next, it can introduce bias as those willing to grant consent represent a skewed population (Nyrén et al., 2014). Finally, it can be difficult to satisfy the informedness criterion: Information about the experiment sometimes can not be communicated in an unambiguous way, or experiments happen at speed that makes enacting informed consent extremely hard (Bird et al., 2016).

The alternative might be a default opt-in policy with a right to withdraw (opt-out). Here, consent can be presumed either in a broad manner—allowing unspecified future research, subject to ethical restrictions—or a tiered manner—allowing certain areas of research but not others (Mittelstadt and Floridi, 2016; Terry, 2012). Since the information about the intended use is no longer uniquely tied to each research case but is more general, this could facilitate the reuse of datasets by several research teams, without the need to ask for consent each time. The success of implementing this approach in practice is likely to depend on public trust and awareness about possible risks and oppor-

³ Additionally, it may be due to organizational skepticism about the effectiveness of sanitization techniques, although it has been shown that automated de-identification systems for English perform on par with manual de-identification (Deleger et al., 2013).

tunities. We also believe that a distinction between academic research and commercial use of clinical data should be implemented, as the public is more willing to allow research than commercial exploitation (Lawrence, 2016; van Staa et al., 2016).

Yet another possibility is open consent, in which individuals make their data publicly available. Initiatives like Personal Genome Project may have an exemplary role, however, they can only provide limited data and they represent a biased population sample (Mittelstadt and Floridi, 2016).

Secure access Since withholding data from researchers would be a dubious way of ensuring confidentiality (Berman, 2002), the research has long been active on secure access and storage of sensitive clinical data, and the balance between the degree of privacy loss and the degree of utility. This is a broad topic that is outside the scope of this article. The interested reader can find the relevant information in Dwork and Pottenger (2013), Malin et al. (2013) and Rindfleisch (1997).

Promotion of knowledge and application of best-of-class approaches to health data is seen as one of the ethical duties of researchers (Duquenoy et al., 2008; Lawrence, 2016). But for this to be put in practice, ways need to be guaranteed (e.g. with government help) to provide researchers with access to the relevant data. Researchers can also go to the data rather than have the data sent to them. It is an open question though whether medical institutions—especially those with less developed research departments—can provide the infrastructure (e.g. enough CPU and GPU power) needed in statistical NLP. Also, granting access to one healthcare organization at a time does not satisfy interoperability (cross-organizational data sharing and research), which can reduce bias by allowing for more complete input data. Interoperability is crucial for epidemiology and rare disease research, where data from one institution can not yield sufficient statistical power (Kaplan et al., 2014).

Are there less sensitive data? One criterion which may have influence on data accessibility is whether the data is about living subjects or not. The HIPAA privacy rule under certain conditions allows disclosure of personal health information of **deceased persons**, without the need to seek IRB agreement and without the need for sanitization (Huser and Cimino, 2014). It is not entirely clear

though how often this possibility has been used in clinical NLP research or broader.

Next, the work on **surrogate data** has recently seen a surge in activity. Increasingly more health-related texts are produced in social media (Abbasi et al., 2014), and patient-generated data are available online. Admittedly, these may not resemble the clinical discourse, yet they bear to the same individuals whose health is documented in the clinical reports. Indeed, linking individuals' health information from online resources to their health records to improve documentation is an active line of research (Padrez et al., 2015). Although it is generally easier to obtain access to social media data, the use of social media still requires similar ethical considerations as in the clinical domain. See for example the influential study on emotional contagion in Facebook posts by Kramer et al. (2014), which has been criticized for not properly gaining prior consent from the users who were involved in the study (Schroeder, 2014).

Another way of reducing sensitivity of data and improving chances for IRB approval is to work on **derived data**. Data that can not be used to reconstruct the original text (and when sanitized, can not directly re-identify the individual) include text fragments, various statistics and trained models. Working on randomized subsets of clinical notes may also improve the chances of obtaining the data. When we only have access to trained models from disparate sources, we can refine them through ensembling and creation of silver standard corpora, cf. Reibholz-Schuhmann et al. (2011).

Finally, clinical NLP is also possible on **veterinary texts**. Records of companion animals are perhaps less likely to involve legal issues, while still amounting to a large pool of data. As an example, around 40M clinical documents from different veterinary clinics in UK and Australia are stored centrally in the VetCompass repository. First NLP steps in this direction were described in the invited talk at the Clinical NLP 2016 workshop (Baldwin, 2016).

3 Social impact and biases

Unlocking knowledge from free text in the health domain has a tremendous societal value. However, discrimination can occur when individuals or groups receive unfair treatment as a result of automated processing, which might be a result of biases in the data that were used to train models.

The question is therefore what the most important biases are and how to overcome them, not only out of ethical but also legal responsibility. Related to the question of bias is so-called *algorithm transparency* (Goodman, 2016; Kamarinou et al., 2016), as this right to explanation requires that influences of bias in training data are charted. In addition to sampling bias, which we introduced in section 2, we discuss in this section further sources of bias. Unlike sampling bias, which is a corpus-level bias, these biases here are already present in documents, and therefore hard to account for by introducing larger corpora.

Data quality Texts produced in the clinical settings do not always tell a complete or accurate patient story (e.g. due to time constraints or due to patient treatment in different hospitals), yet important decisions can be based on them.⁴ As language is situated, a lot of information may be implicit, such as the circumstances in which treatment decisions are made (Hersh et al., 2013). If we fail to detect a medical concept during automated processing, this can not necessarily be a sign of negative evidence.⁵ Work on identifying and imputing missing values holds promise for reducing incompleteness, see Lipton et al. (2016) for an example in sequential modeling applied to diagnosis classification.

Reporting bias Clinical texts may include bias coming from both patient's and clinician's reporting. Clinicians apply their subjective judgments to what is important during the encounter with patients. In other words, there is separation between, on the one side, what is observed by the clinician and communicated by the patient, and on the other, what is noted down. Cases of more serious illness may be more accurately documented as a result of clinician's bias (increased attention) and patient's recall bias. On the other hand, the cases of stigmatized diseases may include suppressed information. In the case of traffic injuries, documentation may even be distorted to avoid legal consequences (Indrayan, 2013).

⁴A way to increase data completeness and reduce selection bias is the use of nationwide patient registries, as known for example in Scandinavian countries (Schmidt et al., 2015).

⁵We can take timing-related "censoring" effects as an example. In event detection, events prior to the start of an observation may be missed or are uncertain, which means that the first appearance of a diagnosis in the clinical record may not coincide with the occurrence of the disease. Similarly, key events after the end of the observation may be missing (e.g. death, when it occurred in another institution).

We need to be aware that clinical notes may reflect health disparities. These can originate from prejudices held by healthcare practitioners which may impact patients' perceptions; they can also originate from communication difficulties in the case of ethnic differences (Zestcott et al., 2016). Finally, societal norms can play a role. Brady et al. (2016) find that obesity is often not documented equally well for both sexes in weight-addressing clinics. Young males are less likely to be recognized as obese, possibly due to societal norms seeing them as "stocky" as opposed to obese. Unless we are aware of such bias, we may draw premature conclusions about the impact of our results.

It is clear that during processing of clinical texts, we should strive to avoid reinforcing the biases. It is difficult to give a solution on how to actually reduce the reporting bias after the fact. One possibility might be to model it. If we see clinical reports as noisy annotations for the patient story in which information is left-out or altered, we could try to decouple the bias from the reports. Inspiration could be drawn, for example, from the work on decoupling reporting bias from annotations in visual concept recognition (Misra et al., 2016).

Observational bias Although variance in health outcome is affected by social, environmental and behavioral factors, these are rarely noted in clinical reports (Kaplan et al., 2014). The bias of missing explanatory factors because they can not be identified within the given experimental setting is also known as *the streetlight effect*. In certain cases, we could obtain important prior knowledge (e.g. demographic characteristics) from data other than clinical notes.

Dual use We have already mentioned linking personal health information from online texts to clinical records as a motivation for exploring surrogate data sources. However, this and many other applications also have potential to be applied in both beneficial and harmful ways. It is easy to imagine how sensitive information from clinical notes can be revealed about an individual who is present in social media with a known identity. More general examples of dual use are when the NLP tools are used to analyze clinical notes with a goal of determining individuals' insurability and employability.

4 Conclusion

In this paper, we reviewed some challenges that we believe are central to the work in clinical NLP.

Difficult access to data due to privacy concerns has been an obstacle to progress in the field. We have discussed how the protection of privacy through sanitization measures and the requirement for informed consent may affect the work in this domain. Perhaps, it is time to rethink the right to privacy in health in the light of recent work in ethics of big data, especially its uneasy relationship to the *right to science*, i.e. being able to benefit from science and participate in it (Tasioulas, 2016; Verbeek, 2014). We also touched upon possible sources of bias that can have an effect on the application of NLP in the health domain, and which can ultimately lead to unfair or harmful treatment.

Acknowledgments

We would like to thank Madhumita and the anonymous reviewers for useful comments. Part of this research was carried out in the framework of the Accumulate IWT SBO project, funded by the government agency for Innovation by Science and Technology (IWT).

References

- Ahmed Abbasi, Donald Adjeroh, Mark Dredze, Michael J. Paul, Fatemeh Mariam Zahedi, Huimin Zhao, Nitin Walia, Hemant Jain, Patrick Sanvanson, Reza Shaker, et al. 2014. Social media analytics for smart health. *IEEE Intelligent Systems*, 29(2):60–80.
- Zubair Afzal, Ewoud Pons, Ning Kang, Miriam C.J.M. Sturkenboom, Martijn J. Schuemie, and Jan A. Kors. 2014. ContextD: an algorithm to identify contextual properties of medical terms in a Dutch clinical corpus. *BMC Bioinformatics*, 15(1):373.
- David B. Agus. 2016. Give Up Your Data to Cure Disease, The New York Times, February 6. <https://goo.gl/0REG0n>.
- D. Albright, A. Lanfranchi, A. Fredriksen, W. F. Styler, C. Warner, J. D. Hwang, J. D. Choi, D. Dligach, R. D. Nielsen, J. Martin, W. Ward, M. Palmer, and G. K. Savova. 2013. Towards comprehensive syntactic and semantic annotations of the clinical narrative. *Journal of the American Medical Informatics Association*, 20(5):922–930.
- Maxime Amblard, Karèn Fort, Michel Musiol, and Manuel Rebuschi. 2014. L'impossibilité de l'anonymat dans le cadre de l'analyse du discours. In *Journée ATALA éthique et TAL*.
- Timothy Baldwin. 2016. VetCompass: Clinical Natural Language Processing for Animal Health. Clinical NLP 2016 keynote. <https://goo.gl/ScGFa2>.

- Regina Barzilay. 2016. How NLP can help cure cancer? NAACL'16 keynote. <https://goo.gl/hi5nrq>.
- Jules J. Berman. 2002. Confidentiality issues for medical data miners. *Artificial Intelligence in Medicine*, 26(1):25–36.
- Sarah Bird, Solon Barocas, Kate Crawford, Fernando Diaz, and Hanna Wallach. 2016. Exploring or Exploiting? Social and Ethical Implications of Autonomous Experimentation in AI. In *Workshop on Fairness, Accountability, and Transparency in Machine Learning*.
- Cassandra C. Brady, Vidhu V. Thaker, Todd Lingren, Jessica G. Woo, Stephanie S. Kennebeck, Bahram Namjou-Khales, Ashton Roach, Jonathan P. Bickel, Nandan Patibandla, Guergana K. Savova, et al. 2016. Suboptimal Clinical Documentation in Young Children with Severe Obesity at Tertiary Care Centers. *International Journal of Pediatrics*, 2016.
- Wendy W. Chapman, Prakash M. Nadkarni, Lynette Hirschman, Leonard W. D'Avolio, Guergana K. Savova, and Özlem Uzuner. 2011. Overcoming barriers to NLP for clinical text: the role of shared tasks and the need for additional creative solutions. *Journal of the American Medical Informatics Association*, 18(5):540–543.
- Kevin Bretonnel Cohen and Dina Demner-Fushman. 2014. *Biomedical natural language processing*. John Benjamins Publishing Company.
- Hugues De Mazancourt, Alain Couillaud, Gilles Adda, and Gaëlle Recourcé. 2015. Faire du TAL sur des données personnelles : un oxymore ? In *TALN 2015*.
- Louise Deleger, Katalin Molnar, Guergana Savova, Fei Xia, Todd Lingren, Qi Li, Keith Marsolo, Anil Jegga, Megan Kaiser, Laura Stoutenborough, and Imre Solti. 2013. Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. *Journal of the American Medical Informatics Association*, 20(1):84.
- Dina Demner-Fushman, Wendy W. Chapman, and Clement J. McDonald. 2009. What can natural language processing do for clinical decision support? *Journal of Biomedical Informatics*, 42(5):760–772.
- Penny Duquenoy, Carlisle George, and Anthony Solomonides. 2008. Considering something ELSE: Ethical, legal and socio-economic factors in medical imaging and medical informatics. *Computer Methods and Programs in Biomedicine*, 92(3):227–237.
- Cynthia Dwork and Rebecca Pottenger. 2013. Toward practicing privacy. *Journal of the American Medical Informatics Association*, 20(1):102–108.
- Joshua Fairfield and Hannah Shtein. 2014. Big data, big problems: Emerging issues in the ethics of data science and journalism. *Journal of Mass Media Ethics*, 29(1):38–51.
- Jung-wei Fan, Rashmi Prasad, Romme M. Yabut, Richard M. Loomis, Daniel S. Zisook, John E. Mattison, and Yang Huang. 2011. Part-of-speech tagging for clinical text: wall or bridge between institutions. In *AMIA Annual Symposium Proceedings*.
- Carol Friedman, Thomas C. Rindfleisch, and Milton Corn. 2013. Natural language processing: state of the art and prospects for significant progress, a workshop sponsored by the National Library of Medicine. *Journal of Biomedical Informatics*, 46(5):765–773.
- Ben Goldacre. 2014. The NHS plan to share our medical data can save lives but must be done right. <https://goo.gl/MH2eC0>.
- Bryce W. Goodman. 2016. A Step Towards Accountable Algorithms?: Algorithmic Discrimination and the European Union General Data Protection. In *NIPS Symposium on Machine Learning and the Law*.
- Moritz Hardt, Eric Price, and Nati Srebro. 2016. Equality of opportunity in supervised learning. In *NIPS*.
- William R. Hersh, Mark G. Weiner, Peter J. Embi, Judith R. Logan, Philip R.O. Payne, Elmer V. Bernstein, Harold P. Lehmann, George Hripcsak, Timothy H. Hartzog, James J. Cimino, and Joel H. Saltz. 2013. Caveats for the use of operational electronic health record data in comparative effectiveness research. *Medical care*, 51(8 0 3).
- Dirk Hovy and Shannon L. Spruit. 2016. The social impact of natural language processing. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 591–598, Berlin, Germany, August. Association for Computational Linguistics.
- George Hripcsak, Carol Friedman, Philip O. Alderson, William DuMouchel, Stephen B. Johnson, and Paul D. Clayton. 1995. Unlocking clinical data from narrative reports: a study of natural language processing. *Annals of internal medicine*, 122(9):681–688.
- Vojtech Huser and James J. Cimino. 2014. Don't take your EHR to heaven, donate it to science: legal and research policies for EHR post mortem. *Journal of the American Medical Informatics Association*, 21(1):8–12.
- Abhaya Indrayan. 2013. Varieties of bias to guard against. <https://goo.gl/SqnuZY>.
- Alistair EW Johnson, Tom J. Pollard, Lu Shen, Liwei H. Lehman, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G. Mark. 2016. MIMIC-III, a freely accessible critical care database. *Scientific data*, 3.
- Dimitra Kamarinou, Christopher Millard, and Jatinder Singh. 2016. Machine Learning with Personal Data. *Queen Mary School of Law Legal Studies Research Paper*, (247/2016).

- Robert M. Kaplan, David A. Chambers, and Russell E. Glasgow. 2014. Big data and large sample size: a cautionary note on the potential for bias. *Clinical and translational science*, 7(4):342–346.
- Youngjun Kim, Ellen Riloff, and John F. Hurdle. 2015. A study of concept extraction across different types of clinical notes. In *AMIA Annual Symposium Proceedings*, volume 2015, page 737. American Medical Informatics Association.
- Adam D.I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. 2014. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790.
- Neil Lawrence. 2016. Data Analysis, NHS and Industrial Partners. <https://goo.gl/rRIcu5>.
- Zachary C. Lipton, David Kale, and Randall Wetzel. 2016. Modeling Missing Data in Clinical Time Series with RNNs. *arXiv preprint arXiv:1606.04130*.
- Bradley A. Malin, Khaled El Emam, and Christine M. O’Keefe. 2013. Biomedical data privacy: problems, perspectives, and recent advances. *Journal of the American Medical Informatics Association*, 20(1):2–6.
- Stéphane M. Meystre, Guergana K. Savova, Karin C. Kipper-Schuler, John F. Hurdle, et al. 2008. Extracting information from textual documents in the electronic health record: a review of recent research. *Yearbook of Medical Informatics*, 35:128–44.
- Ishan Misra, C. Lawrence Zitnick, Margaret Mitchell, and Ross B. Girshick. 2016. Seeing through the Human Reporting Bias: Visual Classifiers from Noisy Human-Centric Labels. In *The IEEE Conference on Computer Vision and Pattern Recognition*.
- Brent Daniel Mittelstadt and Luciano Floridi. 2016. The ethics of big data: current and foreseeable issues in biomedical contexts. *Science and engineering ethics*, 22(2):303–341.
- Travis B. Murdoch and Allan S. Detsky. 2013. The inevitable application of big data to health care. *JAMA*, 309(13):1351–1352.
- Olof Nyrén, Magnus Stenbeck, and Henrik Grönberg. 2014. The European Parliament proposal for the new EU General Data Protection Regulation may severely restrict European epidemiological research. *European Journal of Epidemiology*, 29(4):227–230.
- Kieran C. O’Doherty, Emily Christofides, Jeffery Yen, Heidi Beate Bentzen, Wylie Burke, Nina Hallowell, Barbara A. Koenig, and Donald J. Willison. 2016. If you build it, they will come: unintended future uses of organised health data collections. *BMC Medical Ethics*, 17(1).
- Kevin A. Padrez, Lyle Ungar, Hansen Andrew Schwartz, Robert J. Smith, Shawndra Hill, Tadas Antanavicius, Dana M. Brown, Patrick Crutchley, David A. Asch, and Raina M. Merchant. 2015. Linking social media and medical record data: a study of adults presenting to an academic, urban emergency department. *BMJ quality & safety*.
- Jules Polonetsky, Omer Tene, and Kelsey Finch. 2016. Shades of Gray: Seeing the Full Spectrum of Practical Data De-identification. *Santa Clara Law Review*, 56(3).
- Sameer Pradhan, Noemie Elhadad, Brett R. South, David Martinez, Amy Vogel, Hanna Suominen, Wendy W. Chapman, and Guergana Savova. 2013. Task 1: Share/clef ehealth evaluation lab. In *Online Working Notes of CLEF*.
- Dietrich Reibholz-Schuhmann, Antonio Jimeno Yepes, Chen Li, et al. 2011. Assessment of NER solutions against the first and second CALBC Silver Standard Corpus. *Journal of Biomedical Semantics*, 2(5):S11.
- Thomas C. Rindfleisch. 1997. Privacy, information technology, and health care. *Communications of the ACM*, 40(8):92–100.
- Mohammed Saeed, Mauricio Villarroel, Andrew T. Reisner, Gari Clifford, Li-Wei Lehman, George Moody, Thomas Heldt, Tin H. Kyaw, Benjamin Moody, and Roger G. Mark. 2011. Multiparameter Intelligent Monitoring in Intensive Care II (MIMIC-II): a public-access intensive care unit database. *Critical care medicine*, 39(5):952.
- Morten Schmidt, S.A. Schmidt, Jakob Lynge Sandegaard, Vera Ehrenstein, Lars Pedersen, and Henrik Toft Sørensen. 2015. The danish national patient registry: a review of content, data quality, and research potential. *Clinical Epidemiology*, 7(449):e490.
- Ralph Schroeder. 2014. Big data and the brave new world of social media research. *Big Data & Society*, 1(2):2053951714563194.
- William Styler IV, Steven Bethard, Sean Finan, Martha Palmer, Sameer Pradhan, Piet de Groen, Brad Erickson, Timothy Miller, Chen Lin, Guergana Savova, and James Pustejovsky. 2014. Temporal annotation in the clinical domain. *Transactions of the Association for Computational Linguistics*, 2:143–154.
- John Tasioulas. 2016. Van Hasselt Lecture 2016: Big Data, Human Rights and the Ethics of Scientific Research. <https://goo.gl/QREHUN>.
- Linnet Taylor, Luciano Floridi, and Bart van der Sloot. 2017. *Group Privacy: New Challenges of Data Technologies*. Springer International Publishing.
- Omer Tene. 2011. Privacy: The new generations. *International Data Privacy Law*, 1(1):15–27.

- Nicolas P. Terry. 2012. Protecting patient privacy in the age of big data. *University of Missouri-Kansas City Law Review*, 81(2).
- Özlem Uzuner, Brett R. South, Shuying Shen, and Scott L. DuVall. 2011. 2010 i2b2/va challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association*, 18(5):552–556.
- Tjeerd-Pieter van Staa, Ben Goldacre, Iain Buchan, and Liam Smeeth. 2016. Big health data: the need to earn public trust. *BMJ*, 354:i3636.
- Sumithra Velupillai, D. Mowery, Brett R. South, Maria Kvist, and Hercules Dalianis. 2015. Recent advances in clinical natural language processing in support of semantic analysis. *Yearbook of Medical Informatics*, 10:183–193.
- Peter-Paul Verbeek. 2014. *Op de vleugels van Icarus*. Lemniscaat.
- Colin A. Zestcott, Irene V. Blair, and Jeff Stone. 2016. Examining the presence, consequences, and reduction of implicit bias in health care: A narrative review. *Group Processes & Intergroup Relations*.