Natural Language Processing in Healthcare: A Bird's Eye View

Luca Bacco[1,2,*], Felice Dell'Orletta[2] and Mario Merone[1]

[1]*Department of Engineering, Unit of Computer Systems and Bioinformatics, Campus Bio-Medico University of Rome, Via Alvaro del Portillo, 21, 00128, Rome, Italy*

[2]*ItaliaNLP Lab, National Research Council, Istituto di Linguistica Computazionale "Antonio Zampolli", Via Giuseppe Moruzzi, 1, 56124, Pisa, Italy*

Abstract

The healthcare industry is experiencing an unprecedented era of transformation, driven by the proliferation of Electronic Health Records (EHRs) and the emergence of vast amounts of natural language data from sources like social media. This dynamic landscape presents novel opportunities for companies, healthcare practitioners, and researchers, underpinned by the transformative potential of Natural Language Processing (NLP) technologies. This paper offers a brief overview of the rationale for studying NLP in healthcare, highlighting its foundational importance. It explores the current trends and invaluable resources available in the field, while also delving into the multifaceted challenges that must be addressed to harness the full potential of NLP in healthcare. By consolidating essential insights and bridging the gap between opportunities and challenges, this manuscript serves as a valuable resource for both established researchers and newcomers seeking to navigate the complex terrain of NLP in healthcare, ultimately contributing to the advancement of this critical domain.

Keywords

Natural Language Processing, Healthcare, Review, Electronic Health Records

1. Introduction

When we mention Artificial Intelligence (AI) in healthcare, we often imagine systems adept at interpreting sensor data — such as time series from physiological measurements and medical images — to provide invaluable diagnostic support to healthcare professionals and patients alike. However, a substantial portion of patient information moves through the healthcare ecosystem in the form of natural language narratives [2]. These textual data hold immense potential (and need) for automated processing and analysis.

The past few decades have witnessed a profound digital transformation across industries, including healthcare. This transformation has unleashed a torrent of data from diverse sources alongside new opportunities and challenges. The opportunities offered by Natural Language Processing (NLP) in healthcare are vast and intricate, rendering a comprehensive review im-

NL4AI 2023: Seventh Workshop on Natural Language for Artificial Intelligence, November 6-7th, 2023, Rome, Italy [1]

*Corresponding author.

l.bacco@unicampus.it (L. Bacco); felice.dellorletta@ilc.cnr.it (F. Dell'Orletta); m.merone@unicampus.it (M. Merone)

0000-0001-5462-2727 (L. Bacco); 0000-0003-3454-9387 (F. Dell'Orletta); 0000-0002-9406-2397 (M. Merone)

practical. For newcomers venturing into this field, navigating this complex landscape can be overwhelming.

In recent years, several researchers have undertaken the task of dissecting this field from various angles. Some have focused on the learning techniques employed in the literature, such as machine and deep learning [3, 4]. Others have focused on specific data sources, either electronic health records from healthcare institutions or unstructured data from social media platforms [5, 6, 7]. Among those works, Gao et al. [8] have compiled a valuable resource by cataloging publicly available datasets in clinical NLP, which researchers are exploiting to overcome privacy concerns (see Section 4). Nonetheless, most of these investigations have revolved around specialized subdomains within healthcare and clinical practices, such as mental health [9], orthopedics [10], and dentistry [11].

Thus, this paper aims to provide a concise overview of the panorama of NLP in healthcare, offering condensed insights into data sources, tools, current trends, and persistent challenges, providing a guiding beacon, especially for NLP researchers and practitioners embarking on the healthcare field.

2. Data sources

Electronic Health Records (EHRs) The healthcare industry is in the midst of a rapid digital transformation, marked by the widespread adoption of Electronic Health Records (EHRs), also known as Electronic Medical Records (EMRs), in hospitals and clinics worldwide [12, 13]. These clinical documents often contain approximately 80% of unstructured data [14], primarily as free text. In clinical practice, healthcare professionals exchange patient information as narrative notes and reports [2].

Besides offering a more comprehensive description of a patient's health status compared to quantitative data alone, these natural language reports can be time-consuming to understand and redact. Not surprisingly, the development of NLP in the healthcare sector has paralleled the increasing adoption of EHRs. NLP tools have emerged as essential solutions for efficiently managing unstructured data within these records. Automatically analyzing patients' records not only facilitates access to a wealth of information but also serves the needs of physicians, healthcare companies, patients, and researchers.

The World Wide Web (WWW) In addition to EHRs, the proliferation of the World Wide Web (WWW) has opened up new opportunities for the NLP community to harness data from sources like forums and social media to enhance healthcare services [5]. Given the vast volume of text-based information exchanged daily on the WWW, this digital landscape presents an unprecedented opportunity for NLP researchers and practitioners. It serves as an invaluable source of textual data for training NLP systems, often through web scraping technologies [15]. Beyond general-purpose social media platforms, the WWW has spawned specialized online health communities such as PatientsLikeMe[1] and DailyStrength[2]. These platforms bring together users with similar health-related interests and concerns. Other platforms on the Internet consist

[1]https://www.patientslikeme.com/
[2]https://www.dailystrength.org/

of websites for spreading medical information. NLP researchers are already making use of resources like Wikipedia and SimpleWikipedia[3], and the MSD Manuals[4] [16].

3. Tasks

Within the vast panorama of NLP applied to healthcare, various tasks exist, each offering benefits to multiple stakeholders. Although the division proposed here is not an absolute categorization (i.e., some of the solutions may benefit more stakeholders), it serves as a handy framework for understanding the diverse applications of NLP in this domain.

Tasks for physicians NLP tasks designed for physicians predominantly revolve around the generation/extraction of critical information from EHRs. These tasks serve as tools for lightening the workload of healthcare professionals and enhancing decision support for physicians [6]. For instance, the development of Natural Language Generation (NLG) systems proves invaluable in creating clinically precise structured reports from various clinical documents or condensing the extensive notes generated daily throughout the care process [17]. Conversely, Natural Language Understanding (NLU) systems function as Information Extraction tools, directly harvesting pertinent data from these reports, encompassing diagnoses, medication records, treatment plans, and other essential entities. Through these functionalities, NLP systems show the potential to empower clinicians with evidence-based recommendations and timely alerts, elevating the caliber of medical decision-making. Simultaneously, they alleviate the burdens associated with manual data entry, enhancing efficiency and accuracy in healthcare practice.

Tasks for institutions Automated analysis of text data within EHRs holds tremendous significance in enhancing the management of healthcare services provided by organizations. For instance, healthcare institutions can harness NLP to predict the risk of patients' readmissions. This predictive capability facilitates the implementation of targeted interventions, optimizes resource allocation, reduces hospital readmission rates, and ultimately lowers costs [18].

Another valuable application of text analysis in EHRs involves streamlining administrative tasks for healthcare facilities. It is achieved by automating the coding activities involved in billing processes. Diagnoses and procedures are usually encoded using standardized nomenclatures such as the International Classification of Diseases (ICD), overseen by the World Health Organization (WHO). These codes simplify health data comparison across populations and support epidemiological analyses. Currently, the responsibility for this classification process often falls on trained or even untrained personnel within healthcare institutions. Given the task complexity and time-intensive nature, it poses challenges even for professionally trained staff and is prone to errors. These errors can result in financial losses and potential legal consequences. Consequently, there is a growing interest in developing automatic systems for ICD coding, as evidenced by various research efforts [19] focused on clinical note extraction, such as discharge summaries.

[3]https://en.wikipedia.org/wiki/Main_Page, https://simple.wikipedia.org/wiki/Main_Page
[4]https://www.msdmanuals.com/

Beyond EHRs, the World Wide Web (WWW) is a crucial resource for healthcare institutions. NLP-driven surveillance systems can actively monitor digital sources, enabling the early detection of disease outbreaks, emerging trends, and even signs of suicidal intentions [20, 21], allowing timely and proactive responses.

Moreover, just as consumers seek product or service reviews before making decisions, patients actively seek and share health-related experiences online [22, 23]. Healthcare companies can derive significant benefits from automating the analysis of patients' opinions to identify strengths and weaknesses in their services and treatments.Automated tools offer the advantage of processing a vast number of reviews, eliminating the need for traditional structured surveys and questionnaires that limit patient expressiveness and are costly and time-consuming to design and analyze. This type of analysis is commonly known as sentiment analysis or opinion mining, a task that has garnered extensive attention in various domains within the NLP community, yet remains relatively underexplored in healthcare [24].

Tasks for patients The World Wide Web (WWW) empowers patients to access a wealth of information, one of the Internet's foremost virtues being its democratic provision of easy access to vast knowledge resources. However, the abundance of medical information available online can pose challenges for patients without medical backgrounds, who may misinterpret these resources [25], potentially leading to harmful consequences. In this context, NLP is emerging as a valuable tool for simplifying health-related texts, bridging the expertise gap for patients [16, 26]. However, the Internet, particularly on social media and similar platforms, can be disseminated with medical misinformation. Consequently, several researchers have recently dedicated their efforts to detecting health-related fake news and combatting what has been termed the *infodemic* [27, 28].

Another popular application benefiting patients is chatbots, digital agents designed for interactive conversations with users. These chatbots offer a friendly and engaging means of educating patients to enhance adherence to care treatments and provide real-time support for their decision-making skills. Notably, studies have shown that patient education becomes more effective with increasing interactions with care providers number [29]. In this context, chatbots hold significant promise for helping patients cultivate healthier habits at scale, ultimately reducing hospital admissions, healthcare costs, and time commitments. Within this framework, NLP solutions have already begun to play a pivotal role as *healthbots*: Parmar et al.[30] conducted a review of healthbots available on the major mobile app stores. However, their findings revealed that most healthbots rely on rule-based approaches and finite-state dialogue management, leaving space for the latest advancements in NLP for application in healthcare.

Tasks for biomedical researchers NLP represents an unprecedented opportunity for biomedical researchers. Apart from employing NLP to extract and summarize information from vast biomedical literature databases, researchers can exploit its techniques to ease the recruitment of patients for studies by identifying people meeting the study criteria from the clinical notes [31, 32], perhaps with appropriate adjustments to handle variations in clinical documentation between different institutions [33]. This expedient leads to building larger cohorts with fewer efforts, which is extremely useful for researchers in the biomedical field.

Furthermore, NLP can be employed to annotate patients' health status from their reports, providing *silver labels* to associate with other kinds of data, like images, and then train the principal AI system, which is often data-hungry, in a supervised way [10].

4. Challenges

The integration of NLP in healthcare has encountered and continues to face unique challenges. The main concern is the imperative to safeguard patient privacy, given that narrative reports within healthcare are filled with sensitive information governed by stringent legislation, such as the U.S. Health Insurance Portability and Accountability Act (HIPAA) and the European Union's General Data Protection Regulation (GDPR). The implications of these regulations are profound, demanding the implementation of appropriate measures to protect individuals' privacy. However, the costs and reliability challenges associated with de-identifying health reports present a prominent hurdle for the NLP community in healthcare. These issues constrain access to shared data, inhibiting collaboration and reproducibility among researchers' teams [34]. Not surprisingly, the first shared task for clinical NLP in 2006, the *Integrating Biology and the Bedside* (i2b2), was centered on the automated removal of Private Health Information (PHI) from medical discharge records [35].

To handle this issue, many researchers have focused their efforts on publicly accessible clinical note databases. For example, databases like the Medical Information Mart for Intensive Care (MIMIC)[5] have enabled researchers to work with extensive datasets, surmounting regulatory obstacles. Also, the recent DR.BENCH [36] has emerged as a new benchmark for model development/evaluation for diagnostic reasoning based on EHRs.

Furthermore, medical and clinical languages present their own complexities [37], encompassing numerous clinical, biological, and medical subdomains, each characterized by its own lexicon, acronyms, abbreviations, ambiguous terms, entities, and variations. To address this challenge, some experts have developed resources such as lexicons, ontologies, and tools tailored for medical and clinical languages. Resources like the Unified Medical Language System (UMLS)[6], along with auxiliary tools such as MetaMap[7], cTakes[8] and CLAMP[9], as well as adapted advanced models and representations [38] serve as valuable assets for NLP practitioners. However, applying these tools to analyze diverse sources like social media, which may exhibit significant language variations, can potentially undermine the effectiveness of such tools [39].

In light of the scarcity of data and the intricacies of a domain-specific language, the issue of multilingualism assumes significance. A majority of the research efforts are oriented towards English, likely due to the lack of available resources in other languages [40].

In addition, the imperative of *explainability* emerges as a critical factor of contemporary research, particularly within healthcare. The lack of interpretability undermines the acceptance and trust of AI and NLP systems in this sensitive domain among both patients and healthcare practitioners [41].

[5] https://mimic.mit.edu/
[6] https://www.nlm.nih.gov/research/umls/index.html
[7] https://lhncbc.nlm.nih.gov/ii/tools/MetaMap.html
[8] https://ctakes.apache.org/
[9] http://clamp.uth.edu/index.php

5. Conclusions

In conclusion, this short paper offers a bird's eye view into the multifaceted panorama of NLP in healthcare, showcasing its significant potential to benefit multiple stakeholders. It encompasses patients seeking personalized health information, healthcare institutions working to streamline their operations, physicians needing clinical decision support, and researchers delving into the complexities of diseases, catalyzing a transformative change in the healthcare landscape.

Nonetheless, NLP in healthcare extends beyond pure academic research; it is a thriving field with real-world applications that can significantly enhance patient care. Moreover, as NLP techniques continue to advance and adapt, an exciting frontier full of unexploited opportunities for researchers and practitioners unfolds.

References

[1] E. Bassignana, D. Brunato, M. Polignano, A. Ramponi, Preface to the Seventh Workshop on Natural Language for Artificial Intelligence (NL4AI), in: Proceedings of the Seventh Workshop on Natural Language for Artificial Intelligence (NL4AI 2023) co-located with 22th International Conference of the Italian Association for Artificial Intelligence (AI* IA 2023), 2023.

[2] C. Friedman, N. Elhadad, Natural Language Processing in Health Care and Biomedicine, Springer London, London, 2014, pp. 255–284.

[3] E. H. Houssein, R. E. Mohamed, A. A. Ali, Machine learning techniques for biomedical natural language processing: a comprehensive review, IEEE Access (2021).

[4] S. Wu, K. Roberts, S. Datta, J. Du, Z. Ji, Y. Si, S. Soni, Q. Wang, Q. Wei, Y. Xiang, et al., Deep learning in clinical natural language processing: a methodical review, Journal of the American Medical Informatics Association 27 (2020) 457–470.

[5] G. Gonzalez-Hernandez, A. Sarker, K. O'Connor, G. Savova, Capturing the patient's perspective: a review of advances in natural language processing of health-related text, Yearbook of medical informatics 26 (2017) 214–227.

[6] I. Li, J. Pan, J. Goldwasser, N. Verma, W. P. Wong, M. Y. Nuzumlalı, B. Rosand, Y. Li, M. Zhang, D. Chang, et al., Neural natural language processing for unstructured data in electronic health records: A review, Computer Science Review 46 (2022) 100511.

[7] E. Hossain, R. Rana, N. Higgins, J. Soar, P. D. Barua, A. R. Pisani, et al., Natural language processing in electronic health records in relation to healthcare decision-making: a systematic review, arXiv preprint arXiv:2306.12834 (2023).

[8] Y. Gao, D. Dligach, L. Christensen, S. Tesch, R. Laffin, D. Xu, T. Miller, O. Uzuner, M. M. Churpek, M. Afshar, A scoping review of publicly available language tasks in clinical natural language processing, Journal of the American Medical Informatics Association 29 (2022) 1797–1806.

[9] T. Zhang, A. M. Schoene, S. Ji, S. Ananiadou, Natural language processing applied to mental illness detection: a narrative review, npj Digital Medicine 5 (2022). doi:10.1038/s41746-022-00589-7, cited by: 5; All Open Access, Gold Open Access, Green Open Access.

[10] L. Bacco, F. Russo, L. Ambrosio, F. D'Antoni, L. Vollero, G. Vadalà, F. Dell'Orletta, M. Merone, R. Papalia, V. Denaro, Natural language processing in low back pain and spine diseases: A systematic review, Frontiers in Surgery 9 (2022).
[11] F. Pethani, A. G. Dunn, Natural language processing for clinical notes in dentistry: a systematic review, Journal of Biomedical Informatics (2023) 104282.
[12] Y. Wang, L. Wang, M. Rastegar-Mojarad, S. Moon, F. Shen, N. Afzal, S. Liu, Y. Zeng, S. Mehrabi, S. Sohn, H. Liu, Clinical information extraction applications: A literature review, Journal of Biomedical Informatics 77 (2018) 34–49.
[13] J. Liang, Y. Li, Z. Zhang, D. Shen, J. Xu, X. Zheng, T. Wang, B. Tang, J. Lei, J. Zhang, Adoption of electronic health records (ehrs) in china during the past 10 years: consecutive survey data analysis and comparison of sino-american challenges and experiences, Journal of medical Internet research 23 (2021) e24813.
[14] K. Hyoun-Joong, Managing unstructured big data in healthcare system, hir 25 (2019) 1–2.
[15] B. Zhao, Web scraping, Encyclopedia of big data (2017) 1–3.
[16] B. Ondov, K. Attal, D. Demner-Fushman, A survey of automated methods for biomedical text simplification, Journal of the American Medical Informatics Association 29 (2022) 1976–1988.
[17] R. Pivovarov, N. Elhadad, Automated methods for the summarization of electronic health records, Journal of the American Medical Informatics Association 22 (2015) 938–947.
[18] A. Rumshisky, M. Ghassemi, T. Naumann, P. Szolovits, V. Castro, T. McCoy, R. Perlis, Predicting early psychiatric readmission with natural language processing of narrative discharge summaries, Translational psychiatry 6 (2016) e921–e921.
[19] R. Kaur, J. A. Ginige, O. Obst, A systematic literature review of automated icd coding and classification systems using discharge summaries, 2021.
[20] L. M. Gattepaille, S. Hedfors Vidlin, T. Bergvall, C. E. Pierce, J. Ellenius, Prospective evaluation of adverse event recognition systems in twitter: Results from the web-radr project, Drug safety 43 (2020) 797–808.
[21] R. Skaik, D. Inkpen, Using social media for mental health surveillance: a review, ACM Computing Surveys (CSUR) 53 (2020) 1–31.
[22] F. Greaves, C. Millett, Consistently increasing numbers of online ratings of healthcare in england, J Med Internet Res 14 (2012) e94.
[23] G. G. Gao, J. S. McCullough, R. Agarwal, A. K. Jha, A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period, Journal of medical Internet research 14 (2012) e38.
[24] S. Gohil, S. Vuik, A. Darzi, Sentiment analysis of health care tweets: Review of the methods used, JMIR Public Health Surveill 4 (2018) e43.
[25] R. D. Zielstorff, Controlled vocabularies for consumer health, Journal of Biomedical Informatics 36 (2003) 326–333. Building Nursing Knowledge through Informatics: From Concept Representation to Data Mining.
[26] L. Bacco, F. Dell'Orletta, H. Lai, M. Merone, M. Nissim, A text style transfer system for reducing the physician-patient expertise gap: An analysis with automatic and human evaluations, Expert Systems with Applications (2023) 120874.
[27] M. A. Sager, A. M. Kashyap, M. Tamminga, S. Ravoori, C. Callison-Burch, J. B. Lipoff, Identifying and responding to health misinformation on reddit dermatology forums with

artificially intelligent bots using natural language processing: Design and evaluation study, JMIR Dermatol 4 (2021) e20975.

[28] P. Patwa, S. Sharma, S. Pykl, V. Guptha, G. Kumari, M. S. Akhtar, A. Ekbal, A. Das, T. Chakraborty, Fighting an infodemic: Covid-19 fake news dataset, in: International Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation, Springer, 2021, pp. 21–29.

[29] T. S. Tang, M. M. Funnell, M. B. Brown, J. E. Kurlander, Self-management support in "real-world" settings: An empowerment-based intervention, Patient Education and Counseling 79 (2010) 178–184.

[30] P. Parmar, J. Ryu, S. Pandya, J. Sedoc, S. Agarwal, Health-focused conversational agents in person-centered care: a review of apps, npj Digital Medicine 5 (2022). Cited by: 2; All Open Access, Gold Open Access, Green Open Access.

[31] R. F. Sarmiento, F. Dernoncourt, Improving Patient Cohort Identification Using Natural Language Processing, Springer International Publishing, Cham, 2016, pp. 405–417.

[32] V. M. Castro, D. Dligach, S. Finan, S. Yu, A. Can, M. Abd-El-Barr, V. Gainer, N. A. Shadick, S. Murphy, T. Cai, et al., Large-scale identification of patients with cerebral aneurysms using natural language processing, Neurology 88 (2017) 164–168.

[33] S. Sohn, Y. Wang, C.-I. Wi, E. A. Krusemark, E. Ryu, M. H. Ali, Y. J. Juhn, H. Liu, Clinical documentation variations and nlp system portability: a case study in asthma birth cohorts across institutions, Journal of the American Medical Informatics Association 25 (2018) 353–359.

[34] W. W. Chapman, P. M. Nadkarni, L. Hirschman, L. W. D'Avolio, G. K. Savova, O. Uzuner, 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 (2011) 540–543.

[35] O. Uzuner, Y. Luo, P. Szolovits, Evaluating the State-of-the-Art in Automatic De-identification, Journal of the American Medical Informatics Association 14 (2007) 550–563.

[36] Y. Gao, D. Dligach, T. Miller, J. Caskey, B. Sharma, M. M. Churpek, M. Afshar, Dr.bench: Diagnostic reasoning benchmark for clinical natural language processing, 2022.

[37] C. Friedman, S. B. Johnson, Natural language and text processing in biomedicine, in: Biomedical Informatics, Springer, 2006, pp. 312–343.

[38] E. Alsentzer, J. R. Murphy, W. Boag, W.-H. Weng, D. Jin, T. Naumann, M. McDermott, Publicly available clinical bert embeddings, arXiv preprint arXiv:1904.03323 (2019).

[39] K. Denecke, Information extraction from medical social media, in: Health Web Science, Springer, 2015, pp. 61–73.

[40] A. Névéol, H. Dalianis, S. Velupillai, G. Savova, P. Zweigenbaum, Clinical natural language processing in languages other than english: opportunities and challenges, Journal of biomedical semantics 9 (2018) 1–13.

[41] H. W. Loh, C. P. Ooi, S. Seoni, P. D. Barua, F. Molinari, U. R. Acharya, Application of explainable artificial intelligence for healthcare: A systematic review of the last decade (2011–2022), Computer Methods and Programs in Biomedicine 226 (2022) 107161.