

Review

Applications of Natural Language Processing and Large Language Models for Social Determinants of Health: Systematic Review

Swati Rajwal¹, MS; Avinash Kumar Pandey², MS; Ziyuan Zhang³, MS; Yankai Chen⁴, PhD; Michael X Liu⁴, MBA; Sudeshna Das¹, PhD; Hannah Rogers⁵, MS; Abeed Sarker¹, PhD; Yunyu Xiao⁴, PhD

¹Department of Biomedical Informatics, School of Medicine, Emory University, Atlanta, GA, United States

²Laney Graduate School, Emory University, Atlanta, GA, United States

³Department of Epidemiology, University of California, Berkeley, Berkeley, CA, United States

⁴Department of Population Health Sciences, Cornell University, New York, NY, United States

⁵Woodruff Health Sciences Center Library, Emory University, Atlanta, GA, United States

Corresponding Author:

Swati Rajwal, MS
Department of Biomedical Informatics
School of Medicine, Emory University
101 Woodruff Circle
Atlanta, GA 30322
United States
Phone: 1 4704478469
Email: swati.rajwal@emory.edu

Abstract

Background: Social determinants of health (SDOH) are the social, economic, and environmental conditions that influence health outcomes. SDOH information is often embedded in unstructured text, such as notes in electronic health records and social media posts. Advances in natural language processing (NLP), including emergent large language models (LLMs), offer opportunities to extract, analyze, and interpret SDOH expressions from free text for inclusion in downstream analyses. Existing literature on NLP applications for SDOH is dispersed across disciplines and characterized by methodological heterogeneity and variability in study quality and scope, complicating synthesis and cross-study comparison.

Objective: This study aimed to examine the use of NLP, including LLMs, in SDOH research, and highlight gaps and future research directions.

Methods: We conducted a systematic review following PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines, searching 7 major databases for publications between 2014 and November 2025. We included journal and conference proceedings papers that applied NLP methods to identify, classify, extract, or predict SDOH from text. Three reviewers independently screened studies and extracted data; conflicts were resolved by two senior reviewers. We abstracted study metadata, dataset characteristics, NLP approaches, SDOH domains addressed, and NLP performance metrics. We also conducted risk-of-bias analyses and identified influential studies based on relative citation counts.

Results: 142 studies met the inclusion criteria. Nearly two-thirds (89/142, 62.7%) were published between 2023 and 2025, reflecting rapid recent growth. Most studies relied on electronic health records (93/142, 65.5%) and private datasets (81/142, 57.0%), while only 20.4% (29/142) used publicly available data. Commonly studied SDOH domains were housing instability (72/142, 50.7%), employment (65/142, 45.8%), and financial conditions (63/142, 44.4%); structural factors, such as immigration status (5/142, 3.5%), were rarely examined. Of studies that reported evaluation metrics, most focused on classification (26/83, 31.32%) or extraction (38/83, 45.7%), and used cross-sectional designs. Reported model performances were typically strong, with median F_1 -scores ranging roughly from 0.75 to 0.85 across model categories. Only 49 studies shared code, and fewer than half clearly described model interpretability or reproducibility practices. LLMs (including encoder-decoder models) appeared in 19.7% (28/142) of studies, highlighting emerging interest but also raising new concerns around transparency and governance.

Conclusions: This review provides a timely synthesis of NLP and LLM applications across the SDOH research spectrum, addressing an important gap in a topic receiving increasing research attention. By comparing task formulations, data sources, and performance patterns, the review clarifies the research readiness of current approaches and reveals critical gaps. Our findings advance the field by highlighting the absence of a unified SDOH framework, uneven availability of public benchmarks, and limited evaluation of real-world deployment. Addressing these gaps through transparent, inclusive dataset development and implementation-focused evaluation is essential for translating NLP advances into equitable, real-world health impact.

Trial Registration: PROSPERO CRD42024578082; <https://www.crd.york.ac.uk/PROSPERO/view/CRD42024578082>

International Registered Report Identifier (IRRID): RR2-10.2196/66094

J Med Internet Res 2026;28:e83793; doi: [10.2196/83793](https://doi.org/10.2196/83793)

Keywords: social determinants of health; SDOH; natural language processing; large language models; systematic review; PRISMA

Introduction

Social determinants of health (SDOH) refer to nonmedical factors such as the social, economic, and environmental conditions that shape where people live, work, and age. SDOH are among the most powerful drivers of health outcomes and disparities worldwide [1,2], with influences on health [3] and well-being at the individual and population levels [4]. Contemporary estimates suggest that medical care accounts for only 10%-20% of the modifiable contributors to healthy outcomes, while SDOH-related factors drive the remaining 80%-90% [3,5,6]. Consequently, SDOH factors, such as housing instability, food insecurity, and structural racism, have become central to understanding the persistence of disease, quality of life, and mortality across diverse populations [7,8]. Accurately capturing SDOH and incorporating them into health care is crucial for clinicians, health systems, and policymakers aiming to address structural inequities and enhance care delivery. Clinicians, for example, can work with patients in social prescribing based on affordability or the availability of transportation to the relevant pharmacy. Similarly, with an in-depth understanding of structural SDOH, including place-based contextual metrics of economic, educational, health, and environmental conditions, policymakers may effectively guide health policies [9-11]. While health systems increasingly recognize that addressing social needs is essential for value-based care, the mechanisms to systematically identify and act upon these factors are often underdeveloped [12].

Several frameworks have been proposed to characterize SDOH into groups. The Healthy People 2030 initiative [2] groups SDOH into 5 domains: Economic Stability, Education Access and Quality, Health Care Access and Quality, Neighborhood and Built Environment, and Social and Community Context. The World Health Organization conceptualizes SDOH into 2 interacting groups: Structural and Intermediary, with the former having causal priority in influencing health outcomes. Due to their importance in influencing health, SDOH codes were also introduced into the International Classification of Diseases, 10th Revision, Clinical Modification in 2015. These are categorized as Z-codes (Z55-Z65) and are used to document socioeconomic

and psychosocial circumstances that affect a person's health. Yet, the documentation of Z-codes has been low in health systems [13]. Thus, SDOH information is often buried in unstructured text, such as notes in electronic health records (EHRs) and social media postings, where it may be expressed through nuanced or implicit language, limiting information accessibility through conventional methods and requiring the development of customized text mining approaches [14].

Natural language processing (NLP) is a subfield of artificial intelligence and computer science that enables computers to process, understand, interpret, and generate human language in meaningful and useful ways. NLP encompasses a wide range of tasks, including but not limited to text classification, information extraction/entity recognition, text summarization, and language translation. NLP offers a promising tool to systematically analyze vast amounts of unstructured text data, EHRs, social media [15,16], public health reports, and other sources [17]. While biomedical domain NLP has been widely applied to extract clinical concepts such as diagnoses, medications, and procedures, much less work has focused on nonclinical aspects, such as SDOH and their influence on health outcomes.

In the context of SDOH, NLP can assist in parsing and understanding context, co-references, and the relationships between text parts. Unlike manual review (time-consuming, labor-intensive, and prone to human error), NLP enables the rapid and scalable analysis of large datasets with greater accuracy and consistency [18,19] when models are appropriately adapted to the domain and data, although performance may vary across populations and SDOH categories. A common NLP application is the identification and extraction of SDOH factors from EHRs and other text-based data, enabling the systematic classification of social needs, behavioral drivers, and environmental conditions that influence patient outcomes [20-24]. The versatility of NLP is evident in its application across diverse clinical domains, ranging from pediatric populations [25,26] to patients managing chronic conditions such as Alzheimer disease [22] and lower back pain [27]. Beyond traditional medical notes, NLP workflows have been adapted to mine insights from specialized documentation, including clinical social work notes [28] and emergency medical services

records [29], as well as nonclinical data sources like social media, where it has been used to assess the impact of external crises (eg, COVID-19) on marginalized communities [30]. Underpinning these diverse applications is the rapid evolution of model architectures; specifically, the shift toward transformer-based models and generative pretrained transformers has significantly enhanced the precision of SDOH extraction from free-text data [31,32].

Despite these technological advancements, the literature remains fragmented in the application of NLP to characterize SDOH, due to disparate disciplines, data modalities, and methodological frameworks. A comprehensive NLP workflow for SDOH analysis necessitates a rigorous pipeline: (1) defining the target SDOH elements or categories; (2) selecting the appropriate NLP modeling strategy (eg, classification vs extraction); (3) curating gold-standard annotated datasets for supervision and validation; and (4) deploying the optimal strategy. Currently, no comprehensive synthesis exists that integrates these myriad applications of NLP and large language models (LLMs) for SDOH. Consequently, there is a critical knowledge gap regarding the most effective approaches to modeling SDOH, limiting the translation of these technical capabilities into standardized public health and clinical practice.

To address this gap, we conducted a systematic review of peer-reviewed studies published between 2014 and 2025. This period is marked by the introduction of Word2Vec [33], transformer models [34] such as BERT [35], and generative models like the generative pretrained transformer series [36]. Our review consolidates evidence from health, informatics, and computer science literature to assess the state of the science in applying NLP to SDOH. We broadly categorize SDOH into individual and structural factors, a simplification of more comprehensive frameworks such as Healthy People 2030 and the World Health Organization Commission on SDOH. Individual-level SDOH refer to personal circumstances such as income, education, employment status, housing conditions, and access to health care, which directly influence an individual's health outcomes. Structural SDOH encompass the broader systemic and institutional contexts (such as policies, social norms, economic systems, and structural racism) that shape and constrain the distribution of individual-level resources and opportunities [10]. This systematic review has two major objectives: (1) to characterize NLP techniques, including LLMs, used to analyze SDOH in unstructured individual and public data, including annotation, prediction, detection, and classification; and (2) to assess the effectiveness of such techniques or models, identify potential knowledge gaps, and reveal research questions relevant for future studies.

Methods

Registration and Protocol

The systematic review was registered in PROSPERO (CRD42024578082), and the protocol was published in JMIR Research Protocol [37]. The review followed the

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [38], and the search strategy was reported in accordance with PRISMA-S [39]. We conducted a narrative synthesis without meta-analysis following the Synthesis Without Meta-analysis guidelines [40] due to substantial methodological heterogeneity across studies. Formal subgroup, heterogeneity, sensitivity analyses, reporting bias analysis, or certainty of evidence assessment were not conducted as the review did not involve meta-analysis and instead examined differences across studies descriptively. Synthesis focused on identifying patterns and trends in NLP methods, SDOH domains, data sources, and reported performance rather than pooling quantitative estimates. The interdisciplinary review team consisted of a librarian and researchers with expertise in health informatics, public health, and social sciences.

Ethical Considerations

This work does not involve human participants and does not require approval from the IRB.

Eligibility Criteria

We included peer-reviewed literature (journal articles and full conference papers) published between January 2014 and November 2025. The selection of this timeframe was deliberate, as it strategically captures the entire modern deep learning era in NLP, from the rise of word embeddings (eg, Word2Vec) around 2014 to the development of the transformer architecture in 2017 and the subsequent explosion of LLMs. Only studies in English were included, although the underlying NLP data could be in any language. Eligible studies met the following specific criteria: (1) addressed a research question involving the design, development, and application of NLP in health data analysis for SDOH; (2) used text-based data; and (3) used NLP techniques or LLMs (open-source or commercial) related to SDOH. Studies using NLP in conjunction with other methods were eligible, provided NLP was a core component. Preprints (arXiv/bioRxiv), forewords, prefaces, table of contents, programs, schedules, indexes, call for papers or participation, lists of reviewers, lists of tutorial abstracts, invited talks, appendices, session information, obituaries, book reviews, newsletters, lists of proceedings, lifetime achievement awards, erratum, systematic reviews, scoping reviews, and notes were excluded.

Database and Search Strategy

We systematically searched 7 databases: PubMed, Scopus, Web of Science, PsycINFO, Health Source: Nursing/Academic, ACL Anthology, and IEEE Xplore. Searches were conducted independently in each database rather than through a multi-database platform. No study registries, manual browsing of web-based resources, citation chaining, or direct contact with authors or experts were undertaken as part of the search methodology. The search strategies were developed specifically for this review and were not adapted from previous reviews. No additional information sources or search methods were used.

The search spanned publications from January 1, 2014, to November 2, 2025. A health sciences librarian (HR) developed and iteratively refined the search strategy in consultation with the research team, using both controlled

vocabulary and free-text terms. The final PubMed search strategy, including filters, is detailed in [Textbox 1](#). Comprehensive search strategies for all information sources are provided in [Multimedia Appendix 1](#).

Textbox 1. Search query for PubMed. A comprehensive list of all search queries customized for each database is available in [Multimedia Appendix 1](#).

Query:

- (“Natural Language Processing”[Mesh] OR “natural language”[tw] OR NLP[tw] OR “large LM*”[tw] OR LLM[tw] OR LLMs[tw] OR “large language model*”[tw] OR ChatGPT*[tw] OR “Chat GPT*”[tw] OR GPT4*[tw] OR GPT-4*[tw] OR GPT3*[tw] OR GPT-3*[tw] OR “Generative Pre-trained Transformer*”[tw] OR LLAMA[tw] OR “Claude 3”[tw] OR Mistral[tw] OR MedPaLM*[tw] OR Med-PaLM*[tw] OR “text mining”[tw] OR “text process*”[tw] OR “information retrieval”[tw] OR “information extract*”[tw]) AND (“Social Determinants of Health”[Mesh] OR SDOH[tw] OR SDH[tw] OR SBDH*[tw] OR “determinants of health”[tw] OR “health determina*”[tw] OR “life events”[tw] OR “social determinant*”[tw] OR “socioeconomic determinant*”[tw] OR “socioeconomic factor*”[tw] OR “social determinate*”[tw] OR “social factor*”[tw] OR “social need*”[tw] OR “social prescribing”[tw] OR “social determining factor*”[tw] OR “social risk*”[tw])

Filters:

- Language: English
- Years: 2014-2025
- Exclude: Preprints

Selection Process

Three reviewers (SR, ZZ, and YC) independently screened each study for eligibility by marking it as a “yes” (for inclusion), “no” (for exclusion), or “maybe” (in case of uncertainty about relevance) on the Covidence platform [41]. Two senior reviewers (AS and YX) resolved potential discrepancies during any screening step. Blinded voting ensured that reviewers did not view others’ decisions during the screening process. The reviewers retrieved eligible studies for second-stage review (full-text) using the same inclusion criteria and removed those that did not meet them. The final set of studies to include was approved by consensus of all reviewers.

Data Extraction

Before formal data extraction, 3 reviewers (SR, ZZ, and YC) piloted a structured data extraction form using 5 sample studies to ensure clarity and consistency. Six independent reviewers (SR, ZZ, YC, AKP, ML, and SD) then conducted the full data extraction in teams of 2, with each member independently extracting half of the assigned studies. Ambiguities were resolved within the teams to maintain accuracy and consistency. The final data captured study metadata (eg, year of publication and type), dataset characteristics (eg, source, sample size, and type), NLP approaches (eg, models used and task type), SDOH domains addressed, and performance metrics (eg, precision, F_1 -score, and recall). We treated these performance metrics as effect measures for outcomes of interest (ie, model performance). Due to heterogeneity in study design, NLP tasks, datasets, and evaluation frameworks, no single standardized effect size metric was applicable, and reported metrics were extracted and summarized descriptively without statistical transformation. Model performance metrics reported by individual studies were extracted as reported and treated descriptively. We summarized study characteristics in structured tables

and used visualizations, including a PRISMA flow diagram, frequency plots, heatmaps, and bubble charts, to display study selection, SDOH domains, and NLP model performance patterns. We assessed risk of bias using a short checklist adapted from the Joanna Briggs Institute [42] and covered 6 items: reporting of population demographics (Q1), clarity of study aim (Q2), relevance to SDOH and NLP (Q3), use of a reference standard (Q4), reporting of evaluation measures (Q5), and acknowledgment of study limits (Q6). Each item was rated as yes, no, or unclear. Studies were classified as low (≥ 5 yes), moderate (3-4 yes), or high (≤ 3 yes) risk of bias.

Influential Studies

To assess visibility and identify recurring methodological, collaborative, and reporting features, we calculated a normalized citation ratio (NCR) for each study:

$$\text{NCR} = \frac{\text{Paper's citations}}{\text{Average of its publication year's cohort}}$$

This approach adjusts for citation differences based on publication year.

Results

Study Selection and Characteristics

[Figure 1](#) shows the number of articles included in each phase of the screening process, resulting in the eventual inclusion of 142 studies ([Multimedia Appendix 1](#) shows the search strategy results across databases). [Figure 2](#) presents publication trends over the years, illustrating a growing research interest over time, with the majority of studies being published between 2023 and 2025 (89/142, 62.7%). The majority of the studies were published in journals (109/142, 76.7%), while the remaining were in conference

or workshop proceedings (Figure 2C). Also, as shown in Figure 2B, single-site studies were more common overall, but there has been a recent rise in multi-site studies (data from multiple institutes). Most studies were rated as low risk of bias (116/142), 26 as moderate, and none as high risk (Table 1). Table S5 in Multimedia Appendix 1 presents a

summary of studies. Also, because this review used descriptive narrative synthesis across highly heterogeneous studies, formal heterogeneity analyses, sensitivity analyses, reporting-bias assessment, and certainty-of-evidence grading were not conducted.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram outlining the inclusion and screening processes for study selection. ACL: Association for Computational Linguistics; NLP: natural language processing; SDOH: social determinants of health.

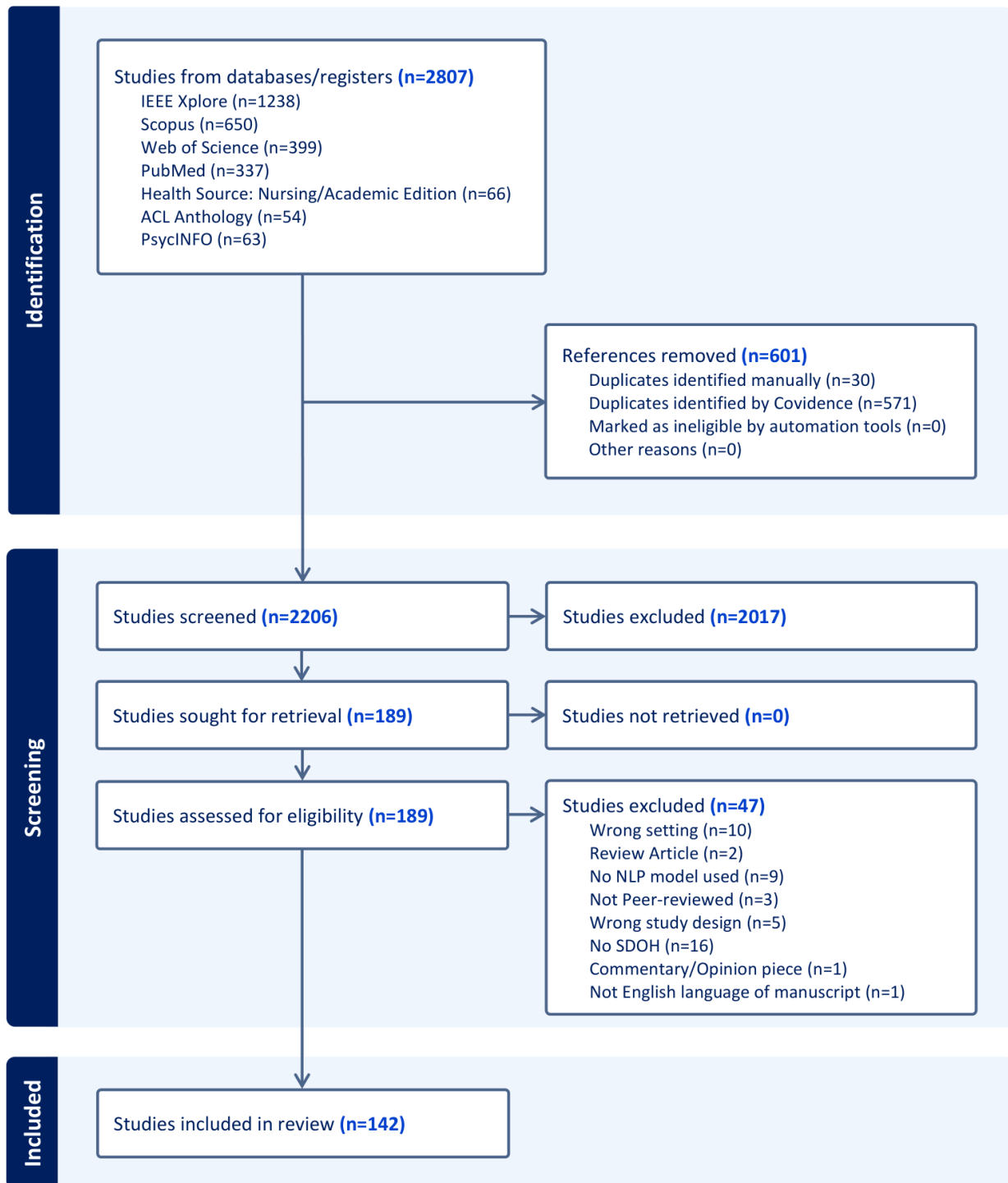


Figure 2. Yearly publication trends in natural language processing for SDOH research illustrate increasing interest over time. (A) SDOH type distribution. (B) Single- versus multi-institution dataset usage. (C) Publication venue distribution. Multi-institution datasets and journal publications increased substantially after 2022. Note that only partial data was available for 2025 at the time of review. SDOH: social determinants of health.

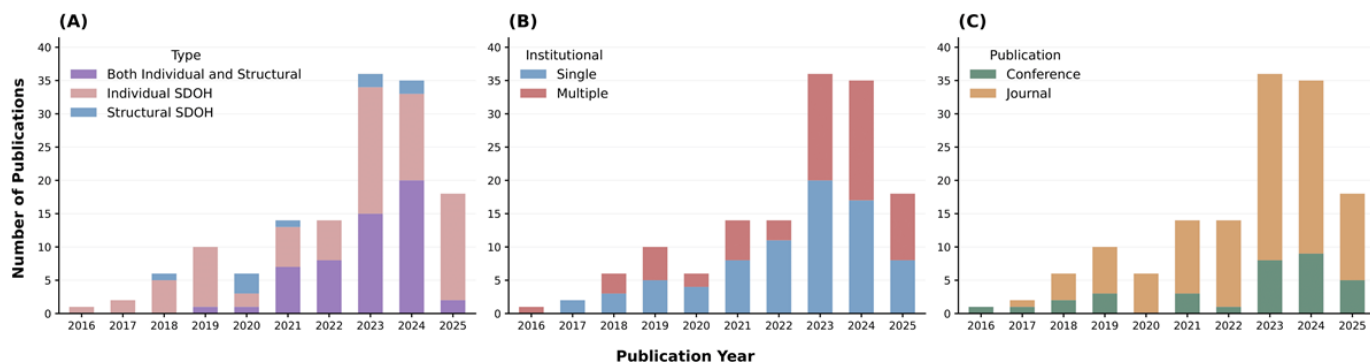


Table 1. Risk of bias assessment for included studies.

Reference	Q1	Q2	Q3	Q4	Q5	Q6	Total “Y”	Risk category
[43]	✓ ^a	✓	✓	✓	✓	✓	6	↓ ^b
[44]	✓	✓	✓	✓	• ^c	✓	5	↓
[45]	✓	✓	✓	✓	•	✓	5	↓
[46]	✓	✓	✓	✓	•	✓	5	↓
[47]	✓	✓	✓	✓	•	✓	5	↓
[48]	✓	✓	✓	× ^d	×	✓	4	≈ ^e
[49]	✓	✓	✓	✓	•	✓	5	↓
[50]	×	✓	✓	✓	•	✓	4	≈
[51]	✓	✓	✓	×	•	✓	4	≈
[26]	×	✓	✓	✓	✓	✓	5	↓
[28]	✓	✓	✓	×	×	✓	4	≈
[52]	✓	✓	✓	✓	✓	✓	6	↓
[53]	×	✓	✓	✓	✓	✓	5	↓
[54]	×	✓	✓	×	•	✓	3	≈
[55]	×	✓	✓	✓	✓	✓	5	↓
[56]	✓	✓	✓	✓	✓	✓	6	↓
[57]	✓	✓	✓	✓	✓	✓	6	↓
[30]	✓	✓	✓	✓	✓	✓	6	↓
[58]	×	✓	✓	✓	✓	✓	5	↓
[59]	×	✓	✓	✓	•	✓	4	≈
[22]	×	✓	✓	✓	✓	✓	5	↓
[60]	✓	✓	✓	✓	•	✓	5	↓
[61]	✓	✓	✓	✓	✓	✓	6	↓
[62]	✓	✓	✓	✓	✓	✓	6	↓
[63]	✓	✓	✓	✓	•	✓	5	↓
[23]	✓	✓	✓	✓	✓	✓	6	↓
[64]	✓	✓	✓	✓	•	✓	5	↓
[65]	✓	✓	✓	✓	•	✓	5	↓
[66]	✓	✓	✓	✓	•	✓	5	↓
[67]	✓	✓	✓	✓	•	✓	5	↓
[29]	✓	✓	✓	✓	•	✓	5	↓
[68]	✓	✓	✓	✓	✓	✓	6	↓
[69]	✓	✓	✓	✓	✓	✓	6	↓
[70]	×	✓	✓	✓	•	✓	4	≈
[31]	×	✓	✓	✓	•	✓	4	≈
[71]	✓	✓	✓	✓	✓	✓	6	↓

Reference	Q1	Q2	Q3	Q4	Q5	Q6	Total “Y”	Risk category
[72]	✓	✓	✓	✓	•	✓	5	↓
[73]	✓	✓	✓	✓	•	✓	5	↓
[74]	×	✓	✓	✓	✓	✓	5	↓
[75]	✓	✓	✓	✓	×	✓	5	↓
[76]	×	✓	✓	✓	✓	✓	5	↓
[77]	×	✓	✓	✓	✓	✓	5	↓
[78]	×	✓	✓	✓	✓	✓	5	↓
[79]	✓	✓	✓	✓	•	✓	5	↓
[80]	×	✓	✓	✓	✓	✓	5	↓
[81]	×	✓	✓	✓	✓	✓	5	↓
[82]	✓	✓	✓	✓	✓	✓	6	↓
[83]	✓	✓	✓	✓	•	✓	5	↓
[32]	×	✓	✓	✓	✓	✓	5	↓
[84]	×	✓	✓	✓	✓	✓	5	↓
[85]	×	✓	✓	✓	✓	✓	5	↓
[86]	×	✓	✓	✓	✓	✓	5	↓
[87]	×	✓	✓	✓	✓	✓	5	↓
[88]	✓	✓	✓	×	•	✓	4	≈
[27]	×	✓	✓	✓	✓	✓	5	↓
[89]	✓	✓	✓	✓	•	✓	5	↓
[25]	✓	✓	✓	✓	•	✓	5	↓
[90]	×	✓	✓	✓	✓	✓	5	↓
[91]	✓	✓	✓	✓	•	✓	5	↓
[92]	×	✓	✓	✓	✓	✓	5	↓
[93]	✓	✓	✓	✓	✓	✓	6	↓
[94]	✓	✓	✓	✓	✓	✓	6	↓
[95]	✓	✓	✓	✓	×	✓	5	↓
[96]	✓	✓	✓	✓	×	✓	5	↓
[97]	✓	✓	✓	✓	•	✓	5	↓
[98]	✓	✓	✓	✓	✓	✓	6	↓
[99]	×	✓	✓	✓	✓	✓	5	↓
[100]	×	✓	✓	✓	✓	✓	5	↓
[101]	×	✓	✓	✓	✓	✓	5	↓
[102]	✓	✓	✓	✓	✓	✓	6	↓
[103]	✓	✓	✓	×	×	✓	4	≈
[104]	×	✓	✓	✓	•	✓	4	≈
[20]	✓	✓	✓	✓	•	✓	5	↓
[105]	✓	✓	✓	✓	×	✓	5	↓
[106]	✓	✓	✓	✓	✓	✓	6	↓
[21]	✓	✓	✓	✓	•	✓	5	↓
[107]	✓	✓	✓	✓	•	✓	5	↓
[108]	✓	✓	✓	✓	✓	✓	6	↓
[109]	×	✓	✓	✓	×	✓	4	≈
[110]	×	✓	✓	✓	•	✓	4	≈
[111]	×	✓	✓	✓	✓	✓	5	↓
[112]	×	✓	✓	×	×	✓	3	≈
[113]	×	✓	✓		×	✓	3	≈
[114]	✓	✓	✓	✓	✓	✓	6	↓
[115]	×	✓	✓	✓	•	✓	4	≈

Reference	Q1	Q2	Q3	Q4	Q5	Q6	Total "Y"	Risk category
[116]	x	✓	✓	✓	✓	✓	5	↓
[117]	✓	✓	✓	✓	•	✓	5	↓
[118]	x	✓	✓	✓	✓	✓	5	↓
[12]	x	✓	✓	✓	✓	✓	5	↓
[119]	✓	✓	✓	✓	✓	✓	6	↓
[120]	x	✓	✓	✓	•	✓	4	≈
[121]	✓	✓	✓	✓	•	✓	5	↓
[122]	x	✓	✓	✓	✓	✓	5	↓
[123]	x	✓	✓	✓	✓	✓	5	↓
[124]	x	✓	✓	✓	x	✓	4	≈
[125]	x	✓	✓	✓	✓	✓	5	↓
[126]	x	✓	✓	✓	✓	✓	5	↓
[127]	x	✓	✓	✓	•	✓	4	≈
[128]	x	✓	✓	✓	✓	✓	5	↓
[129]	x	✓	✓	x	•	✓	3	≈
[130]	✓	✓	✓	✓	✓	✓	6	↓
[131]	x	✓	✓	✓	•	✓	4	≈
[132]	✓	✓	✓	✓	✓	✓	6	↓
[133]	✓	✓	✓	✓	•	✓	5	↓
[134]	✓	✓	✓	✓	✓	✓	6	↓
[135]	x	✓	✓	✓	✓	✓	5	↓
[136]	✓	✓	✓	✓	•	✓	5	↓
[137]	x	✓	✓	✓	✓	✓	5	↓
[138]	x	✓	✓	✓	•	✓	4	≈
[139]	x	✓	✓	✓	✓	✓	5	↓
[140]	x	✓	✓	✓	•	✓	4	≈
[141]	✓	✓	✓	✓	✓	✓	6	↓
[142]	✓	✓	✓	✓	✓	✓	6	↓
[143]	x	✓	✓	✓	✓	✓	5	↓
[144]	x	✓	✓	✓	✓	✓	5	↓
[145]	x	✓	✓	✓	✓	✓	5	↓
[146]	✓	✓	✓	✓	•	✓	5	↓
[147]	x	✓	✓	✓	✓	✓	5	↓
[148]	✓	✓	✓	✓	✓	✓	6	↓
[149]	x	✓	✓	✓	✓	✓	5	↓
[150]	✓	✓	✓	✓	✓	✓	6	↓
[151]	✓	✓	✓	✓	✓	✓	6	↓
[152]	x	✓	✓	✓	✓	✓	5	↓
[153]	✓	✓	✓	✓	✓	✓	6	↓
[154]	x	✓	✓	✓	•	✓	4	≈
[155]	x	✓	✓	✓	✓	✓	5	↓
[156]	x	✓	✓	✓	✓	✓	5	↓
[157]	✓	✓	✓	✓	✓	✓	6	↓
[158]	✓	✓	✓	✓	✓	✓	6	↓
[24]	✓	✓	✓	✓	✓	✓	6	↓
[159]	x	✓	✓	✓	✓	✓	5	↓
[160]	x	✓	✓	✓	✓	✓	5	↓
[161]	x	✓	✓	✓	✓	✓	5	↓
[162]	x	✓	✓	✓	✓	✓	5	↓

Reference	Q1	Q2	Q3	Q4	Q5	Q6	Total "Y"	Risk category
[163]	✓	✓	✓	✓	✓	✓	6	↓
[164]	✓	✓	✓	✓	✓	✓	6	↓
[165]	✓	✓	✓	✓	✓	✓	6	↓
[166]	✓	✓	✓	✓	•	✓	5	↓
[167]	✓	✓	✓	✓	✓	✓	6	↓
[168]	✓	✓	✓	✓	✓	✓	6	↓
[169]	✓	✓	✓	✓	•	✓	5	↓
[170]	✓	✓	✓	✓	•	✓	5	↓

^a✓: Yes.

^b↓: Low risk of bias.

^c•: Unclear.

^d×: No.

^e≈: Moderate risk of bias.

Data Sources and Characteristics

Across the 142 studies reviewed, over half used private datasets (81/142, 57.04%), while publicly accessible datasets were used in 29 studies (20.42%). A smaller number of studies used datasets that were available via specific data use agreements (7/142, 4.93%), accessible only for shared tasks (5/142, 3.52%), or available in partial form (subset available, 5/142, 3.52%). The most common data type was EHRs, reported in 93 studies. Eight studies used national datasets (Primary Land Use Tax Lot Output [PLUTO], CalEnviroScreen 4.0, etc). Social media sources, including Twitter, Reddit, and Facebook, were leveraged in 9 studies (5.1%). Other, less frequently reported types included research abstracts, web content, and interviews.

The majority of private datasets included institution-specific EHRs (eg, Columbia University Irving Medical Center; University of California, San Francisco; Johns Hopkins; Eskenazi; Medical University of South Carolina; UF Health; UNC; Kaiser Permanente Southern California) and clinical data warehouses tied to academic or health systems. The Veterans Affairs health system was particularly prominent, covering datasets such as: Veterans Affairs Corporate Data Warehouse [46,47,75,83,89,90,94,135], Veteran Health Administration notes [114], Supportive Services for Veteran Families [75], administrative data [136], Veterans Aging Cohort Study [141]. Several datasets incorporated augmented sources to enrich available information, such as LexisNexis, geospatial SDOH (from diverse government sources [47]), and patient-reported tools (eg, Timeline Follow-Back [83]). Publicly available datasets included Medical Information Mart for Intensive Care III (MIMIC-III) [31,32,55,57,58,74,77,78,82,85,86,95,97,104,118,125,128,130], n2c2 2018/2022 [55,58,80,85,128], SemEval-2015 [128], LGBTQ+ Minority

Stress on Social Media (MiSSoM+) dataset [92], and others. MIMIC-III and its derivatives (eg, social and behavioral determinants of health-MIMIC) were frequently used for training [85,86,95], validation [82], or annotation purposes [32].

SDOH Factors

Figure 3 illustrates the most commonly studied SDOH factors across the years. We show factors that appeared in at least 15 papers. Table 2 shows the SDOH type (individual vs structural) studied in our cohort of papers. The most studied SDOH factors were housing instability (50.7%) and employment (45.8%), appearing in nearly half of the 142 papers, followed by financial context (44.4%), substance use (37.3%), and social isolation (34.5%). Education and living circumstances each appeared in 27.5% and 28.9% of studies, respectively. At the lower end, justice system involvement and language literacy each appeared in approximately 8% of studies. The least studied factors included immigration status (3.5%) and 4 factors that appeared in less than 3 papers each: access to lethal means, acculturation, digital divide, and military sexual trauma. Collectively, these results show that individual SDOH (substance use, social connection/isolation, etc) received more research attention than structural SDOH (transportation, insurance, etc) over the years. The uneven focus on certain SDOH could be due to data and methodological constraints rather than importance alone. Individual-level factors such as housing and employment are more explicitly documented and easier to annotate, whereas structural determinants (eg, immigration status or the digital divide) are often inconsistently recorded or absent from clinical text.

Figure 3. Distribution of the most studied social determinants of health over the years. Bubble size is proportional to the number of studies. Individual-level determinants are frequently studied compared with structural determinants. Table S2 in [Multimedia Appendix 1](#) for the category dictionary. *partial data for 2025.

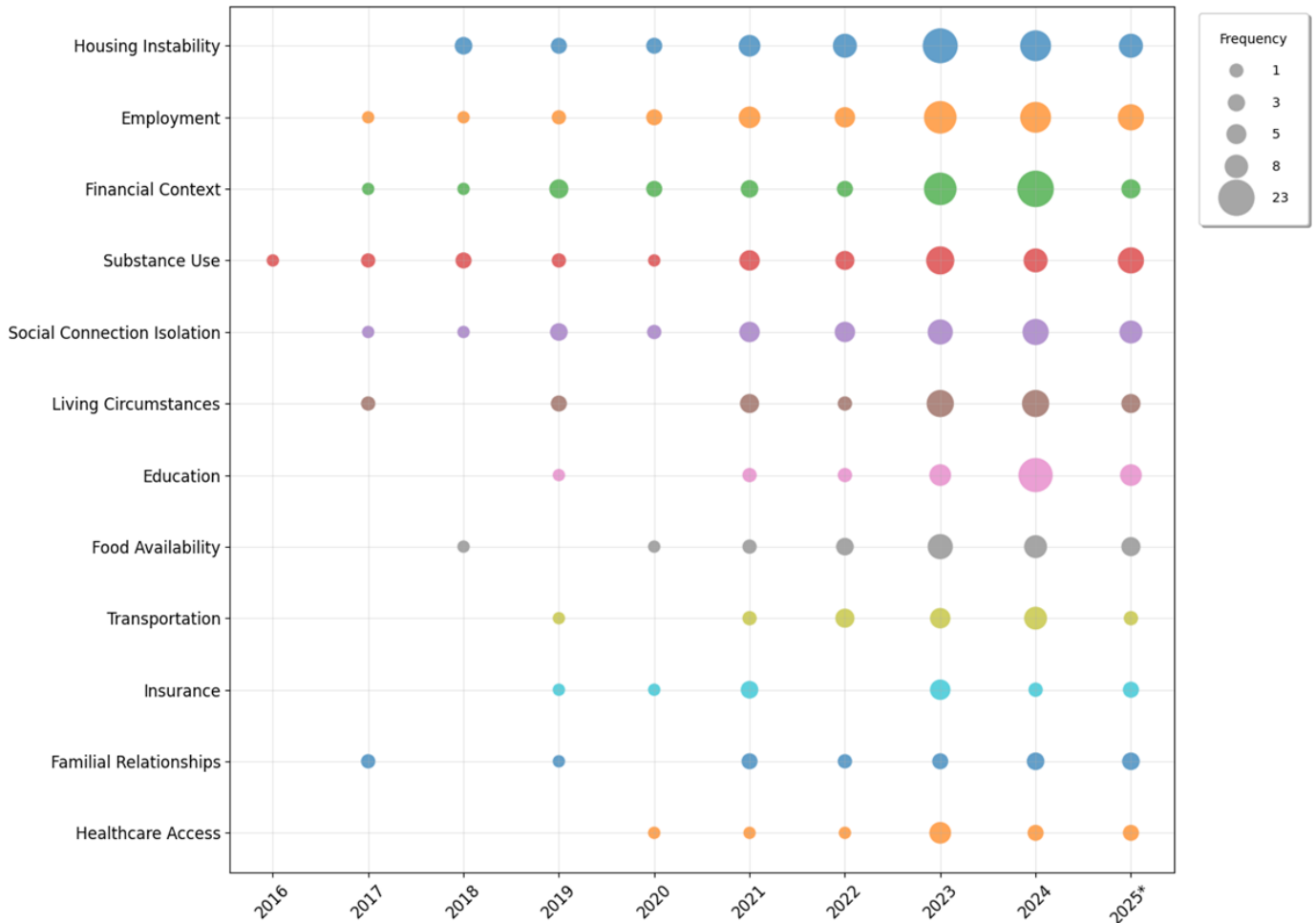


Table 2. Social Determinants of Health (SDOH) type distribution.

SDOH type	Reference
Individual (n=79)	[43,111,115-117,119,122-125,127,138,143,144], [12,20,87,93,94,98,99,101,114,118,121,130,136,141,142], [22,23,29,31,45,47,55,57,58,62,70,71,73,74,78,80-83,85,86,90,132-134,137,145-147,149,150,152], [24,135,148,155-166,168-170]
Structural (n=9)	[48,53,59,88,110,112,120,129,126]
Both (n=54)	[21,44,46,95,96,100,102-109,113], [25,27,32,52,56,63,64,67,75-77,79,84,89,91,92,97], [26,28,30,49-51,54,60,61,65,66,68,69,72,128,131,139,140,151,153,154,167]

NLP Methods and Performance

Figure 4A presents a heatmap illustrating the temporal distribution and frequency of the dominant NLP methods identified in this review. Among the included studies, the majority (n=83) quantified model performance using standard metrics, including F_1 -score, precision, and recall. Figure 4B delineates the distribution of F_1 -scores across the 5 categories, highlighting significant performance variations

between rule-based, traditional machine learning, and deep learning approaches. Complementing this, Figure 5 provides a bubble chart depicting the mean precision and recall values for each model category. Collectively, the visualized performance trends show the general capabilities of different modeling approaches for SDOH tasks, though performance may vary based on task complexity, dataset characteristics, and annotation quality.

Figure 4. (A) Frequency heatmap representative of the best-performing NLP methodologies across years. Interestingly, transformer- and large language model-based studies surged over the last 3 years. (B) Box plot showing F_1 -score distribution by model category. Models with fewer than 20 samples are excluded, and extreme values are removed for readability. Table S3 in Multimedia Appendix 1 details the category dictionary. *partial data for 2025. NLP: natural language processing.

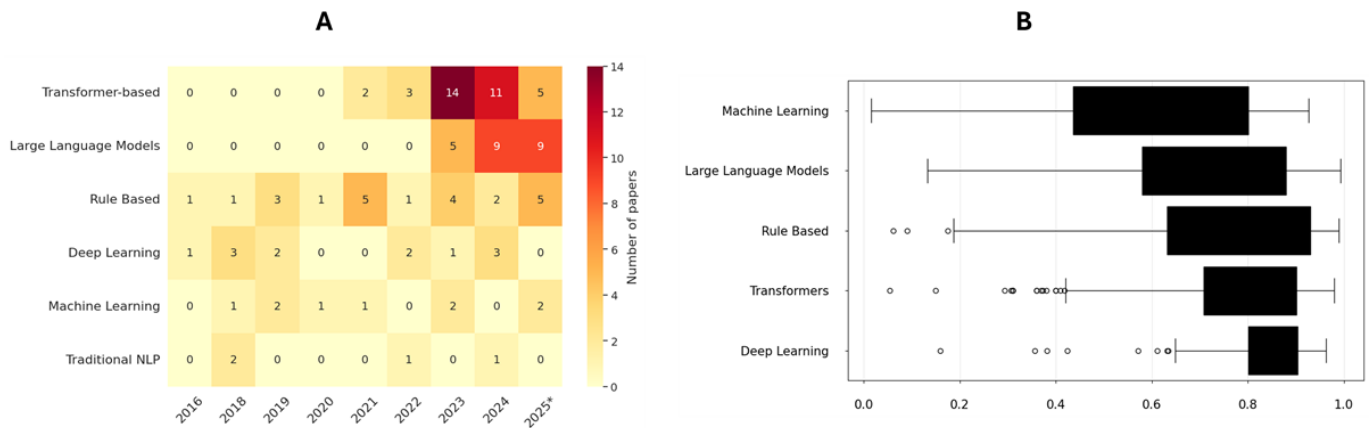
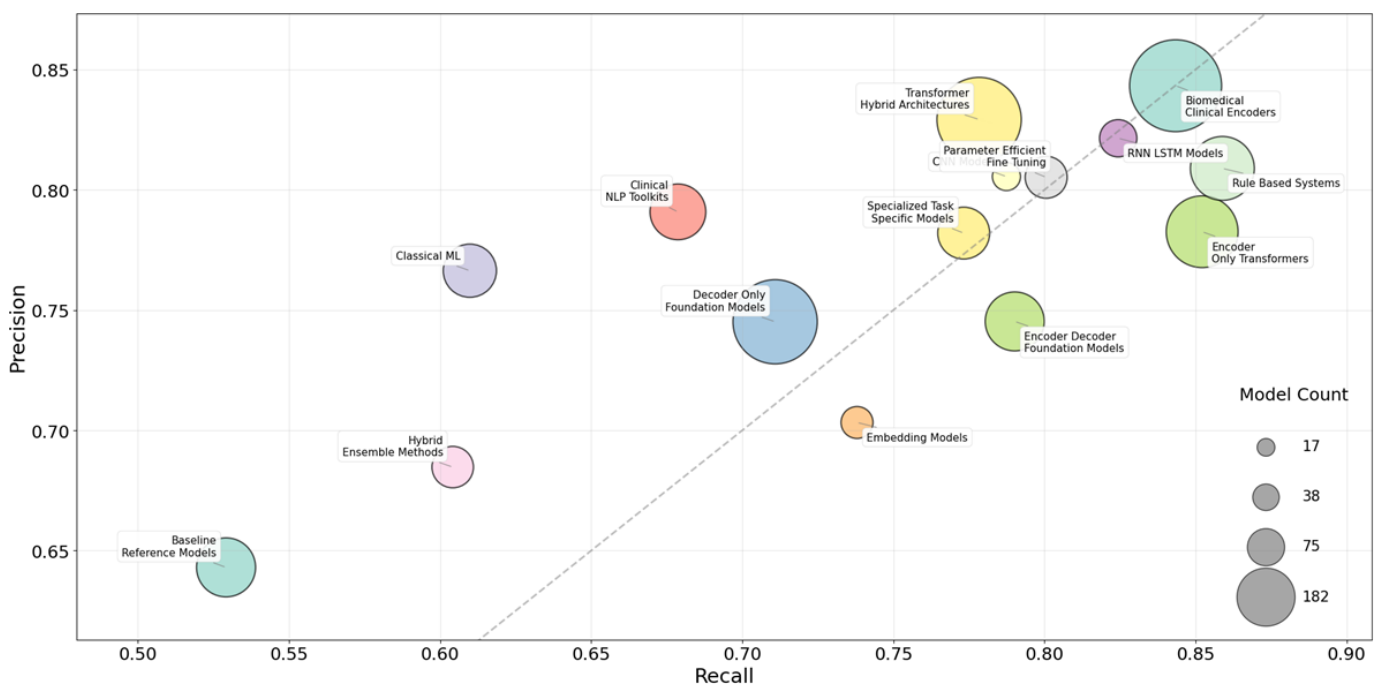


Figure 5. Bubble chart representing mean values of precision and recall for different NLP model categories. The dashed line represents the equal-precision and recall boundary. Bubble size is proportional to the number of models; only categories with ≥ 7 observations are shown. Performance metrics are from heterogeneous evaluation contexts, representative of aggregate trends, and are not individually comparable across studies. Table S4 in Multimedia Appendix 1 details NLP method categorization. LSTM: long short-term memory; ML: machine learning; NLP: natural language processing; RNN: recurrent neural network.



Overall, the results highlight that transformer-based encoder-only models such as BERT and generative models (encoder-decoders) trained on specialized clinical tasks currently represent the state of the art in terms of SDOH classification and extraction performance. LLMs (19.7% of reviewed studies, including encoder-decoders) and their capabilities represent an emerging area of research. The limited adoption of LLMs likely reflects practical and ethical concerns rather than a lack of capability. Challenges related to hallucinations, limited reproducibility, fairness and bias risks, and uncertainty around data leakage and governance (especially in clinical settings) may constrain their broader use despite strong performance in controlled evaluations. The majority of studies included in this review ($\approx 90\%$) use a cross-sectional design, focusing on extracting or classifying SDOH

information from a single point in time, providing a valuable snapshot. While useful, such approaches cannot capture the changes in SDOH factors among study subjects. Only a handful of studies [66,75,83] adopt a longitudinal approach, following individuals' records (like EHRs) over time to track changes in SDOH and their impact on health outcomes.

Model Interpretation Techniques

Model interpretation is essential for understanding not just how well a model performs, but also the basis for its decisions. A total of 82 studies reported on using interpretability techniques to gain insights into model performance. These included qualitative content analysis [29,59,131], Shapley additive explanations [47,71,163], Local Interpretable Model-Agnostic Explanations [23], attention visualization

or neuron activation analysis [54], and ablation studies [57,87]. One common approach found across studies was error analysis, which plays a key role in understanding how and why models make specific mistakes. It involves examining misclassifications or incorrect outputs to identify patterns of failure. Several studies in our review performed manual (domain expert-based) error analysis to identify model weaknesses. For instance, one paper found that their model struggled to distinguish between general and specific substance-related terms by incorrectly extracting the generic verb “smokes” as the key information when the actual target was the more specific noun “cigarettes” [77]. Another study showed that soft prompting (using trainable continuous vectors instead of manually crafted text instructions) improved the extraction of overlapping or nested SDOH concepts, especially when prompt length was tuned [58]. Error analyses were particularly prominent in studies using LLMs such as ChatGPT, Llama, and Gemini [26,31,57,59,61,155-157,159] because these models (being black-box) can produce outputs that appear coherent and convincing even when they contain mistakes. In addition, confusion matrices were used to visualize class-level misclassification patterns and highlight systematic errors [62].

Influential Studies

Analysis of the 20 influential studies published between 2018 and 2025 reveals a methodological landscape dominated by advanced NLP architectures, including BERT [60,78,80,118,130] and LLMs [57], alongside some rule-based [22,127] and hybrid approaches [27,109,168]. Notably, three-quarters (15/20) of these studies relied on manual annotation or clinician input to establish ground truth, highlighting a continued reliance on human expertise to ensure data quality and study robustness. The prevalence of multi-institutional collaborations in these studies (n=11) suggests that diverse teams and broader data sources are key drivers of research impact, likely facilitating increased generalizability and resource access. Regarding performance, the majority of studies (13/20) reported strong F_1 -scores, though results varied by task complexity. High benchmarks included concept extraction scores up to 0.9118, and relation extraction at 0.8332 [58], with Seq2Seq models achieving approximately 0.889 [80] and other systems reaching 0.86 [78]. Performance heterogeneity was evident in class-specific tasks; for example, one study reported attribute performance (eg, substance use and employment) ranging from 0.81 to 0.93 [118], while another reported a wider spread from 0.491 for non-SDOH factors to 0.774 for the best-performing class (occupation) using the same model [130]. Beyond standard metrics, one study involving EHR data reported that NLP-based methodology captured SDOH in 80.03% of cases compared with 38.17% using structured fields only [89]. Surprisingly, only 9 out of these 20 studies provided publicly accessible code. Although this represents a higher proportion than the overall set (49/142, 34.5% papers shared their code), public code availability remains limited even among the most influential studies.

Publishing Venues and Funding Sources

The most frequent publication venue was the Journal of the American Medical Informatics Association (12/142, 8.5%), followed by Journal of the American Medical Informatics Association Open (4.9%), Journal of Medical Internet Research (4.2%), and Journal of Biomedical Informatics (4.2%). Conference contributions were most often from the IEEE and AMIA Annual Symposium Proceedings. A majority of studies (110/142, 77.46%) acknowledged funding. The National Institutes of Health and its affiliated institutes (National Library of Medicine, National Center for Advancing Translational Sciences, National Cancer Institute, National Institute on Aging, and NIDA) were the leading funders (64/110, 58.2% of funded studies). Other key funders included the National Science Foundation (n=10), Agency for Healthcare Research and Quality (n=6), and international agencies such as the National Institute for Health Research and Canadian Institutes of Health Research. There is an increase in the number of studies reporting federal funding for SDOH-related NLP research in recent years. The percentage of funded studies more than doubled from 33% to 67% during 2018-2020, to 78%-83% during 2021-2025.

Discussion

Principal Findings

This systematic review synthesizes evidence from 142 studies, addressing 2 primary objectives: to characterize NLP techniques, including LLMs, used to analyze SDOH in text-based data, and to assess their reported effectiveness, gaps, and future research needs. Overall, the review reveals rapid methodological advancement and improvement in NLP performance in recent years, marked by a shift toward transformer-based architectures and emerging use of LLMs, with most applications focused on extracting or classifying individual-level SDOH from clinical and social text. Despite technical progress, the review shows that challenges related to data availability, performance reproducibility, model interpretability, and translation into real-world clinical or public health practice remain. Key findings are elaborated in the following paragraphs.

Across studies, model performance in identifying individual SDOH from unstructured text has generally improved over time, reflecting advances in representation learning and modern NLP systems' ability to capture abstract and context-dependent social concepts [171]. These advances are likely driven by both increased recognition from the research community of the role of SDOH in shaping health outcomes [172] and by the maturation of transformer-based models that perform well on complex linguistic tasks [173]. Strong reported performance, however, does not consistently translate into reproducible findings or application-ready systems [174]. Many studies relied on private, institution-specific EHR datasets, which limit independent replication and cross-site validation [175]. Although manual and clinician-led annotation was common across the reviewed studies, detailed documentation of annotation guidelines and procedures was

often lacking, making it difficult to assess consistency in how SDOH concepts were defined and applied across studies. Limited transparency around data and annotation practices also has important equity implications [176]. Reliance on proprietary EHR data favors well-resourced institutions and restricts broader participation in SDOH-focused NLP research [177].

This review also sheds light on the gap between technical development and practical implementation, and limitations on model interpretability. While many studies report promising model performance, few detail pathways for integrating NLP-derived SDOH insights into clinical workflows or public health interventions, which may be due to systemic, technical, and regulatory barriers [178]. As a result, the potential of these methods to inform equity-focused decision-making remains largely unrealized, a finding supported by prior research [176]. NLP/LLM model interpretability has received more research attention in recent years [179]. A little over half of all reviewed studies used some variant of interpretability or explainability methods, but their use was inconsistent and often inadequately discussed. This gap is relatively more critical for SDOH applications, where language may reflect complex social, cultural, and structural nuances, with uninterpretable errors disproportionately affecting marginalized groups [180]. Limited interpretability also undermines trust in systems' ability to inform care or policy, both within and outside the sphere of SDOH [181].

Future Research Directions

Based on the review, we outline several priorities for advancing NLP and LLM applications in the study of SDOH. Most of the existing work used cross-sectional designs, offering only static snapshots of social context. Future research should adopt longitudinal designs capable of capturing the dynamic nature of social risk factors, their accumulation over time, and their interaction with health trajectories [182]. Such longitudinal approaches can include continuous monitoring as well as repeated snapshot or point-in-time measures collected at predetermined intervals [183], for tracking meaningful changes in SDOH. Future work may combine unstructured narratives with structured EHR variables, imaging, and community-level data to produce more holistic and actionable models [184, 185]. Effective multimodal approaches must account for heterogeneity in SDOH-containing data [186] to maximize information capture and analytical utility. Advancing the field will also require stronger commitments to transparency and reproducibility. Establishing open benchmarks, shared tasks, and community challenges has driven NLP method development on many topics [80,128], and should be the focus of future work to promote transparency and reproducibility in SDOH-related NLP. Greater attention is also needed for structural determinants of health. Fewer than one-fifth of reviewed studies focused on structural SDOH factors such as policy, racism, or socioeconomic stratification [187]. Capturing structural determinants presents unique methodological challenges, such as patient-reported data on systemic factors being limited by awareness and subjectivity [188]. Linking individual-level health narratives with community-

and policy-level data may offer a more comprehensive approach to understanding how structural forces shape health outcomes in future research [189]. Bias and ethical considerations also remain underexplored. Few studies systematically examined algorithmic bias, despite well-documented risks of inequities in automated decision-making [190,191]. Future systems must embed safeguards for fairness, inclusivity, and accountability to avoid reinforcing disparities. Recent work has proposed actionable frameworks and techniques for mitigating bias in artificial intelligence systems [192-194]. Adopting such approaches will be essential for ensuring SDOH-focused models promote rather than perpetuate health inequities. In addition, while clinical concepts are supported by structured hierarchies such as the Unified Medical Language System [195], no equivalent framework exists for social (or nonclinical) determinants. Establishing such frameworks would reduce inconsistencies, enable interoperability, and support the creation of benchmark datasets. Finally, improving model generalizability remains critical, as models trained on single-institution EHRs often underperform in external settings, reflecting the lexical, demographic, and contextual variability of clinical narratives [174,196]. Future work should therefore move beyond single-site training by explicitly benchmarking models in cross-institution evaluations [197]. Training on pooled multi-site data and testing on held-out institutions, with performance reported by SDOH category and population subgroup, may provide actionable evidence of real-world generalizability.

Limitations

This review has several limitations. Determining whether studies were directly relevant to SDOH required subjective judgment, as many did not explicitly frame their research in SDOH terms. This introduces potential selection bias, despite a structured review process. To balance comprehensiveness and feasibility, we adhered to predefined selection criteria, focusing primarily on studies with a clear connection to SDOH. Although this approach helped maintain relevance, it may have led to the exclusion of studies with indirect but meaningful contributions to the field. Future work could explore more systematic methods, such as automated screening tools or expert consensus frameworks, to enhance the consistency and reproducibility of the selection process. We also acknowledge that our synthesis is constrained by heterogeneity across study designs, SDOH definitions, annotation schemes, datasets, and evaluation metrics, which limits direct comparability of reported performance across studies.

Conclusions

In this systematic review, we provide a comprehensive synthesis of NLP and LLM applications for SDOH by systematically surveying methodologies, data sources, evaluation practices, and translational readiness across 142 studies. In contrast to prior reviews that covered subdomains of SDOH, clinical contexts, or individual modeling techniques, ours provides a comparison of tasks, architectures, performance patterns, interpretability practices, and reproducibility considerations across the research spectrum,

spanning both individual and structural SDOH. By discussing NLP/LLM technical performance in the broader context of data accessibility, bias, generalizability, and implementation, this review advances the field beyond proof-of-concept model development toward a clearer understanding of what is required for real-world deployment. The findings highlight that, while NLP-based SDOH extraction is technically mature for certain use cases, its impact remains limited

by fragmented SDOH representations, a paucity of public benchmarks, and insufficient evaluation in clinical and public health workflows. In light of our findings, future work should prioritize public benchmarks, clearer reporting, and more inclusive datasets to advance this emerging field. Addressing these gaps has direct real-world implications by facilitating inclusive data practices, interpretable and reproducible methods, and implementation-focused evaluations.

Acknowledgments

Grammarly and Claude were used for grammar checking and Overleaf LaTeX formatting, respectively. We thank the anonymous reviewers and editor for their constructive suggestions during the review period that improved the quality of this manuscript.

Funding

YX is supported by the National Institute of Mental Health (RF1MH134649), the American Foundation for Suicide Prevention (YIG-2-133-22), and Google. AS is supported by the National Institute on Drug Abuse (R01DA057599). The content in the paper does not necessarily represent the official views of any of the institutions mentioned. The sponsors had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

Data for this study were sourced from reviewed articles referenced in this manuscript. Literature search string queries are available in [Multimedia Appendix 1](#).

Authors' Contributions

SR contributed to the conceptualization, methodology, investigation, data curation, and writing of the original draft, as well as review and editing of the manuscript. AKP contributed to data curation, visualization, and review and editing. ZZ and YC contributed to data curation and review and editing of the manuscript. ML and SD contributed to data curation, while HR contributed to resources and data curation. AS and YX contributed to conceptualization, supervision, and project administration. All authors participated in the review and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search queries, results, and categorization; summary of studies.

[\[PDF File \(Adobe File\), 403 KB-Multimedia Appendix 1\]](#)

Checklist 1

PRISMA checklist.

[\[PDF File \(Adobe File\), 124 KB-Checklist 1\]](#)

References

1. Social determinants of health. World Health Organization. 2025. URL: <https://www.who.int/health-topics/social-determinants-of-health> [Accessed 2026-03-03]
2. Social determinants of health (healthy people 2030). US Department of Health and Human Services, Office of Disease Prevention and Health Promotion. 2025. URL: <https://odphp.health.gov/healthypeople/priority-areas/social-determinants-health> [Accessed 2026-03-03]
3. Hood CM, Gennuso KP, Swain GR, Catlin BB. County health rankings: relationships between determinant factors and health outcomes. *Am J Prev Med*. Feb 2016;50(2):129-135. [doi: [10.1016/j.amepre.2015.08.024](https://doi.org/10.1016/j.amepre.2015.08.024)] [Medline: [26526164](https://pubmed.ncbi.nlm.nih.gov/26526164/)]
4. Singh GK, Daus GP, Allender M, et al. Social determinants of health in the United States: addressing major health inequality trends for the nation, 1935-2016. *Int J MCH AIDS*. 2017;6(2):139-164. [doi: [10.21106/ijma.236](https://doi.org/10.21106/ijma.236)] [Medline: [29367890](https://pubmed.ncbi.nlm.nih.gov/29367890/)]
5. Magnan S. Social determinants of health 101 for health care: five plus five. NAM perspectives discussion paper. National Academy of Medicine; 2017. URL: <https://nam.edu/social-determinants-of-health-101-for-health-care-five-plus-five/> [Accessed 2026-03-03]
6. Artiga S, Hinton E. Beyond health care: the role of social determinants in promoting health and health equity. Kaiser Family Foundation. 2018. URL: <https://www.kff.org/racial-equity-and-health-policy/beyond-health-care-the-role-of-social-determinants-in-promoting-health-and-health-equity/> [Accessed 2026-04-08]

7. Marmot M, Allen J, Boyce T, Goldblatt P, Morrison J. Health equity in England: the marmot review 10 years on. The Health Foundation; 2020. URL: <https://www.health.org.uk/reports-and-analysis/reports/health-equity-in-england-the-marmot-review-10-years-on-0> [Accessed 2026-04-08]
8. Bundy JD, Mills KT, He H, et al. Social determinants of health and premature death among adults in the USA from 1999 to 2018: a national cohort study. *Lancet Public Health*. Jun 2023;8(6):e422-e431. [doi: [10.1016/S2468-2667\(23\)00081-6](https://doi.org/10.1016/S2468-2667(23)00081-6)] [Medline: [37244672](https://pubmed.ncbi.nlm.nih.gov/37244672/)]
9. Xiao Y, Mann JJ, Chow JCC, et al. Patterns of social determinants of health and child mental health, cognition, and physical health. *JAMA Pediatr*. Dec 1, 2023;177(12):1294-1305. [doi: [10.1001/jamapediatrics.2023.4218](https://doi.org/10.1001/jamapediatrics.2023.4218)] [Medline: [37843837](https://pubmed.ncbi.nlm.nih.gov/37843837/)]
10. Xiao Y, Yip PSF, Pathak J, Mann JJ. Association of social determinants of health and vaccinations with child mental health during the COVID-19 pandemic in the US. *JAMA Psychiatry*. Jun 1, 2022;79(6):610-621. [doi: [10.1001/jamapsychiatry.2022.0818](https://doi.org/10.1001/jamapsychiatry.2022.0818)] [Medline: [35475851](https://pubmed.ncbi.nlm.nih.gov/35475851/)]
11. Xiao Y, Brown TT, Snowden LR, Chow JCC, Mann JJ. COVID-19 policies, pandemic disruptions, and changes in child mental health and sleep in the United States. *JAMA Netw Open*. Mar 1, 2023;6(3):e232716. [doi: [10.1001/jamanetworkopen.2023.2716](https://doi.org/10.1001/jamanetworkopen.2023.2716)] [Medline: [36912834](https://pubmed.ncbi.nlm.nih.gov/36912834/)]
12. Feller DJ, Bear Don't Walk Iv OJ, Zucker J, Yin MT, Gordon P, Elhadad N. Detecting social and behavioral determinants of health with structured and free-text clinical data. *Appl Clin Inform*. Jan 2020;11(1):172-181. [doi: [10.1055/s-0040-1702214](https://doi.org/10.1055/s-0040-1702214)] [Medline: [32131117](https://pubmed.ncbi.nlm.nih.gov/32131117/)]
13. Truong HP, Luke AA, Hammond G, Wadhwa RK, Reidhead M, Joynt Maddox KE. Utilization of social determinants of health ICD-10 Z-codes among hospitalized patients in the United States, 2016-2017. *Med Care*. Dec 2020;58(12):1037-1043. [doi: [10.1097/MLR.0000000000001418](https://doi.org/10.1097/MLR.0000000000001418)] [Medline: [32925453](https://pubmed.ncbi.nlm.nih.gov/32925453/)]
14. Park Y, Mulligan N, Gleize M, Kristiansen M, Bettencourt-Silva JH. Discovering associations between social determinants and health outcomes: merging knowledge graphs from literature and electronic health data. *AMIA Annu Symp Proc*. 2021;2021:940-949. [Medline: [35308956](https://pubmed.ncbi.nlm.nih.gov/35308956/)]
15. Rajwal S, Pandey AK, Han Z, Sarker A. Unveiling voices: identification of concerns in a social media breast cancer cohort via natural language processing. Presented at: Proceedings of the First Workshop on Patient-Oriented Language Processing (CL4Health) @ LREC-COLING 2024; May 20, 2024; Torino, Italia. URL: <https://aclanthology.org/2024.cl4health-1.32/> [Accessed 2026-03-03]
16. Guo Y, Rajwal S, Lakamana S, et al. Generalizable natural language processing framework for migraine reporting from social media. *AMIA Jt Summits Transl Sci Proc*. 2023;2023:261-270. [Medline: [37350878](https://pubmed.ncbi.nlm.nih.gov/37350878/)]
17. Gonzalez-Hernandez G, Sarker A, O'Connor K, Savova G. Capturing the patient's perspective: a review of advances in natural language processing of health-related text. *Yearb Med Inform*. Aug 2017;26(1):214-227. [doi: [10.15266/IY-2017-029](https://doi.org/10.15266/IY-2017-029)] [Medline: [29063568](https://pubmed.ncbi.nlm.nih.gov/29063568/)]
18. Gauthier MP, Law JH, Le LW, et al. Automating access to real-world evidence. *JTO Clin Res Rep*. Jun 2022;3(6):100340. [doi: [10.1016/j.jtocrr.2022.100340](https://doi.org/10.1016/j.jtocrr.2022.100340)] [Medline: [35719866](https://pubmed.ncbi.nlm.nih.gov/35719866/)]
19. Yu S, Le A, Feld E, et al. A natural language processing-assisted extraction system for Gleason scores: development and usability study. *JMIR Cancer*. Jul 2, 2021;7(3):e27970. [doi: [10.2196/27970](https://doi.org/10.2196/27970)] [Medline: [34255641](https://pubmed.ncbi.nlm.nih.gov/34255641/)]
20. Rouillard CJ, Nasser MA, Hu H, Roblin DW. Evaluation of a natural language processing approach to identify social determinants of health in electronic health records in a diverse community cohort. *Med Care*. Mar 1, 2022;60(3):248-255. [doi: [10.1097/MLR.0000000000001683](https://doi.org/10.1097/MLR.0000000000001683)] [Medline: [34984989](https://pubmed.ncbi.nlm.nih.gov/34984989/)]
21. Stemerman R, Arguello J, Brice J, Krishnamurthy A, Houston M, Kitzmiller R. Identification of social determinants of health using multi-label classification of electronic health record clinical notes. *JAMIA Open*. Jul 2021;4(3):o0aa069. [doi: [10.1093/jamiaopen/o0aa069](https://doi.org/10.1093/jamiaopen/o0aa069)] [Medline: [34514351](https://pubmed.ncbi.nlm.nih.gov/34514351/)]
22. Wu W, Holkeboer KJ, Kolawole TO, Carbone L, Mahmoudi E. Natural language processing to identify social determinants of health in Alzheimer's disease and related dementia from electronic health records. *Health Serv Res*. Dec 2023;58(6):1292-1302. [doi: [10.1111/1475-6773.14210](https://doi.org/10.1111/1475-6773.14210)] [Medline: [37534741](https://pubmed.ncbi.nlm.nih.gov/37534741/)]
23. Gray GM, Zirikly A, Ahumada LM, et al. Application of natural language processing to identify social needs from patient medical notes: development and assessment of a scalable, performant, and rule-based model in an integrated healthcare delivery system. *JAMIA Open*. Dec 2023;6(4):o0ad085. [doi: [10.1093/jamiaopen/o0ad085](https://doi.org/10.1093/jamiaopen/o0ad085)] [Medline: [37799347](https://pubmed.ncbi.nlm.nih.gov/37799347/)]
24. Soley N, Bentil M, Shah J, Rouhizadeh M, Taylor CO. Unveiling social determinants of health impact on adverse pregnancy outcomes through natural language processing. *Sci Rep*. Aug 9, 2025;15(1):29183. [doi: [10.1038/s41598-025-13542-x](https://doi.org/10.1038/s41598-025-13542-x)] [Medline: [40783439](https://pubmed.ncbi.nlm.nih.gov/40783439/)]
25. Lowery B, D'Acunto S, Crowe RP, Fishe JN. Using natural language processing to examine social determinants of health in prehospital pediatric encounters and associations with EMS transport decisions. *Prehosp Emerg Care*. 2023;27(2):246-251. [doi: [10.1080/10903127.2022.2072984](https://doi.org/10.1080/10903127.2022.2072984)] [Medline: [35500212](https://pubmed.ncbi.nlm.nih.gov/35500212/)]

26. Fu Y, Ramachandran GK, Dobbins NJ, et al. Extracting social determinants of health from pediatric patient notes using large language models: novel corpus and methods. In: Calzolari N, Kan MY, Hoste V, Lenci A, Sakti S, Xue N, editors. Presented at: Joint International Conference on Computational Linguistics, Language Resources and Evaluation; May 20-25, 2024:7045-7056; Torino, Italy. URL: <https://aclanthology.org/2024.lrec-main.618/> [Accessed 2026-03-03]
27. Lituiev DS, Lacar B, Pak S, Abramowitsch PL, De Marchis EH, Peterson TA. Automatic extraction of social determinants of health from medical notes of chronic lower back pain patients. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1438-1447. [doi: [10.1093/jamia/ocad054](https://doi.org/10.1093/jamia/ocad054)] [Medline: [37080559](https://pubmed.ncbi.nlm.nih.gov/37080559/)]
28. Sun S, Zack T, Williams CYK, Sushil M, Butte AJ. Topic modeling on clinical social work notes for exploring social determinants of health factors. *JAMIA Open*. Apr 2024;7(1):ooad112. [doi: [10.1093/jamiaopen/ooad112](https://doi.org/10.1093/jamiaopen/ooad112)] [Medline: [38223407](https://pubmed.ncbi.nlm.nih.gov/38223407/)]
29. Burnett SJ, Stemerma R, Innes JC, Kaisler MC, Crowe RP, Clemency BM. Social determinants of health in EMS records: a mixed-methods analysis using natural language processing and qualitative content analysis. *West J Emerg Med*. Sep 2023;24(5):878-887. [doi: [10.5811/westjem.59070](https://doi.org/10.5811/westjem.59070)] [Medline: [37788028](https://pubmed.ncbi.nlm.nih.gov/37788028/)]
30. Whitfield C, Liu Y, Anwar M. Impact of COVID-19 pandemic on social determinants of health issues of marginalized Black and Asian communities: a social media analysis empowered by natural language processing. *J Racial Ethn Health Disparities*. Jun 2025;12(3):1641-1656. [doi: [10.1007/s40615-024-01996-0](https://doi.org/10.1007/s40615-024-01996-0)] [Medline: [38625665](https://pubmed.ncbi.nlm.nih.gov/38625665/)]
31. Shah-Mohammadi F, Finkelstein J. AI-powered social determinants of health extraction from patient records: a GPT-based investigation. Presented at: 2024 IEEE First International Conference on Artificial Intelligence for Medicine, Health and Care (AIMHC); Feb 5-7, 2024:109-112; Laguna Hills, CA, USA. [doi: [10.1109/AIMHC59811.2024.00028](https://doi.org/10.1109/AIMHC59811.2024.00028)]
32. Wang X, Gupta D, Killian M, He Z. Benchmarking transformer-based models for identifying social determinants of health in clinical notes. *Proc IEEE Int Conf Healthc Inform*. Jun 2023;2023:570-574. [doi: [10.1109/ichi57859.2023.00102](https://doi.org/10.1109/ichi57859.2023.00102)] [Medline: [38239824](https://pubmed.ncbi.nlm.nih.gov/38239824/)]
33. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. Preprint posted online on Jan 16, 2013. URL: <https://arxiv.org/abs/1301.3781v3> [Accessed 2026-03-03]
34. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Preprint posted online on Aug 2, 2023. URL: <http://arxiv.org/abs/1706.03762> [Accessed 2026-03-03]
35. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Burstein J, Doran C, Solorio T, editors. Presented at: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; Jun 2-7, 2019; Minneapolis, Minnesota. [doi: [10.18653/v1/N19-1423](https://doi.org/10.18653/v1/N19-1423)]
36. Radford A, Narasimhan K, Salimans T, Sutskever I. Improving language understanding by generative pre-training. *Semantic Scholar*. 2018. URL: <https://www.semanticscholar.org/paper/Improving-Language-Understanding-by-Generative-Radford-Narasimhan/cd18800a0fe0b668a1cc19f2ec95b5003d0a5035> [Accessed 2026-04-08]
37. Rajwal S, Zhang Z, Chen Y, Rogers H, Sarker A, Xiao Y. Applications of natural language processing and large language models for social determinants of health: protocol for a systematic review. *JMIR Res Protoc*. Jan 21, 2025;14:e66094. [doi: [10.2196/66094](https://doi.org/10.2196/66094)] [Medline: [39836952](https://pubmed.ncbi.nlm.nih.gov/39836952/)]
38. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Tricco AC, Welch VA, Whiting P, Moher D, editors. BMJ*. Mar 29, 2021;372:n71. [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
39. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev*. Jan 26, 2021;10(1):39. [doi: [10.1186/s13643-020-01542-z](https://doi.org/10.1186/s13643-020-01542-z)] [Medline: [33499930](https://pubmed.ncbi.nlm.nih.gov/33499930/)]
40. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. *BMJ*. Jan 16, 2020;368:l6890. [doi: [10.1136/bmj.l6890](https://doi.org/10.1136/bmj.l6890)] [Medline: [31948937](https://pubmed.ncbi.nlm.nih.gov/31948937/)]
41. The world's #1 systematic review tool. *Covidence*. URL: <https://www.covidence.org/> [Accessed 2026-03-03]
42. Aromataris E, Fernandez R, Godfrey CM, Holly C, Khalil H, Tungpunkom P. Summarizing systematic reviews: methodological development, conduct and reporting of an umbrella review approach. *Int J Evid Based Healthc*. Sep 2015;13(3):132-140. [doi: [10.1097/XEB.0000000000000055](https://doi.org/10.1097/XEB.0000000000000055)] [Medline: [26360830](https://pubmed.ncbi.nlm.nih.gov/26360830/)]
43. Feller DJ, Zucker J, Yin MT, Gordon P, Elhadad N. Using clinical notes and natural language processing for automated HIV risk assessment. *JAIDS*. 2018;77(2):160-166. [doi: [10.1097/QAI.0000000000001580](https://doi.org/10.1097/QAI.0000000000001580)]
44. Bako AT, Walter-McCabe H, Kasthurirathne SN, Halverson PK, Vest JR. Reasons for social work referrals in an urban safety-net population: a natural language processing and market basket analysis approach. *J Soc Serv Res*. May 4, 2021;47(3):414-425. [doi: [10.1080/01488376.2020.1817834](https://doi.org/10.1080/01488376.2020.1817834)]
45. Zhu J, Yalamanchi N, Jin R, Kenne DR, Phan N. Investigating COVID-19's impact on mental health: trend and thematic analysis of reddit users' discourse. *J Med Internet Res*. Jul 12, 2023;25:e46867. [doi: [10.2196/46867](https://doi.org/10.2196/46867)] [Medline: [37436793](https://pubmed.ncbi.nlm.nih.gov/37436793/)]

46. Morrow D, Zamora-Resendiz R, Beckham JC, et al. A case for developing domain-specific vocabularies for extracting suicide factors from healthcare notes. *J Psychiatr Res*. Jul 2022;151:328-338. [doi: [10.1016/j.jpsychires.2022.04.009](https://doi.org/10.1016/j.jpsychires.2022.04.009)] [Medline: [35533516](https://pubmed.ncbi.nlm.nih.gov/35533516/)]
47. Kessler RC, Bauer MS, Bishop TM, et al. Evaluation of a model to target high-risk psychiatric inpatients for an intensive postdischarge suicide prevention intervention. *JAMA Psychiatry*. Mar 1, 2023;80(3):230-240. [doi: [10.1001/jamapsychiatry.2022.4634](https://doi.org/10.1001/jamapsychiatry.2022.4634)] [Medline: [36652267](https://pubmed.ncbi.nlm.nih.gov/36652267/)]
48. Wang T, Codling D, Bhugra D, et al. Unraveling ethnic disparities in antipsychotic prescribing among patients with psychosis: a retrospective cohort study based on electronic clinical records. *Schizophr Res*. Oct 2023;260:168-179. [doi: [10.1016/j.schres.2023.08.024](https://doi.org/10.1016/j.schres.2023.08.024)] [Medline: [37669576](https://pubmed.ncbi.nlm.nih.gov/37669576/)]
49. Chilman N, Laporte D, Dorrington S, et al. Understanding social and clinical associations with unemployment for people with schizophrenia and bipolar disorders: large-scale health records study. *Soc Psychiatry Psychiatr Epidemiol*. Oct 2024;59(10):1709-1719. [doi: [10.1007/s00127-024-02620-6](https://doi.org/10.1007/s00127-024-02620-6)] [Medline: [38378812](https://pubmed.ncbi.nlm.nih.gov/38378812/)]
50. Wang L, Guo Y, Yin X, Wang Y, Tong R. Exploring the determinants of health-promoting behaviors among miners: a text mining and meta-analysis. *Appl Psychol Health Well Being*. Feb 2024;16(1):3-24. [doi: [10.1111/aphw.12465](https://doi.org/10.1111/aphw.12465)] [Medline: [37339782](https://pubmed.ncbi.nlm.nih.gov/37339782/)]
51. Goldstein EV, Bailey EV, Wilson FA. Poverty and suicidal ideation among hispanic mental health care patients leading up to the COVID-19 pandemic. *Hispanic Health Care Int*. Mar 2024;22(1):6-10. [doi: [10.1177/15404153231181110](https://doi.org/10.1177/15404153231181110)]
52. Xie F, Wang S, Viveros L, et al. Using natural language processing to identify the status of homelessness and housing instability among serious illness patients from clinical notes in an integrated healthcare system. *JAMIA Open*. Jul 4, 2023;6(3). [doi: [10.1093/jamiaopen/ooad082](https://doi.org/10.1093/jamiaopen/ooad082)]
53. Harris DR, Fu S, Wen A, et al. The ENACT network is acting on housing instability and the unhoused using the open health natural language processing toolkit. *J Clin Transl Sci*. 2024;8(1):e98. [doi: [10.1017/cts.2024.543](https://doi.org/10.1017/cts.2024.543)] [Medline: [39655040](https://pubmed.ncbi.nlm.nih.gov/39655040/)]
54. Ashayeri M, Abbasabadi N. Unraveling energy justice in NYC urban buildings through social media sentiment analysis and transformer deep learning. *Energy Build*. Mar 2024;306:113914. [doi: [10.1016/j.enbuild.2024.113914](https://doi.org/10.1016/j.enbuild.2024.113914)]
55. Peng C, Yang X, Yu Z, Bian J, Hogan WR, Wu Y. Clinical concept and relation extraction using prompt-based machine reading comprehension. *J Am Med Inform Assoc*. Aug 18, 2023;30(9):1486-1493. [doi: [10.1093/jamia/ocad107](https://doi.org/10.1093/jamia/ocad107)]
56. Hobensack M, Song J, Oh S, et al. Social risk factors are associated with risk for hospitalization in home health care: a natural language processing study. *J Am Med Dir Assoc*. Dec 2023;24(12):1874-1880. [doi: [10.1016/j.jamda.2023.06.031](https://doi.org/10.1016/j.jamda.2023.06.031)]
57. Guevara M, Chen S, Thomas S, et al. Large language models to identify social determinants of health in electronic health records. *NPJ Digit Med*. Jan 11, 2024;7(1):6. [doi: [10.1038/s41746-023-00970-0](https://doi.org/10.1038/s41746-023-00970-0)] [Medline: [38200151](https://pubmed.ncbi.nlm.nih.gov/38200151/)]
58. Peng C, Yang X, Smith KE, et al. Model tuning or prompt tuning? a study of large language models for clinical concept and relation extraction. *J Biomed Inform*. May 2024;153:104630. [doi: [10.1016/j.jbi.2024.104630](https://doi.org/10.1016/j.jbi.2024.104630)] [Medline: [38548007](https://pubmed.ncbi.nlm.nih.gov/38548007/)]
59. Hu Z, Zhang Y, Rossi R, Yu T, Kim S, Pan S. Are Large Language Models Capable of Causal Reasoning for Sensing Data Analysis. Association for Computing Machinery, Inc; 2024:24-29. [doi: [10.1145/3662006.3662064](https://doi.org/10.1145/3662006.3662064)]
60. Savcicens G, Eliassi-Rad T, Hansen LK, et al. Using sequences of life-events to predict human lives. *Nat Comput Sci*. Jan 2024;4(1):43-56. [doi: [10.1038/s43588-023-00573-5](https://doi.org/10.1038/s43588-023-00573-5)] [Medline: [38177491](https://pubmed.ncbi.nlm.nih.gov/38177491/)]
61. Luo X, Tahabi FM, Marc T, Hauer LA, Storey S. Zero-shot learning to extract assessment criteria and medical services from the preventive healthcare guidelines using large language models. *J Am Med Inform Assoc*. Aug 1, 2024;31(8):1743-1753. [doi: [10.1093/jamia/ocae145](https://doi.org/10.1093/jamia/ocae145)]
62. Kwon S, Wang X, Liu W, et al. ODD: a benchmark dataset for the natural language processing based opioid related aberrant behavior detection. Presented at: Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics; Jun 16-21, 2024:4338-4359; Mexico City, Mexico. [doi: [10.18653/v1/2024.naacl-long.244](https://doi.org/10.18653/v1/2024.naacl-long.244)]
63. Vest JR, Mazurenko O. Non-response bias in social risk factor screening among adult emergency department patients. *J Med Syst*. Jul 22, 2023;47(1):78. [doi: [10.1007/s10916-023-01975-8](https://doi.org/10.1007/s10916-023-01975-8)] [Medline: [37480515](https://pubmed.ncbi.nlm.nih.gov/37480515/)]
64. Harris DR, Anthony N, Quesinberry D, Delcher C. Evidence of housing instability identified by addresses, clinical notes, and diagnostic codes in a real-world population with substance use disorders. *J Clin Transl Sci*. 2023;7(1):e196. [doi: [10.1017/cts.2023.626](https://doi.org/10.1017/cts.2023.626)] [Medline: [37771412](https://pubmed.ncbi.nlm.nih.gov/37771412/)]
65. Cybulski L, Chilman N, Jewell A, et al. Improving our understanding of the social determinants of mental health: a data linkage study of mental health records and the 2011 UK census. *BMJ Open*. Jan 29, 2024;14(1):e073582. [doi: [10.1136/bmjopen-2023-073582](https://doi.org/10.1136/bmjopen-2023-073582)] [Medline: [38286672](https://pubmed.ncbi.nlm.nih.gov/38286672/)]
66. Xie K, Ojemann WKS, Gallagher RS, et al. Disparities in seizure outcomes revealed by large language models. *J Am Med Inform Assoc*. May 20, 2024;31(6):1348-1355. [doi: [10.1093/jamia/ocae047](https://doi.org/10.1093/jamia/ocae047)]

67. Shafer PR, Davis A, Clark JA. Finding social need-les in a haystack: ascertaining social needs of medicare patients recorded in the notes of care managers. *BMC Health Serv Res*. Dec 12, 2023;23(1):1400. [doi: [10.1186/s12913-023-10446-2](https://doi.org/10.1186/s12913-023-10446-2)] [Medline: [38087286](https://pubmed.ncbi.nlm.nih.gov/38087286/)]
68. Miranda O, Fan P, Qi X, et al. DeepBiomarker2: prediction of alcohol and substance use disorder risk in post-traumatic stress disorder patients using electronic medical records and multiple social determinants of health. *J Pers Med*. Jan 14, 2024;14(1):2024. [doi: [10.3390/jpm14010094](https://doi.org/10.3390/jpm14010094)] [Medline: [38248795](https://pubmed.ncbi.nlm.nih.gov/38248795/)]
69. Yu Z, Peng C, Yang X, et al. Identifying social determinants of health from clinical narratives: a study of performance, documentation ratio, and potential bias. *J Biomed Inform*. May 2024;153:104642. [doi: [10.1016/j.jbi.2024.104642](https://doi.org/10.1016/j.jbi.2024.104642)] [Medline: [38621641](https://pubmed.ncbi.nlm.nih.gov/38621641/)]
70. Magoc T, Allen KS, McDonnell C, et al. Generalizability and portability of natural language processing system to extract individual social risk factors. *Int J Med Inform*. Sep 2023;177:105115. [doi: [10.1016/j.ijmedinf.2023.105115](https://doi.org/10.1016/j.ijmedinf.2023.105115)] [Medline: [37302362](https://pubmed.ncbi.nlm.nih.gov/37302362/)]
71. Huang T, Socrates V, Gilson A, et al. Identifying incarceration status in the electronic health record using large language models in emergency department settings. *J Clin Transl Sci*. 2024;8(1):e53. [doi: [10.1017/cts.2024.496](https://doi.org/10.1017/cts.2024.496)] [Medline: [38544748](https://pubmed.ncbi.nlm.nih.gov/38544748/)]
72. Bhanot K, Erickson JS, Bennett KP. MortalityMinder: visualization and AI interpretations of social determinants of premature mortality in the United States. *Information*. 2024;15(5):254. [doi: [10.3390/info15050254](https://doi.org/10.3390/info15050254)]
73. Mehta S, Lyles CR, Rubinsky AD, et al. Social determinants of health documentation in structured and unstructured clinical data of patients with diabetes: comparative analysis. *JMIR Med Inform*. Aug 22, 2023;11:e46159. [doi: [10.2196/46159](https://doi.org/10.2196/46159)] [Medline: [37621203](https://pubmed.ncbi.nlm.nih.gov/37621203/)]
74. Richie R, Ruiz VM, Han S, Shi L, Tsui FR. Extracting social determinants of health events with transformer-based multitask, multilabel named entity recognition. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1379-1388. [doi: [10.1093/jamia/ocad046](https://doi.org/10.1093/jamia/ocad046)] [Medline: [37002953](https://pubmed.ncbi.nlm.nih.gov/37002953/)]
75. Chapman AB, Scharfstein DO, Montgomery AE, et al. Using natural language processing to study homelessness longitudinally with electronic health record data subject to irregular observations. *AMIA Annu Symp Proc*. 2023;2023:894-903. [Medline: [38222404](https://pubmed.ncbi.nlm.nih.gov/38222404/)]
76. Ravichandran U, Jungst D, Kwan E. Implementing an NLP Tool to Address SDOH Needs. Institute of Electrical and Electronics Engineers Inc; 2023. [doi: [10.1109/ICHI57859.2023.00091](https://doi.org/10.1109/ICHI57859.2023.00091)]
77. Zhao X, Rios A. A marker-based neural network system for extracting social determinants of health. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1398-1407. [doi: [10.1093/jamia/ocad041](https://doi.org/10.1093/jamia/ocad041)] [Medline: [37011635](https://pubmed.ncbi.nlm.nih.gov/37011635/)]
78. Lybarger K, Dobbins NJ, Long R, et al. Leveraging natural language processing to augment structured social determinants of health data in the electronic health record. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1389-1397. [doi: [10.1093/jamia/ocad073](https://doi.org/10.1093/jamia/ocad073)] [Medline: [37130345](https://pubmed.ncbi.nlm.nih.gov/37130345/)]
79. Stewart de Ramirez S, Shallat J, McClure K, Foulger R, Barenblat L. Screening for social determinants of health: active and passive information retrieval methods. *Popul Health Manag*. Dec 2022;25(6):781-788. [doi: [10.1089/pop.2022.0228](https://doi.org/10.1089/pop.2022.0228)]
80. Lybarger K, Yetisgen M, Uzuner Ö. The 2022 n2c2/UW shared task on extracting social determinants of health. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1367-1378. [doi: [10.1093/jamia/ocad012](https://doi.org/10.1093/jamia/ocad012)] [Medline: [36795066](https://pubmed.ncbi.nlm.nih.gov/36795066/)]
81. Wang S, Dang Y, Sun Z, et al. An NLP approach to identify SDOH-related circumstance and suicide crisis from death investigation narratives. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1408-1417. [doi: [10.1093/jamia/ocad068](https://doi.org/10.1093/jamia/ocad068)]
82. Sajdeya R, Mardini MT, Tighe PJ, et al. Developing and validating a natural language processing algorithm to extract preoperative cannabis use status documentation from unstructured narrative clinical notes. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1418-1428. [doi: [10.1093/jamia/ocad080](https://doi.org/10.1093/jamia/ocad080)] [Medline: [37178155](https://pubmed.ncbi.nlm.nih.gov/37178155/)]
83. Chapman AB, Cordasco K, Chassman S, et al. Assessing longitudinal housing status using electronic health record data: a comparison of natural language processing, structured data, and patient-reported history. *Front Artif Intell*. 2023;6. [doi: [10.3389/frai.2023.1187501](https://doi.org/10.3389/frai.2023.1187501)]
84. Lelkes AD, Loreaux E, Schuster T, Chen MJ, Rajkomar A. SDOH-NLI: a dataset for inferring social determinants of health from clinical notes. Presented at: Findings of the Association for Computational Linguistics; Dec 6-10, 2023:4789-4798; Singapore. [doi: [10.18653/v1/2023.findings-emnlp.317](https://doi.org/10.18653/v1/2023.findings-emnlp.317)]
85. Romanowski B, Ben Abacha A, Fan Y. Extracting social determinants of health from clinical note text with classification and sequence-to-sequence approaches. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1448-1455. [doi: [10.1093/jamia/ocad071](https://doi.org/10.1093/jamia/ocad071)] [Medline: [37100768](https://pubmed.ncbi.nlm.nih.gov/37100768/)]
86. Ramachandran GK, Fu Y, Han B, et al. Prompt-based extraction of social determinants of health using few-shot learning. Presented at: Proceedings of the 5th Clinical Natural Language Processing Workshop; Jul 14, 2023; Toronto, Canada. [doi: [10.18653/v1/2023.clinicalnlp-1.41](https://doi.org/10.18653/v1/2023.clinicalnlp-1.41)]

87. Bashir SR, Raza S, Kocaman V, Qamar U. Clinical application of detecting COVID-19 risks: a natural language processing approach. *Viruses*. Dec 11, 2022;14(12):2761. [doi: [10.3390/v14122761](https://doi.org/10.3390/v14122761)] [Medline: [36560764](https://pubmed.ncbi.nlm.nih.gov/36560764/)]
88. Lee K, Han S, Suh HS. Early impact of COVID-19 social distancing on social determinants of health and their effects on mental health and quality of life of Korean undergraduate students. *Front Public Health*. 2023;11. [doi: [10.3389/fpubh.2023.1197143](https://doi.org/10.3389/fpubh.2023.1197143)]
89. Mitra A, Pradhan R, Melamed RD, et al. Associations between natural language processing-enriched social determinants of health and suicide death among US veterans. *JAMA Netw Open*. Mar 1, 2023;6(3):e233079. [doi: [10.1001/jamanetworkopen.2023.3079](https://doi.org/10.1001/jamanetworkopen.2023.3079)] [Medline: [36920391](https://pubmed.ncbi.nlm.nih.gov/36920391/)]
90. Yao Z, Tsai J, Liu W, et al. Automated identification of eviction status from electronic health record notes. *J Am Med Inform Assoc*. Jul 19, 2023;30(8):1429-1437. [doi: [10.1093/jamia/ocad081](https://doi.org/10.1093/jamia/ocad081)] [Medline: [37203429](https://pubmed.ncbi.nlm.nih.gov/37203429/)]
91. Allen KS, Hood DR, Cummins J, Kasturi S, Mendonca EA, Vest JR. Natural language processing-driven state machines to extract social factors from unstructured clinical documentation. *JAMIA Open*. Apr 6, 2023;6(2). [doi: [10.1093/jamiaopen/ooad024](https://doi.org/10.1093/jamiaopen/ooad024)]
92. Cascalheira CJ, Chapagain S, Flinn RE, et al. Predicting linguistically sophisticated social determinants of health disparities with neural networks: the case of LGBTQ+ minority stress. Presented at: 2023 IEEE International Conference on Big Data (BigData); Dec 15-18, 2023; Sorrento, Italy. [doi: [10.1109/BigData59044.2023.10386882](https://doi.org/10.1109/BigData59044.2023.10386882)]
93. Chilman N, Song X, Roberts A, et al. Text mining occupations from the mental health electronic health record: a natural language processing approach using records from the clinical record interactive search (CRIS) platform in south London, UK. *BMJ Open*. Mar 2021;11(3):e042274. [doi: [10.1136/bmjopen-2020-042274](https://doi.org/10.1136/bmjopen-2020-042274)]
94. Chapman AB, Jones A, Kelley AT, et al. ReHouSED: a novel measurement of veteran housing stability using natural language processing. *J Biomed Inform*. Oct 2021;122:103903. [doi: [10.1016/j.jbi.2021.103903](https://doi.org/10.1016/j.jbi.2021.103903)] [Medline: [34474188](https://pubmed.ncbi.nlm.nih.gov/34474188/)]
95. Mitra A, Ahsan H, Li W, et al. Risk factors associated with nonfatal opioid overdose leading to intensive care unit admission: a cross-sectional study. *JMIR Med Inform*. 2021;9(11):e32851. [doi: [10.2196/32851](https://doi.org/10.2196/32851)]
96. Sy B, Wassil M, Connelly H, Hassan A. Behavioral predictive analytics towards personalization for self-management: a use case on linking health-related social needs. *SN Comput Sci*. 2022;3(3):237. [doi: [10.1007/s42979-022-01092-2](https://doi.org/10.1007/s42979-022-01092-2)] [Medline: [35493988](https://pubmed.ncbi.nlm.nih.gov/35493988/)]
97. Teng A, Wilcox A. Simplified data science approach to extract social and behavioural determinants: a retrospective chart review. *BMJ Open*. Jan 18, 2022;12(1):e048397. [doi: [10.1136/bmjopen-2020-048397](https://doi.org/10.1136/bmjopen-2020-048397)] [Medline: [35042703](https://pubmed.ncbi.nlm.nih.gov/35042703/)]
98. Hatf E, Singh Deol G, Rouhizadeh M, et al. Measuring the value of a practical text mining approach to identify patients with housing issues in the free-text notes in electronic health record: findings of a retrospective cohort study. *Front Public Health*. 2021;9:697501. [doi: [10.3389/fpubh.2021.697501](https://doi.org/10.3389/fpubh.2021.697501)] [Medline: [34513783](https://pubmed.ncbi.nlm.nih.gov/34513783/)]
99. Hatf E, Rouhizadeh M, Nau C, et al. Development and assessment of a natural language processing model to identify residential instability in electronic health records' unstructured data: a comparison of 3 integrated healthcare delivery systems. *JAMIA Open*. Apr 2022;5(1):ooac006. [doi: [10.1093/jamiaopen/ooac006](https://doi.org/10.1093/jamiaopen/ooac006)] [Medline: [35224458](https://pubmed.ncbi.nlm.nih.gov/35224458/)]
100. Raza S, Schwartz B. Detecting biomedical named entities in covid-19 texts. Presented at: Proceedings of the 1st Workshop on Healthcare AI and COVID-19; Jul 22, 2022:117-126; Hawaii. URL: <https://proceedings.mlr.press/v184/raza22a.html>
101. Datar S, Lindemann EA, Silverman G, et al. Identifying Mentions of Life Stressors in Clinical Notes. Institute of Electrical and Electronics Engineers Inc; 2021:153-160. [doi: [10.1109/ICHI52183.2021.00033](https://doi.org/10.1109/ICHI52183.2021.00033)]
102. Reeves RM, Christensen L, Brown JR, et al. Adaptation of an NLP system to a new healthcare environment to identify social determinants of health. *J Biomed Inform*. Aug 2021;120:103851. [doi: [10.1016/j.jbi.2021.103851](https://doi.org/10.1016/j.jbi.2021.103851)] [Medline: [34174396](https://pubmed.ncbi.nlm.nih.gov/34174396/)]
103. Olusanya OA, Ammar N, Davis RL, Bednarczyk RA, Shaban-Nejad A. A digital personal health library for enabling precision health promotion to prevent human papilloma virus-associated cancers. *Front Digit Health*. 2021;3:683161. [doi: [10.3389/fdgh.2021.683161](https://doi.org/10.3389/fdgh.2021.683161)] [Medline: [34713154](https://pubmed.ncbi.nlm.nih.gov/34713154/)]
104. Poulsen MN, Freda PJ, Troiani V, Davoudi A, Mowery DL. Classifying characteristics of opioid use disorder from hospital discharge summaries using natural language processing. *Front Public Health*. 2022;10:850619. [doi: [10.3389/fpubh.2022.850619](https://doi.org/10.3389/fpubh.2022.850619)] [Medline: [35615042](https://pubmed.ncbi.nlm.nih.gov/35615042/)]
105. Dorr DA, Quiñones AR, King T, Wei MY, White K, Bejan CA. Prediction of future health care utilization through note-extracted psychosocial factors. *Med Care*. Aug 1, 2022;60(8):570-578. [doi: [10.1097/MLR.0000000000001742](https://doi.org/10.1097/MLR.0000000000001742)] [Medline: [35658116](https://pubmed.ncbi.nlm.nih.gov/35658116/)]
106. Yu Z, Yang X, Dang C, et al. A study of social and behavioral determinants of health in lung cancer patients using transformers-based natural language processing models. *AMIA Annu Symp Proc*. 2021;2021:1225-1233. [Medline: [35309014](https://pubmed.ncbi.nlm.nih.gov/35309014/)]
107. Tsui FR, Shi L, Ruiz V, et al. Natural language processing and machine learning of electronic health records for prediction of first-time suicide attempts. *JAMIA Open*. Jan 2021;4(1):ooab011. [doi: [10.1093/jamiaopen/ooab011](https://doi.org/10.1093/jamiaopen/ooab011)] [Medline: [33758800](https://pubmed.ncbi.nlm.nih.gov/33758800/)]

108. Yu Z, Yang X, Guo Y, Bian J, Wu Y. Assessing the documentation of social determinants of health for lung cancer patients in clinical narratives. *Front Public Health*. 2022;10:778463. [doi: [10.3389/fpubh.2022.778463](https://doi.org/10.3389/fpubh.2022.778463)] [Medline: [35419333](https://pubmed.ncbi.nlm.nih.gov/35419333/)]
109. Hatéf E, Rouhizadeh M, Tia I, et al. Assessing the availability of data on social and behavioral determinants in structured and unstructured electronic health records: a retrospective analysis of a multilevel health care system. *JMIR Med Inform*. Aug 2, 2019;7(3):e13802. [doi: [10.2196/13802](https://doi.org/10.2196/13802)] [Medline: [31376277](https://pubmed.ncbi.nlm.nih.gov/31376277/)]
110. Bettencourt-Silva JH, Mulligan N, Sbdio M, et al. Discovering New Social Determinants of Health Concepts from Unstructured Data: Framework and Evaluation. IOS Press; 2020:173-177. [doi: [10.3233/SHTI200145](https://doi.org/10.3233/SHTI200145)]
111. Bucher BT, Shi J, Pettit RJ, Ferraro J, Chapman WW, Gundlapalli A. Determination of marital status of patients from structured and unstructured electronic healthcare data. *AMIA Annu Symp Proc*. 2019;2019:267-274. [Medline: [32308819](https://pubmed.ncbi.nlm.nih.gov/32308819/)]
112. Sebestyén V, Domokos E, Abonyi J. Focal points for sustainable development strategies-text mining-based comparative analysis of voluntary national reviews. *J Environ Manage*. Jun 1, 2020;263:110414. [doi: [10.1016/j.jenvman.2020.110414](https://doi.org/10.1016/j.jenvman.2020.110414)] [Medline: [32174539](https://pubmed.ncbi.nlm.nih.gov/32174539/)]
113. Cho SM, Park CU, Song M. The evolution of social health research topics: a data-driven analysis. *Soc Sci Med*. Nov 2020;265:113299. [doi: [10.1016/j.socscimed.2020.113299](https://doi.org/10.1016/j.socscimed.2020.113299)] [Medline: [32905964](https://pubmed.ncbi.nlm.nih.gov/32905964/)]
114. Conway M, Keyhani S, Christensen L, et al. Moonstone: a novel natural language processing system for inferring social risk from clinical narratives. *J Biomed Semant*. Dec 2019;10(1):2019. [doi: [10.1186/s13326-019-0198-0](https://doi.org/10.1186/s13326-019-0198-0)]
115. Dorr D, Bejan CA, Pizzimenti C, Singh S, Storer M, Quinones A. Identifying Patients with Significant Problems Related to Social Determinants of Health with Natural Language Processing. IOS Press; 2019:1456-1457. [doi: [10.3233/SHTI190482](https://doi.org/10.3233/SHTI190482)]
116. Du X, Bian J, Prosperi M. An Operational Deep Learning Pipeline for Classifying Life Events from Individual Tweets. Springer Verlag; 2019:54-66. [doi: [10.1007/978-3-030-11680-4_7](https://doi.org/10.1007/978-3-030-11680-4_7)]
117. DuBay DA, Su Z, Morinelli TA, et al. Development and future deployment of a 5 years allograft survival model for kidney transplantation. *Nephrology (Carlton)*. Aug 2019;24(8):855-862. [doi: [10.1111/nep.13488](https://doi.org/10.1111/nep.13488)] [Medline: [30198104](https://pubmed.ncbi.nlm.nih.gov/30198104/)]
118. Lybarger K, Ostendorf M, Yetisgen M. Annotating social determinants of health using active learning, and characterizing determinants using neural event extraction. *J Biomed Inform*. Jan 2021;113:103631. [doi: [10.1016/j.jbi.2020.103631](https://doi.org/10.1016/j.jbi.2020.103631)]
119. Zhu VJ, Lenert LA, Bunnell BE, Obeid JS, Jefferson M, Halbert CH. Automatically identifying social isolation from clinical narratives for patients with prostate cancer. *BMC Med Inform Decis Mak*. Mar 14, 2019;19(1):43. [doi: [10.1186/s12911-019-0795-y](https://doi.org/10.1186/s12911-019-0795-y)] [Medline: [30871518](https://pubmed.ncbi.nlm.nih.gov/30871518/)]
120. Bettencourt-Silva JH, Mulligan N, Jochim C, et al. Exploring the Social Drivers of Health during a Pandemic: Leveraging Knowledge Graphs and Population Trends in COVID-19. IOS Press BV; 2020:6-11. [doi: [10.3233/SHTI200684](https://doi.org/10.3233/SHTI200684)]
121. Zhang X, Bellolio MF, Medrano-Gracia P, Werys K, Yang S, Mahajan P. Use of natural language processing to improve predictive models for imaging utilization in children presenting to the emergency department. *BMC Med Inform Decis Mak*. Dec 30, 2019;19(1):287. [doi: [10.1186/s12911-019-1006-6](https://doi.org/10.1186/s12911-019-1006-6)] [Medline: [31888609](https://pubmed.ncbi.nlm.nih.gov/31888609/)]
122. Feller DJ, Zucker J, Don't Walk OB 4th, et al. Towards the inference of social and behavioral determinants of sexual health: development of a gold-standard corpus with semi-supervised learning. *AMIA Annu Symp Proc*. 2018;2018:422-429. [Medline: [30815082](https://pubmed.ncbi.nlm.nih.gov/30815082/)]
123. Bejan CA, Angiolillo J, Conway D, et al. Mining 100 million notes to find homelessness and adverse childhood experiences: 2 case studies of rare and severe social determinants of health in electronic health records. *J Am Med Inform Assoc*. Jan 1, 2018;25(1):61-71. [doi: [10.1093/jamia/ocx059](https://doi.org/10.1093/jamia/ocx059)]
124. Lindemann EA, Chen ES, Wang Y, Skube SJ, Melton GB. Representation of social history factors across age groups: a topic analysis of free-text social documentation. *AMIA Annu Symp Proc*. 2017;2017:1169-1178. [Medline: [29854185](https://pubmed.ncbi.nlm.nih.gov/29854185/)]
125. Velupillai S, Mowery DL, Conway M, Hurdle J, Kiouss B. Vocabulary development to support information extraction of substance abuse from psychiatry notes. Presented at: Proceedings of the 15th Workshop on Biomedical Natural Language Processing; Aug 12, 2016; Berlin, Germany. [doi: [10.18653/v1/W16-2912](https://doi.org/10.18653/v1/W16-2912)]
126. Relia K, Akbari M, Duncan D, Chunara R. Socio-spatial self-organizing maps: using social media to assess relevant geographies for exposure to social processes. *Proc ACM Hum Comput Interact*. Nov 2018;2:145. [doi: [10.1145/3274414](https://doi.org/10.1145/3274414)] [Medline: [30957076](https://pubmed.ncbi.nlm.nih.gov/30957076/)]
127. Navathe AS, Zhong F, Lei VJ, et al. Hospital readmission and social risk factors identified from physician notes. *Health Serv Res*. Apr 2018;53(2):1110-1136. [doi: [10.1111/1475-6773.12670](https://doi.org/10.1111/1475-6773.12670)] [Medline: [28295260](https://pubmed.ncbi.nlm.nih.gov/28295260/)]
128. Peng C, Yang X, Chen A, et al. Generative large language models are all-purpose text analytics engines: text-to-text learning is all your need. *J Am Med Inform Assoc*. Sep 1, 2024;31(9):1892-1903. [doi: [10.1093/jamia/ocae078](https://doi.org/10.1093/jamia/ocae078)] [Medline: [38630580](https://pubmed.ncbi.nlm.nih.gov/38630580/)]
129. Gleize M, Mulligan N, Bettencourt-Silva JH. Social Determinant Trends of COVID-19: An Analysis Using Knowledge Graphs from Published Evidence and Online Trends. 2021:744-748. [doi: [10.3233/SHTI210271](https://doi.org/10.3233/SHTI210271)]

130. Han S, Zhang RF, Shi L, et al. Classifying social determinants of health from unstructured electronic health records using deep learning-based natural language processing. *J Biomed Inform.* Mar 2022;127:103984. [doi: [10.1016/j.jbi.2021.103984](https://doi.org/10.1016/j.jbi.2021.103984)]
131. Suda-King C, Winch L, Tucker JM, Zuehlke AD, Hunter C, Simmons JM. Representation of social determinants of health terminology in medical subject headings: impact of added terms. *J Am Med Inform Assoc.* Nov 1, 2024;31(11):2595-2604. [doi: [10.1093/jamia/ocae191](https://doi.org/10.1093/jamia/ocae191)] [Medline: [39047296](https://pubmed.ncbi.nlm.nih.gov/39047296/)]
132. Bejan CA, Ripperger M, Wilimitis D, et al. Improving ascertainment of suicidal ideation and suicide attempt with natural language processing. *Sci Rep.* Sep 7, 2022;12(1):15146. [doi: [10.1038/s41598-022-19358-3](https://doi.org/10.1038/s41598-022-19358-3)] [Medline: [36071081](https://pubmed.ncbi.nlm.nih.gov/36071081/)]
133. Chandler MT, Cai TR, Santacroce L, Ulysse S, Liao KP, Feldman CH. Classifying individuals with rheumatic conditions as financially insecure using electronic health record data and natural language processing: algorithm derivation and validation. *ACR OPEN Rheumatol.* Aug 2024;6(8):481-488. [doi: [10.1002/acr2.11675](https://doi.org/10.1002/acr2.11675)] [Medline: [38747148](https://pubmed.ncbi.nlm.nih.gov/38747148/)]
134. Boch S, Hussain SA, Bambach S, DeShetler C, Chisolm D, Linwood S. Locating youth exposed to parental justice involvement in the electronic health record: development of a natural language processing model. *JMIR Pediatr Parent.* Mar 21, 2022;5(1):e33614. [doi: [10.2196/33614](https://doi.org/10.2196/33614)] [Medline: [35311681](https://pubmed.ncbi.nlm.nih.gov/35311681/)]
135. Mahbub M, Goethert I, Danciu I, et al. Question-answering system extracts information on injection drug use from clinical notes. *Commun Med.* 2024;4(1). [doi: [10.1038/s43856-024-00470-6](https://doi.org/10.1038/s43856-024-00470-6)]
136. Wray CM, Vali M, Walter LC, et al. Examining the interfacility variation of social determinants of health in the veterans health administration. *Fed Pract.* Jan 2021;38(1):15-19. [doi: [10.12788/fp.0080](https://doi.org/10.12788/fp.0080)] [Medline: [33574644](https://pubmed.ncbi.nlm.nih.gov/33574644/)]
137. Bhate NJ, Mittal A, He Z, Luo X. Zero-shot learning with minimum instruction to extract social determinants and family history from clinical notes using GPT model. *Proc IEEE Int Conf Big Data.* Dec 2023;2023:1476-1480. [doi: [10.1109/BigData59044.2023.10386811](https://doi.org/10.1109/BigData59044.2023.10386811)] [Medline: [39101057](https://pubmed.ncbi.nlm.nih.gov/39101057/)]
138. Dai HJ, Su ECY, Uddin M, Jonnagaddala J, Wu CS, Syed-Abdul S. Exploring associations of clinical and social parameters with violent behaviors among psychiatric patients. *J Biomed Inform.* Nov 2017;75S:S149-S159. [doi: [10.1016/j.jbi.2017.08.009](https://doi.org/10.1016/j.jbi.2017.08.009)] [Medline: [28822857](https://pubmed.ncbi.nlm.nih.gov/28822857/)]
139. Purpura A, Mulligan N, Kartoun U, Koski E, Anand V, Bettencourt-Silva J. Investigating cross-domain binary relation classification in biomedical natural language processing. *AMIA Jt Summits Transl Sci Proc.* 2024;2024:384-390. [Medline: [38827064](https://pubmed.ncbi.nlm.nih.gov/38827064/)]
140. Krishnamoorthy R, Nagarajan V, Pour H, et al. Voice-enabled response analysis agent (VERAA): leveraging large language models to map voice responses in SDoH survey. *AMIA Jt Summits Transl Sci Proc.* 2024;2024:258-265. [Medline: [38827075](https://pubmed.ncbi.nlm.nih.gov/38827075/)]
141. Wang EA, Long JB, McGinnis KA, et al. Measuring exposure to incarceration using the electronic health record. *Med Care.* Jun 2019;57 Suppl 6 Suppl 2(Suppl 6 2):S157-S163. [doi: [10.1097/MLR.0000000000001049](https://doi.org/10.1097/MLR.0000000000001049)] [Medline: [31095055](https://pubmed.ncbi.nlm.nih.gov/31095055/)]
142. Zhu V, Lenert L, Bunnell B, Obeid J, Jefferson M, Halbert CH. Automatically identifying financial stress information from clinical notes for patients with prostate cancer. *Cancer Res Rep.* 2020;1(1):102. [doi: [10.61545/crr-1-102](https://doi.org/10.61545/crr-1-102)] [Medline: [38317775](https://pubmed.ncbi.nlm.nih.gov/38317775/)]
143. Subramani S, Wang H, Vu HQ, Li G. Domestic violence crisis identification from facebook posts based on deep learning. *IEEE Access.* 2018;6:54075-54085. [doi: [10.1109/ACCESS.2018.2871446](https://doi.org/10.1109/ACCESS.2018.2871446)]
144. Subramani S, Michalska S, Wang H, Du J, Zhang Y, Shakeel H. Deep learning for multi-class identification from domestic violence online posts. *IEEE Access.* 2019;7:46210-46224. [doi: [10.1109/ACCESS.2019.2908827](https://doi.org/10.1109/ACCESS.2019.2908827)]
145. Liyanage C, Garg M, Mago V, Sohn S. Augmenting Reddit Posts to Determine Wellness Dimensions Impacting Mental Health. *Association for Computational Linguistics*; 2023:306-312. [doi: [10.18653/v1/2023.bionlp-1.27](https://doi.org/10.18653/v1/2023.bionlp-1.27)]
146. Martin EA, D'Souza AG, Saini V, Tang K, Quan H, Eastwood CA. Extracting social determinants of health from inpatient electronic medical records using natural language processing. *J Epidemiol Popul Health.* Dec 2024;72(6):202791. [doi: [10.1016/j.jep.2024.202791](https://doi.org/10.1016/j.jep.2024.202791)] [Medline: [39546940](https://pubmed.ncbi.nlm.nih.gov/39546940/)]
147. Gabriel RA, Litake O, Simpson S, Burton BN, Waterman RS, Macias AA. On the development and validation of large language model-based classifiers for identifying social determinants of health. *Proc Natl Acad Sci USA.* Sep 24, 2024;121(39). [doi: [10.1073/pnas.2320716121](https://doi.org/10.1073/pnas.2320716121)]
148. Ralevski A, Taiyab N, Nossal M, Mico L, Piekos S, Hadlock J. Using large language models to abstract complex social determinants of health from original and deidentified medical notes: development and validation study. *J Med Internet Res.* 2024;26:e63445. [doi: [10.2196/63445](https://doi.org/10.2196/63445)]
149. Gong L, Bresnick J, Zhang A, Wu C, Jha K. Boosting social determinants of health extraction with semantic knowledge augmented large language model. *AMIA Annu Symp Proc.* 2024;2024:453-462. [Medline: [40417469](https://pubmed.ncbi.nlm.nih.gov/40417469/)]
150. Moon S, Wu Y, Doughty JB, et al. Automated identification of patients' unmet social needs in clinical text using natural language processing. *Mayo Clin Proc Digit Health.* Sep 2024;2(3):411-420. [doi: [10.1016/j.mcpdig.2024.06.008](https://doi.org/10.1016/j.mcpdig.2024.06.008)] [Medline: [39324128](https://pubmed.ncbi.nlm.nih.gov/39324128/)]

151. Roy S, Morrell S, Zhao L, Homayouni R. Large-scale identification of social and behavioral determinants of health from clinical notes: comparison of Latent Semantic Indexing and Generative Pretrained Transformer (GPT) models. *BMC Med Inform Decis Mak*. Oct 10, 2024;24(1):296. [doi: [10.1186/s12911-024-02705-x](https://doi.org/10.1186/s12911-024-02705-x)] [Medline: [39390479](https://pubmed.ncbi.nlm.nih.gov/39390479/)]
152. Gong L, Shor A, Zhang A, Jha K. Context-specific feature augmentation for improving social determinants of health extraction. Presented at: 2024 IEEE International Conference on Big Data (BigData); Dec 15-18, 2024:1736-1745; Washington, DC, USA. [doi: [10.1109/BigData62323.2024.10825225](https://doi.org/10.1109/BigData62323.2024.10825225)]
153. Roosan D, Chok J, Li Y, Khou T. Utilizing quantum computing-based large language transformer models to identify social determinants of health from electronic health records. Presented at: 2024 International Conference on Electrical, Computer and Energy Technologies (ICECET); Jul 25-27, 2024; Sydney, Australia. [doi: [10.1109/ICECET61485.2024.10698600](https://doi.org/10.1109/ICECET61485.2024.10698600)]
154. Shang T, Yang S, He W, et al. Leveraging social determinants of health in Alzheimer's research using LLM-augmented literature mining and knowledge graphs. *AMIA Jt Summits Transl Sci Proc*. 2025;2025:491-500. [Medline: [40502260](https://pubmed.ncbi.nlm.nih.gov/40502260/)]
155. Keloth VK, Selek S, Chen Q, et al. Social determinants of health extraction from clinical notes across institutions using large language models. *NPJ Digit Med*. May 17, 2025;8(1):287. [doi: [10.1038/s41746-025-01645-8](https://doi.org/10.1038/s41746-025-01645-8)] [Medline: [40379919](https://pubmed.ncbi.nlm.nih.gov/40379919/)]
156. Consoli B, Wang H, Wu X, et al. SDoH-GPT: using large language models to extract social determinants of health. *J Am Med Inform Assoc*. Jan 1, 2026;33(1):67-78. [doi: [10.1093/jamia/ocaf094](https://doi.org/10.1093/jamia/ocaf094)] [Medline: [40493530](https://pubmed.ncbi.nlm.nih.gov/40493530/)]
157. Gu Z, He L, Naeem A, et al. SBDH-Reader: a large language model-powered method for extracting social and behavioral determinants of health from clinical notes. *J Am Med Inform Assoc*. Oct 1, 2025;32(10):1570-1580. [doi: [10.1093/jamia/ocaf124](https://doi.org/10.1093/jamia/ocaf124)] [Medline: [40720632](https://pubmed.ncbi.nlm.nih.gov/40720632/)]
158. Chen Z, Lasserre P, Lin A, Rajapakshe R. Extraction of social determinants of health from electronic health records using natural language processing. *JCO Clin Cancer Inform*. Jul 2025;9:e2400317. [doi: [10.1200/CCI-24-00317](https://doi.org/10.1200/CCI-24-00317)] [Medline: [40700678](https://pubmed.ncbi.nlm.nih.gov/40700678/)]
159. Gu B, Shao V, Liao Z, et al. Scalable information extraction from free text electronic health records using large language models. *BMC Med Res Methodol*. 2025;25(1). [doi: [10.1186/s12874-025-02470-z](https://doi.org/10.1186/s12874-025-02470-z)]
160. Wang Y, Hilsman J, Li C, et al. Development and validation of natural language processing algorithms in the national ENACT network. *J Clin Trans Sci*. 2025;9(1). [doi: [10.1017/cts.2025.10116](https://doi.org/10.1017/cts.2025.10116)]
161. Peng C, Yu Z, Smith KE, Lo-Ciganic WH, Bian J, Wu Y. Enhancing cross-domain generalizability in social determinants of health extraction with prompt-tuning large language models. *AMIA Jt Summits Transl Sci Proc*. 2025;2025:432-440. [Medline: [40502248](https://pubmed.ncbi.nlm.nih.gov/40502248/)]
162. Shao M, Kang Y, Hu X, Kwak HG, Yang C, Lu J. Mining social determinants of health for heart failure patient 30-day readmission via large language model. *Stud Health Technol Inform*. Aug 7, 2025;329:1902-1903. [doi: [10.3233/SHTI251272](https://doi.org/10.3233/SHTI251272)] [Medline: [40776289](https://pubmed.ncbi.nlm.nih.gov/40776289/)]
163. Zaribafzadeh H, Henson JB, Chan NW, et al. Development of a natural language processing algorithm to extract social determinants of health from clinician notes. *Am J Transplant*. Jun 2025;25(6):1306-1318. [doi: [10.1016/j.ajt.2025.02.019](https://doi.org/10.1016/j.ajt.2025.02.019)]
164. Wang S, Wei Y, Ma H, et al. A multi-stage large language model framework for extracting suicide-related social determinants of health. *Commun Med*. 2025;5(1). [doi: [10.1038/s43856-025-01114-z](https://doi.org/10.1038/s43856-025-01114-z)]
165. Chapman AB, Panadero T, Dalrymple R, et al. Studying Veteran food insecurity longitudinally using electronic health record data and natural language processing. *AMIA Jt Summits Transl Sci Proc*. Jan 16, 2025:124-133. [Medline: [40502216](https://pubmed.ncbi.nlm.nih.gov/40502216/)]
166. Scherbakov D, Heider PM, Wehbe R, Alekseyenko AV, Lenert LA, Obeid JS. Using large language models for extracting stressful life events to assess their impact on preventive colon cancer screening adherence. *BMC Public Health*. 2025;25(1). [doi: [10.1186/s12889-024-21123-2](https://doi.org/10.1186/s12889-024-21123-2)]
167. Wasser LM, Liang HW, Li C, et al. Identifying transportation needs in ophthalmology clinic notes using natural language processing: retrospective, cross-sectional study. *JMIR Med Inform*. Sep 5, 2025;13:e69216. [doi: [10.2196/69216](https://doi.org/10.2196/69216)] [Medline: [40913248](https://pubmed.ncbi.nlm.nih.gov/40913248/)]
168. Patra BG, Lepow LA, Kasi Reddy Jagadeesh Kumar P, et al. Extracting social support and social isolation information from clinical psychiatry notes: comparing a rule-based natural language processing system and a large language model. *J Am Med Inform Assoc*. Jan 1, 2025;32(1):218-226. [doi: [10.1093/jamia/ocae260](https://doi.org/10.1093/jamia/ocae260)] [Medline: [39423850](https://pubmed.ncbi.nlm.nih.gov/39423850/)]
169. Kim MH, Miramontes S, Mehta S, et al. Extracting housing and food insecurity information from clinical notes using cTAKES. *Health Serv Res*. May 2025;60 Suppl 3(Suppl 3):e14440. [doi: [10.1111/1475-6773.14440](https://doi.org/10.1111/1475-6773.14440)] [Medline: [39871689](https://pubmed.ncbi.nlm.nih.gov/39871689/)]
170. Sakib FA, Zhu Z, Grace KT, Yetisgen M, Uzuner O. Spurious Correlations and Beyond: Understanding and Mitigating Shortcut Learning in SDOH Extraction with Large Language Models. *Association for Computational Linguistics*; 2025:1097-1106. [doi: [10.18653/v1/2025.acl-short.86](https://doi.org/10.18653/v1/2025.acl-short.86)]
171. Xu N, Zhang Q, Du C, et al. Revealing emergent human-like conceptual representations from language prediction. *Proc Natl Acad Sci U S A*. Nov 4, 2025;122(44):e2512514122. [doi: [10.1073/pnas.2512514122](https://doi.org/10.1073/pnas.2512514122)] [Medline: [41171846](https://pubmed.ncbi.nlm.nih.gov/41171846/)]

172. Li C, Guo J, Bian J, Becich MJ. Advancing social determinants of health research and practice: data, tools, and implementation. *J Clin Transl Sci*. 2025;9(1):e58. [doi: [10.1017/cts.2024.657](https://doi.org/10.1017/cts.2024.657)] [Medline: [40201637](https://pubmed.ncbi.nlm.nih.gov/40201637/)]
173. Islam S, Elmekki H, Elsebai A, et al. A comprehensive survey on applications of transformers for deep learning tasks. *Expert Syst Appl*. May 2024;241:122666. [doi: [10.1016/j.eswa.2023.122666](https://doi.org/10.1016/j.eswa.2023.122666)]
174. Gong EJ, Bang CS, Lee JJ, Baik GH. Knowledge-practice performance gap in clinical large language models: systematic review of 39 benchmarks. *J Med Internet Res*. Dec 1, 2025;27:e84120. [doi: [10.2196/84120](https://doi.org/10.2196/84120)] [Medline: [41325597](https://pubmed.ncbi.nlm.nih.gov/41325597/)]
175. Goldstein ND, Olivieri-Mui B, Burstyn I. Are aggregated electronic health record datasets good for research? *J Gen Intern Med*. Nov 2025;40(15):3743-3749. [doi: [10.1007/s11606-025-09808-9](https://doi.org/10.1007/s11606-025-09808-9)] [Medline: [40794368](https://pubmed.ncbi.nlm.nih.gov/40794368/)]
176. Li C, Mowery DL, Ma X, et al. Realizing the potential of social determinants data in EHR systems: a scoping review of approaches for screening, linkage, extraction, analysis, and interventions. *J Clin Transl Sci*. 2024;8(1):e147. [doi: [10.1017/cts.2024.571](https://doi.org/10.1017/cts.2024.571)] [Medline: [39478779](https://pubmed.ncbi.nlm.nih.gov/39478779/)]
177. Alberto IRI, Alberto NRI, Ghosh AK, et al. The impact of commercial health datasets on medical research and health-care algorithms. *Lancet Digit Health*. May 2023;5(5):e288-e294. [doi: [10.1016/S2589-7500\(23\)00025-0](https://doi.org/10.1016/S2589-7500(23)00025-0)] [Medline: [37100543](https://pubmed.ncbi.nlm.nih.gov/37100543/)]
178. Artsi Y, Sorin V, Glicksberg BS, Korfiatis P, Nadkarni GN, Klang E. Large language models in real-world clinical workflows: a systematic review of applications and implementation. *Front Digit Health*. 2025;7:1659134. [doi: [10.3389/fgdth.2025.1659134](https://doi.org/10.3389/fgdth.2025.1659134)] [Medline: [41098649](https://pubmed.ncbi.nlm.nih.gov/41098649/)]
179. Calderon N, Reichart R. On behalf of the stakeholders: trends in nlp model interpretability in the era of llms. Presented at: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics; Apr 29 to May 4, 2025:656-693; Albuquerque, New Mexico. URL: <https://aclanthology.org/2025.naacl-long> [doi: [10.18653/v1/2025.naacl-long.29](https://doi.org/10.18653/v1/2025.naacl-long.29)]
180. Senghor AS, Bright TJ, Kakim S, et al. A community-based approach to ethical decision-making in artificial intelligence for health care. *JAMIA Open*. Aug 2025;8(4):ooaf076. [doi: [10.1093/jamiaopen/ooaf076](https://doi.org/10.1093/jamiaopen/ooaf076)] [Medline: [40799930](https://pubmed.ncbi.nlm.nih.gov/40799930/)]
181. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. *J Biomed Inform*. Jan 2021;113:103655. [doi: [10.1016/j.jbi.2020.103655](https://doi.org/10.1016/j.jbi.2020.103655)] [Medline: [33309898](https://pubmed.ncbi.nlm.nih.gov/33309898/)]
182. Feller DJ, Zucker J, Yin MT, Gordon P, Elhadad N, Walk OBD. Longitudinal analysis of social and behavioral determinants of health in the ehr: exploring the impact of patient trajectories and documentation practices. Presented at: Proceedings of the AMIA Annual Symposium; Mar 4, 2020:399-407; Washington, DC. [Medline: [32308833](https://pubmed.ncbi.nlm.nih.gov/32308833/)]
183. Hernán MA, McAdams M, McGrath N, Lanoy E, Costagliola D. Observation plans in longitudinal studies with time-varying treatments. *Stat Methods Med Res*. Feb 2009;18(1):27-52. [doi: [10.1177/0962280208092345](https://doi.org/10.1177/0962280208092345)] [Medline: [19036915](https://pubmed.ncbi.nlm.nih.gov/19036915/)]
184. Chiu CC, Wu CM, Chien TN, Kao LJ, Li C, Chu CM. Integrating structured and unstructured EHR data for predicting mortality by machine learning and latent dirichlet allocation method. *Int J Environ Res Public Health*. Feb 28, 2023;20(5):4340. [doi: [10.3390/ijerph20054340](https://doi.org/10.3390/ijerph20054340)] [Medline: [36901354](https://pubmed.ncbi.nlm.nih.gov/36901354/)]
185. Simon BD, Ozyoruk KB, Gelikman DG, Harmon SA, Türkbey B. The future of multimodal artificial intelligence models for integrating imaging and clinical metadata: a narrative review. *Diagn Interv Radiol*. Jul 8, 2025;31(4):303-312. [doi: [10.4274/dir.2024.242631](https://doi.org/10.4274/dir.2024.242631)] [Medline: [39354728](https://pubmed.ncbi.nlm.nih.gov/39354728/)]
186. Berg K, Doktorchik C, Quan H, Saini V. Automating data collection methods in electronic health record systems: a social determinant of health (SDOH) viewpoint. *Health Syst (Basingstoke)*. 2023;12(4):472-480. [doi: [10.1080/20476965.2022.2075796](https://doi.org/10.1080/20476965.2022.2075796)] [Medline: [38235302](https://pubmed.ncbi.nlm.nih.gov/38235302/)]
187. Egede LE, Walker RJ, Williams JS. Addressing structural inequalities, structural racism, and social determinants of health: a vision for the future. *J Gen Intern Med*. Feb 2024;39(3):487-491. [doi: [10.1007/s11606-023-08426-7](https://doi.org/10.1007/s11606-023-08426-7)] [Medline: [37740168](https://pubmed.ncbi.nlm.nih.gov/37740168/)]
188. Yelton B, Rumthao JR, Sakhuja M, et al. Assessment and documentation of social determinants of health among health care providers: qualitative study. *JMIR Form Res*. Jul 3, 2023;7:e47461. [doi: [10.2196/47461](https://doi.org/10.2196/47461)] [Medline: [37399064](https://pubmed.ncbi.nlm.nih.gov/37399064/)]
189. Eisinger-Mathason TSK, Leshin J, Lahoti V, Fridsma DB, Mucaj V, Kho AN. Data linkage multiplies research insights across diverse healthcare sectors. *Commun Med (Lond)*. Mar 4, 2025;5(1):58. [doi: [10.1038/s43856-025-00769-y](https://doi.org/10.1038/s43856-025-00769-y)] [Medline: [40038513](https://pubmed.ncbi.nlm.nih.gov/40038513/)]
190. Cross JL, Choma MA, Onofrey JA. Bias in medical AI: implications for clinical decision-making. *PLOS Digit Health*. Nov 2024;3(11):e0000651. [doi: [10.1371/journal.pdig.0000651](https://doi.org/10.1371/journal.pdig.0000651)] [Medline: [39509461](https://pubmed.ncbi.nlm.nih.gov/39509461/)]
191. Chen Y, Clayton EW, Novak LL, Anders S, Malin B. Human-centered design to address biases in artificial intelligence. *J Med Internet Res*. Mar 24, 2023;25:e43251. [doi: [10.2196/43251](https://doi.org/10.2196/43251)] [Medline: [36961506](https://pubmed.ncbi.nlm.nih.gov/36961506/)]
192. Sasseville M, Ouellet S, Rhéaume C, et al. Bias mitigation in primary health care artificial intelligence models: scoping review. *J Med Internet Res*. Jan 7, 2025;27:e60269. [doi: [10.2196/60269](https://doi.org/10.2196/60269)] [Medline: [39773888](https://pubmed.ncbi.nlm.nih.gov/39773888/)]
193. Hasanzadeh F, Josephson CB, Waters G, Adedinsewo D, Azizi Z, White JA. Bias recognition and mitigation strategies in artificial intelligence healthcare applications. *NPJ Digit Med*. Mar 11, 2025;8(1):154. [doi: [10.1038/s41746-025-01503-7](https://doi.org/10.1038/s41746-025-01503-7)] [Medline: [40069303](https://pubmed.ncbi.nlm.nih.gov/40069303/)]

194. Rajwal S, Garg S, Abdel-Salam R, Zayed A. Do biased models have biased thoughts. Presented at: Second Conference on Language Modeling; Oct 7-10, 2025; Montreal, Canada. URL: <https://openreview.net/forum?id=vDr0RV3590> [Accessed 2026-03-03]
195. Humphreys BL, Del Fiol G, Xu H. The UMLS knowledge sources at 30: indispensable to current research and applications in biomedical informatics. *J Am Med Inform Assoc*. Oct 1, 2020;27(10):1499-1501. [doi: [10.1093/jamia/ocaa208](https://doi.org/10.1093/jamia/ocaa208)] [Medline: [33059366](https://pubmed.ncbi.nlm.nih.gov/33059366/)]
196. Miller K, Moon S, Fu S, Liu H. Contextual variation of clinical notes induced by EHR migration. *AMIA Annu Symp Proc*. 2023;2023:1155-1164. [Medline: [38222426](https://pubmed.ncbi.nlm.nih.gov/38222426/)]
197. Rockenschaub P, Hilbert A, Kossen T, et al. The impact of multi-institution datasets on the generalizability of machine learning prediction models in the ICU. *Crit Care Med*. 2024;52(11):1710-1721. [doi: [10.1097/CCM.0000000000006359](https://doi.org/10.1097/CCM.0000000000006359)]

Abbreviations

EHR: electronic health record

LLM: large language model

MIMIC: Medical Information Mart for Intensive Care

MiSSoM+: Minority Stress on Social Media

NCR: normalized citation ratio

NLP: natural language processing

PLUTO: Primary Land Use Tax Lot Output

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

SDOH: social determinants of health

Edited by Stefano Brini; peer-reviewed by Leela Prasad Gorrepati, Oluranti Akinsola, Robert Marshall, Roland Abi; submitted 08.Sep.2025; final revised version received 09.Mar.2026; accepted 09.Mar.2026; published 28.Apr.2026

Please cite as:

Rajwal S, Pandey AK, Zhang Z, Chen Y, Liu MX, Das S, Rogers H, Sarker A, Xiao Y

Applications of Natural Language Processing and Large Language Models for Social Determinants of Health: Systematic Review

J Med Internet Res 2026;28:e83793

URL: <https://www.jmir.org/2026/1/e83793>

doi: [10.2196/83793](https://doi.org/10.2196/83793)

© Swati Rajwal, Avinash Kumar Pandey, Ziyuan Zhang, Yankai Chen, Michael X Liu, Sudeshna Das, Hannah Rogers, Abeer Sarker, Yunyu Xiao. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 28.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.