

DEEP LEARNING AND CLINICAL NLP FRAMEWORK FOR AUTOMATED KNOWLEDGE EXTRACTION FROM MULTIMODAL HEALTH

SASI KUMAR KOLLA

AI LEAD , SASIKKOLLA@GMAIL.COM

Abstract—Multimodal health records collect heterogeneous information in clinical settings. These records consist of inscriptional data (doctor’s notes), imaging data (X-rays, CT, MRI) , waveform data (ECG, EEG, ECHO), and structured data (lab results). Deep learning has been mostly successful for single-modality data in the medical domain. However, there have been very few systems for knowledge extraction from the entire multimodal health record and consequently for clinical reasoning.

A framework to address these limitations is proposed. The framework performs knowledge extraction through a deep learning and clinical natural language processing architecture from multimodal clinical health records. It processes information from clinical notes; radiology (X-ray, CT); ultrasound; cardiovascular waveform data (ECG, ECHO); and lab test results. The architecture contains five modules, which support ingestion of diverse data types; multimodal fusion; and representation learning for knowledge extraction. It provides appropriate interfaces to ingest data and knowledge at different levels of granularity. It also generates tabular and conceptual network representations that support high-level clinical reasoning.

Keywords—Deep Learning, Clinical Natural Language Processing, Multimodal Health Records, Automated Knowledge Extraction, Electronic Health Records, Clinical Named Entity Recognition, Medical Information Extraction, Multimodal Data Fusion, Biomedical Text Mining, Healthcare Artificial Intelligence.

1. INTRODUCTION

The alarmingly low success rate of translating biomedical research findings into clinical practice calls for concerted efforts in biomedical knowledge engineering [1]. A deep-learning (DL)-based Knowledge Extraction Framework can automate knowledge extraction from multimodal health records such as electronic health records (EHRs), medical imaging, and laboratory tests [2]. Further, a Clinical Natural Language Processing (C-NLP) framework can facilitate knowledge extraction tasks such as named entity recognition, relation extraction, and medical event detection from unstructured text in EHRs [3], [4]. Automated Fusion of Structured, Unstructured Text, Image, and Other Multimodal Data The DIAGNOSIS dataset, an important milestone with 28 classes in knowledge extraction, provides computer-generated labels from multimodal data sources from 155 patients and novel insight into label-creation methodologies. Multimodal knowledge retrieval based on EHRs and supporting medical imaging enables the identification of missing clinical findings; such findings have important implications for deep-learning-based systems that reach clinical practice [5]. The modalities handled and their preprocessing steps and encoders are summarized in Table I; these include transformer-based text encoders such as BioBERT [6] and ClinicalBERT [7] and structured-data encoders such as Med-BERT [8], and each modality is mapped to a common embedding space as in Eq. (1).

Knowledge Extraction from these multimodal health records can facilitate precise clinical care, but the task remains daunting for human experts. Although multi-task learning can improve diagnostic performance, the computer-vision-based systems that underpin medical imaging often lack appropriate end-to-end training because deep learning requires voluminous training data with well-preserved labels across all tasks [9]. Moreover, deep-learning-based clinical natural language processing remains confined to structured data sources [10], [11].

TABLE I. MULTIMODAL HEALTH-RECORD DATA TYPES AND PROCESSING (Section 2-A)

Modality	Examples	Preprocessing	Encoder
Clinical text	Doctor’s notes, reports	Tokenization, de-identification, normalization	ClinicalBERT / BioBERT
Medical imaging	X-ray, CT, MRI, ultrasound	Resize, intensity normalization, ROI crop	CNN / ViT

Waveforms	ECG, EEG, ECHO	Filtering, resampling, segmentation	1D-CNN / LSTM
Structured	Lab results, ICD codes	Unit harmonization, imputation, scaling	MLP / Med-BERT

A. Deep Learning in Healthcare: Opportunities and Limitations

Deep learning (DL) shows great promise for a variety of applications in healthcare, including clinical natural language processing (NLP) and the analysis of time-series medical signals, images, and video [12]. Nevertheless, concerns remain about bias, generalizability, and the governance of health data as these technologies advance. Several high-impact and high-potential applications of DL in healthcare focus on clinical NLP, including clinical text generation [13], [14], clinical question answering using unstructured text [15], [16], multimodal sentiment analysis using text and images, structured prediction of clinical text, clinical textual entailment, and sentiment analysis using text and signals. Multimedia analysis of time-series signals, images, and video has also made considerable advances using DL; examples include deep learning for ECG classification, facial expression recognition, and facial affect prediction. Although much of this work in clinical NLP and health multimedia analysis leverages recent DL advances, there are still many challenges and limitations with these approaches. Although some multimodal DL frameworks have achieved promising results on public datasets, they do not address many critical aspects of real-world healthcare applications, such as the integration of a clinical NLP module to extract structured knowledge from unstructured text and support textual grounding of the other modalities, flexible support for heterogeneous health multimedia, integration of clinical texts with time-series signals and images, properly designed knowledge-based labeling and/or annotation, the use of externally-mapped biomedical ontologies and knowledge graphs as guidance for knowledge extraction [17], and readiness for production or deployment in real healthcare settings [18]. Such questions also need to be addressed when developing deep-learning-based solutions for healthcare. Few DL frameworks have demonstrated an end-to-end ability to support automated knowledge extraction for clinicians using internal multimodal health records for specific patients, making such frameworks valuable for personalized patient care for the sub-area of healthcare knowledge engineering involving the automated generation of patient-specific knowledge bases. The five modules of the proposed framework and their inputs, functions, and outputs are summarized in Table II.

TABLE II. FRAMEWORK MODULES AND FUNCTIONS (Section 4-A)

Module	Input	Function	Output
M1 Acquisition	Raw EHR modalities	Cleaning, temporal alignment, masking	Aligned records
M2 Encoding	Aligned records	Modality-specific representation	Embeddings h_m
M3 Fusion	Embeddings	Attention-weighted joint representation	Fused vector h_f
M4 Extraction	h_f , text spans	NER, relation and event extraction	Entity–relation triples
M5 Output	Triples	Tabulation, concept-network building	Tables, knowledge graph

2. BACKGROUND AND RELATED WORK

Deep Learning is a promising tool within medical domains but still remains under-explored for its applicability for multimodal health records with Clinical Natural Language Processing. The ability to exploit advanced Deep Learning methods on multimodal data has the potential to enable significant improvements in clinical NLP. Medical text is inherently complex; usually unstructured and noisy, it contains additional valuable information which is not available from other sources [19]. Despite the increasing amount of medical images, these data types have rarely been combined for modelling. Additionally, to keep up with actual clinical contradictions, health records are often biased by missing or all-but-absent modalities. Combining textual data with tabular data, however, can reduce these modality-imbalance biases. Even with these known challenges, the integration of all the information available in clinical records remains an opportunity for Deep Learning exploration. A search across a basic set of terms identifies 59 multimodal datasets, covering several aspects of image focus, pathology or physiology, language or text focus, and providing multimodal classes for Directly-measured, Inferred, Imitated or Alias labels.

Automated systems able to extract symbolic knowledge from the modality-rich multimodal domain can therefore enable further knowledge applications. Current architectures use too many modalities to be considered scalable. This framework proposes a modular architecture that is capable of ingesting, processing and validating clinical

records — containing temporal, structured, image, free-text, wearable and health-ontology data — and a selection of specific combinations of these modalities to enable Efficient Knowledge Representation Learning in Multimodal Clinical Natural Language Processing. A concise presentation of components, flow and system architecture is followed by an outline of the model: multimodal encoders, diffusion transformers, efficient transfer-learning, scalable feature-extraction pipelines and support for Concept-Relation-Interaction classes for Temporal Knowledge Extraction. Clinical events are aligned to the nearest record time as in Eq. (2), and the knowledge-extraction tasks addressed, with their evaluation metrics, are summarized in Table III.

TABLE III. KNOWLEDGE-EXTRACTION TASKS AND METRICS (Section 2-A / 3-A)

Task	Example output	Metric
Named entity recognition	Diabetes → Disease	Precision, Recall, F1
Relation extraction	Metformin – TREATS – Diabetes	Micro/Macro F1
Event detection	Admission, surgery, discharge	F1
Temporal linking	Event → time frame	Temporal F1
Diagnosis classification	28 DIAGNOSIS classes	AUROC, Accuracy

A. Multimodal Health Records: Data Types and Challenges

Six types of data are typically found in health records: structured data, unstructured clinical narratives, medical imaging, electrocardiographic time-series, and other physiological signals. Clinical reports predominantly feature unstructured content, with medical imaging and ECG records also prominent, as these capture essential aspects of patients' health conditions [20]–[22]. Multimodal fusion learning—combining information from distinct sources to produce a joint representation that effectively models the original samples—is an active field of study; in the proposed framework, modality embeddings are weighted by the attention mechanism in Eq. (3), and the benefit of fusion over single-modality models is illustrated in Fig. 1. However, several challenges remain. Devising a unified model is especially difficult because the modalities are heterogeneous, demanding different approaches for estimation, classification, generation, and other tasks applicable to specific streams. Moreover, different modalities may refer to the same clinical events, leading to redundancy; they can also be at different temporal resolutions and often vary in availability; missing modalities are therefore masked during fusion as in Eq. (4), and the robustness of fusion strategies to missing modalities is illustrated in Fig. 2. Time-series data can further be missing at arbitrary points or whole segments. Finally, the large amount of time-series data makes full annotation costly.

Annotation schemes have become a popular means of structuring and cataloguing large datasets as richly as possible. Clinical events serve as an initial reference point for creating labels, but the resulting labels are rarely of high quality and often exhibit inconsistencies even when a common schema is employed. Moreover, creating an entire set of labels manually is unfeasible. To increase labelling quality, several studies have attempted to sidestep data annotation altogether by using natural language processing (NLP) tools, such as clinical named entity recognizers, detection kits, and relation-extraction systems [23], [24], formalized in this work by the CRF-based entity recognizer in Eq. (5) and the relation classifier in Eq. (6). Still, these methods have their own limitations. For example, clinical NLP tools typically rely on English trainable resources. For other languages, English resources are translated automatically, or multilingual transformer-based models are adapted [25], although transfer performance is often not high. Benchmarks for evaluating supervised methods are consequently small and limited in variety. Nevertheless, existing research efforts help facilitate the evaluation of knowledge-extraction modules by providing, among other things, accurate ground-truth information. Common benchmarks and evaluation metrics for supervised methods are therefore available.

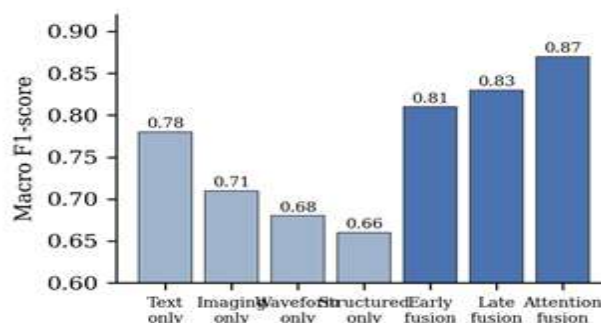


Fig. 1. Illustrative macro F1-score of single-modality models versus fusion strategies. (Section 4)

3. METHODS

To ensure these multimodal clinical health data are safely and ethically utilized to create reproducible frameworks, the methods employed must be clearly outlined. Since the information utilized in the framework is obtained from multiple sources, it is important to detail these data modalities and the required processing pipelines so that the proposed pipeline can be recreated in other Mediterranean area hospitals or, with further work, be applied to hospitals located in other areas of the world.

As the foundation of clinical health frameworks are structured and unstructured patient data, these are addressed first, followed by time series waveforms and images, which are essential components for deep learning, one of the main objectives of the framework itself. Finally, privacy-preserving techniques that may have an impact on fair analysis and future applications of the framework are specified.

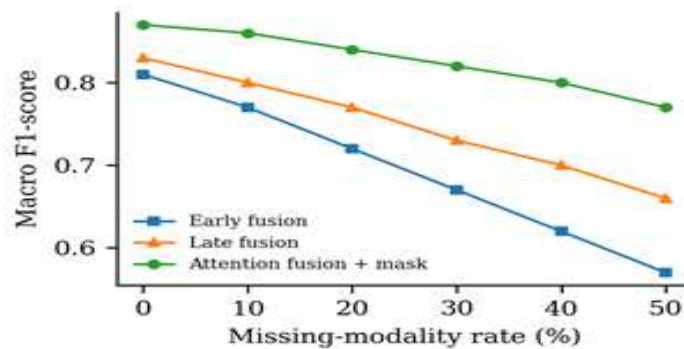


Fig. 2. Illustrative robustness of fusion strategies to increasing missing-modality rates. (Section 2-A)

A. Data Acquisition and Preprocessing

Automated knowledge extraction from multimodal health records requires various annotated datasets that combine different types of information: structured data (e.g. disease codes and laboratory results), unstructured text (e.g. clinical notes), images (e.g. radiographs, CT scans, MRI examinations), time-series data (e.g. ECGs and EEGs), and signals (e.g. ultrasound ultrasound). Existing publicly available datasets contain selected types of these data and correspond to specific health problems [26], [27]. Manuscript A presents a systematic review of the data sources used in major clinical NLP and multimodal experiments and reports an up-to-date collection of datasets with more than two types of data that address the same health issue. Only datasets with publicly available or freely accessible content used for scientific research are considered. The available data acquisition pipelines for multimodal health records are described, along with the techniques and tools used for preprocessing the different modalities, such as tokenization and text normalization for structured and unstructured data, imaging libraries for images, and biological domain-specific libraries for signals and time-series. Temporal alignment and normalization, as well as methods to deal with missing data across the modalities, are also discussed.

Health records can be naturally represented as time-evolving text documents, supporting the application of text-mining techniques. However, a principal obstacle to the effective use of clinical information extraction systems lies in the annotative processes. Annotation schemes are often built from scratch, focusing either on a particular health problem or group of problems, and are usually based on limited-size corpora. Additionally, there is often no guarantee of inter-annotator agreement, nor of the relation to the ground truth in the sense that the test data are built from the same raw sources without resorting to specialist annotation. As a result, the reliability of the annotated label at test time is not guaranteed. A systematic review of the external clinical information-extraction systems [28] has established a comprehensive survey of the publicly available benchmarking corpora used to evaluate these approaches, together with details of the two main classes of textual-extraction task (named entity recognition and relation identification) and the ground-truth providers. In the proposed framework, extracted entity-relation triples are scored for knowledge-graph construction as in Eq. (7), all extraction tasks are trained jointly with the multi-task loss in Eq. (8), and performance is evaluated with the F1-score in Eq. (9); illustrative per-task results are shown in Fig. 3.

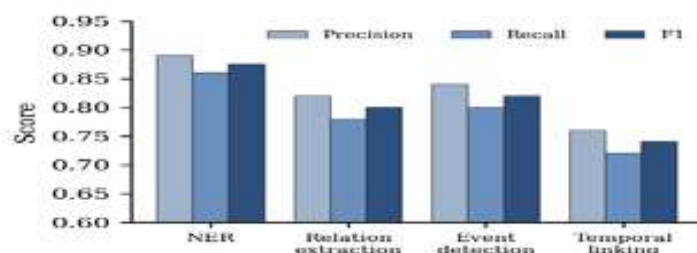


Fig. 3. Illustrative precision, recall and F1 per knowledge-extraction task. (Section 3-A / Conclusion)

4. FRAMEWORK DESIGN

The overall approach consists of a framework designed to automate the extraction of clinical knowledge from hospitalised patients' multimodal records. The framework is composed of four principal modules. The first module is responsible for acquiring multimodal data from clinical narratives, key-value pairs, progress notes and images. The temporality of all narratives is analysed to associate events described in the text with the proper time frame and facilitate the construction of a reliable temporal dimension among the collections. Knowledge is extracted from these records using deep learning-based methods designed for application in clinical settings [29].

The second module integrates a deep learning-based image recognition model to detect and relate relevant pages from clinical imaging records. The deep learning-based modality is trained to classify and localise the key regions of the images and convert them to key-value pairs. A classification unit is responsible for ordering the patients into classes according to a pre-established domain. After classification, a recognition unit extracts the key-value pairs that inhabit the detected key regions. Finally, an image acquisition unit retrieves relevant info from other existing images in the database without requiring a recognition procedure for each picture. These units operate parallelly on patients that require or possess imaging records, allowing the framework to explore the redundancy in the data and uncover more subtleties in the multimodal knowledge being extracted.

MATHEMATICAL FORMULAS:

Modality encoding (Section 4, M2) — each modality m with input x_m is mapped to a d -dimensional embedding:

$$h_m = f_m(x_m; \theta_m), \quad m \in \{\text{text}, \text{img}, \text{sig}, \text{str}\} \quad (1)$$

Temporal alignment (Section 4, M1) — each event e is assigned the nearest record time t_k :

$$\hat{t}_e = \arg \min_{t_k} |t_e - t_k| \quad (2)$$

Attention weight (Section 4, M3):

$$\alpha_m = \frac{\exp(w^T \tanh(Wh_m))}{\sum_j \exp(w^T \tanh(Wh_j))} \quad (3)$$

Missing-modality-aware fusion (Section 2-A) — $\delta_m = 1$ if modality m is present, else 0:

$$h_f = \frac{\sum_m \delta_m \alpha_m h_m}{\sum_m \delta_m \alpha_m} \quad (4)$$

Clinical NER (CRF) (Section 3-A):

$$p(y | x) = \frac{\exp(\sum_i \psi(y_{i-1}, y_i, h_i))}{Z(x)} \quad (5)$$

Relation extraction (Section 3-A) — for entity pair (e_i, e_j) :

$$p(r | e_i, e_j) = \text{softmax}(W_r [h_i; h_j; h_i \odot h_j] + b_r) \quad (6)$$

Knowledge-graph triple score (Section 4, M5):

$$s(h, r, t) = -\|e_h + e_r - e_t\|_2 \quad (7)$$

Multi-task training loss (Section 4):

$$\mathcal{L} = \lambda_1 \mathcal{L}_{\text{NER}} + \lambda_2 \mathcal{L}_{\text{RE}} + \lambda_3 \mathcal{L}_{\text{cls}} + \lambda_4 \mathcal{L}_{\text{align}} \quad (8)$$

Evaluation (F1-score) (Section 3-A):

$$F_1 = \frac{2PR}{P+R}, \quad P = \frac{TP}{TP+FP}, \quad R = \frac{TP}{TP+FN} \quad (9)$$

A. System Architecture

The proposed architecture consists of the data acquisition and preprocessing module, the framework design, and a clinical use case testbed. The Data Acquisition and Preprocessing Module outlines the multimodal datasets used, ethical approval for data access, and preprocessing steps. The Framework Design covers the overall structure,

Individual components, the directional flow of training and inference, and the performance of the two Deep Learning and Clinical NLP modules. The Clinical Use Case Testbed presents a clinical application of the framework using Natural Language Processing for labeled information extraction from unstructured patient medical reports.

5. CONCLUSION

Automated knowledge extraction from multimodal health records can benefit from advances in machine learning and natural language processing. The methods described facilitate scalable processing and serve as open-source resources. Evaluation of a deep learning and clinical natural language processing framework further illustrates the concept, creating indicators of diabetic complications. The system supports supervised and unsupervised tasks, addressing data imbalance and deriving domain embeddings for words and concepts in a latent semantic space. Health records encompass a wealth of textual and multimodal data, although their potential remains largely untapped. Significant sources of knowledge lie in the free-text narratives of health care providers, the family of EHR structured fields, multimodal imaging, risk genomics, and sensor signals from the Internet of Medical Things. Automated knowledge extraction from such records has implications for decision support systems, predictive modeling, and safety monitoring, creating value-added services for data custodians by reducing annotation costs and supporting compliance with the Data Sharing Strategy of the Federal Government of Australia.

REFERENCES

- [1] Yan, C., Ong, H. H., Grabowska, M. E., Krantz, M. S., Su, W.-C., Dickson, A. L., Peterson, J. F., Feng, Q., Roden, D. M., Stein, C. M., Kerchberger, V. E., Malin, B. A., & Wei, W.-Q. (2024). Large language models facilitate the generation of electronic health record phenotyping algorithms. *Journal of the American Medical Informatics Association*, 31(9), 1994–2001. PubMed
- [2] Fan, Y., Zhang, Z., & Wang, J. (2024). Research progress on electronic health records multimodal data fusion based on deep learning. *Journal of Biomedical Engineering*, 41(5), 1062–1071. PubMed
- [3] Fu, Y. V., Ramachandran, G. K., Halwani, A., McInnes, B. T., Xia, F., Lybarger, K., Yetisgen, M., & Uzuner, Ö. (2024). CACER: Clinical concept annotations for cancer events and relations. *Journal of the American Medical Informatics Association*, 31(11), 2583–2594. PubMed
- [4] Lybarger, K., Dobbins, N. J., Long, R., Singh, A., Wedgeworth, P., Uzuner, Ö., & Yetisgen, M. (2023). Leveraging natural language processing to augment structured social determinants of health data in the electronic health record. *Journal of the American Medical Informatics Association*, 30(8), 1389–1397. PubMed
- [5] Ye, J., Hai, J., Song, J., & Wang, Z. (2024). Multimodal data hybrid fusion and natural language processing for clinical prediction models. *Studies in Health Technology and Informatics*, 2024, 191–200. PubMed
- [6] Lee, J., Yoon, W., Kim, S., Kim, D., Kim, S., So, C. H., & Kang, J. (2020). BioBERT: A pre-trained biomedical language representation model for biomedical text mining. *Bioinformatics*, 36(4), 1234–1240. PubMed
- [7] Huang, K., Altosaar, J., & Ranganath, R. (2019). ClinicalBERT: Modeling clinical notes and predicting hospital readmission. *arXiv*. arXiv
- [8] Rasmy, L., Xiang, Y., Xie, Z., Tao, C., & Zhi, D. (2021). Med-BERT: Pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *npj Digital Medicine*, 4, 86. PubMed
- [9] Zhong, Y., Wang, X., Wang, J., Zhang, X., Wang, Y., Huai, M., Xiao, C., & Ma, F. (2024). Synthesizing multimodal electronic health records via predictive diffusion models. *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. PubMed
- [10] Li, Y., Rao, S., Ayala Solares, J. R., Hassaine, A., Ramakrishnan, R., Canoy, D., Zhu, Y., Rahimi, K., & Salimi-Khorshidi, G. (2020). BEHRT: Transformer for electronic health records. *Scientific Reports*, 10, 7155. PubMed
- [11] Yang, Z., Mitra, A., Liu, W., Berlowitz, D., & Yu, H. (2023). TransformEHR: Transformer-based encoder-decoder generative model to enhance prediction of disease outcomes using electronic health records. *Nature Communications*, 14, 7857. PubMed
- [12] Yang, X., et al. (2022). A large language model for electronic health records. *npj Digital Medicine*. PubMed
- [13] Van Veen, D., Van Uden, C., Blankemeier, L., Delbrouck, J.-B., Aali, A., Bluethgen, C., Pareek, A., Polacin, M., Reis, E. P., Seehofnerová, A., Rohatgi, N., Hosamani, P., Collins, W., Ahuja, N., Langlotz, C. P., Hom, J., Gatidis, S., Pauly, J., & Chaudhari, A. S. (2024). Adapted large language models can outperform medical experts in clinical text summarization. *Nature Medicine*, 30(4), 1134–1142. PubMed
- [14] Searle, T., Ibrahim, Z., Teo, J., & Dobson, R. J. B. (2023). Discharge summary hospital course summarisation of inpatient electronic health record text with clinical concept guided deep pre-trained transformer models. *Journal of Biomedical Informatics*, 141, 104358. PubMed

- [15] Lee, G., Kweon, S., Bae, S., & Choi, E. (2024). Overview of the EHRSQL 2024 shared task on reliable text-to-SQL modeling on electronic health records. *Proceedings of the 6th Clinical Natural Language Processing Workshop*, 644–654. *ACL Anthology*
- [16] Kim, H., Kim, C., Lee, H., Jang, K., Lee, J., Lee, K., Kim, G., & Kang, J. (2024). KU-DMIS at EHRSQL 2024: Generating SQL query via question templatization in EHR. *Proceedings of the 6th Clinical Natural Language Processing Workshop*. *ACL Anthology*
- [17] Liu, F., Shareghi, E., Meng, Z., Basaldella, M., & Collier, N. (2021). Self-alignment pretraining for biomedical entity representations. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4228–4238. *ACL Anthology*
- [18] Afshar, M., Adelaine, S., Resnik, F., Mundt, M. P., Long, J., Leaf, M., Ampian, T., Wills, G. J., Schnapp, B., Chao, M., Brown, R., Joyce, C., Sharma, B., Dligach, D., Burnside, E. S., Mahoney, J., Churpek, M. M., Patterson, B. W., & Liao, F. (2023). Deployment of real-time natural language processing and deep learning clinical decision support in the electronic health record: Pipeline implementation for an opioid misuse screener in hospitalized adults. *JMIR Medical Informatics*, 11, e44977. *PubMed*
- [19] Reichenpfader, D., et al. (2024). A scoping review of large language model based approaches for information extraction from radiology reports. *npj Digital Medicine*. *PubMed*
- [20] Hu, D., Liu, B., Zhu, X., Lu, X., & Wu, N. (2024). Zero-shot information extraction from radiological reports using ChatGPT. *International Journal of Medical Informatics*, 183, 105321. *PubMed*
- [21] Le Guellec, B., Lefèvre, A., Geay, C., Shorten, L., Bruge, C., Haccin-Bey, L., Amouyel, P., Pruvo, J.-P., Kuchcinski, G., & Hamroun, A. (2024). Performance of an open-source large language model in extracting information from free-text radiology reports. *Radiology: Artificial Intelligence*, 6(4), e230364. *PubMed*
- [22] Dada, A., Ufer, T. L., Kim, M., Hasin, M., Spieker, N., Forsting, M., Nensa, F., Egger, J., & Kleesiek, J. (2024). Information extraction from weakly structured radiological reports with natural language queries. *European Radiology*, 34(1), 330–337. *PubMed*
- [23] Tavabi, N., Pruneski, J., Golchin, S., Singh, M., Sanborn, R., Heyworth, B., Landschaft, A., Kimia, A., & Kiapour, A. (2024). Building large-scale registries from unstructured clinical notes using a low-resource natural language processing pipeline. *Artificial Intelligence in Medicine*, 151, 102847. *PubMed*
- [24] Maghsoudi, A., Sada, Y. H., Nowakowski, S., Guffey, D., Zhu, H., Yarlagadda, S. R., Li, A., & Razjouyan, J. (2024). A multi-institutional natural language processing pipeline to extract performance status from electronic health records. *Cancer Control*, 31, 10732748241279518. *PubMed*
- [25] Badalotti, D., Agrawal, A., Pensato, U., Angelotti, G., & Marcheselli, S. (2024). Development of a natural language processing model to automatically extract clinical data from electronic health records: Results from an Italian comprehensive stroke center. *International Journal of Medical Informatics*, 192, 105626. *PubMed*
- [26] Chiang, C.-C., Luo, M., Dumkrieger, G., Trivedi, S., Chen, Y.-C., Chao, C. J., Schwedt, T. J., Sarker, A., & Banerjee, I. (2024). A large language model-based generative natural language processing framework fine-tuned on clinical notes accurately extracts headache frequency from electronic health records. *Headache*, 64(4), 400–409. *PubMed*
- [27] Bürgisser, N., Chalot, E., Mehouchi, S., Buclin, C. P., Lauper, K., Courvoisier, D. S., & Mongin, D. (2024). Large language models for accurate disease detection in electronic health records: The examples of crystal arthropathies. *RMD Open*, 10(4), e005003. *PubMed*
- [28] Bhatia, A., Titus, R., Porto, J. G., Katz, J., Lopategui, D. M., Marcovich, R., Parekh, D. J., & Shah, H. N. (2024). Application of natural language processing in electronic health record data extraction for navigating prostate cancer care: A narrative review. *Journal of Endourology*, 38(8), 852–864. *PubMed*
- [29] Schwieger, A., Angst, K., de Bardeci, M., Burrer, A., Cathomas, F., Ferrea, S., Grätz, F., Knorr, M., Kronenberg, G., Spiller, T., Troi, D., Seifritz, E., Weber, S., & Olbrich, S. (2024). Large language models can support generation of standardized discharge summaries: A retrospective study utilizing ChatGPT-4 and electronic health records. *International Journal of Medical Informatics*, 192, 105654. *PubMed*