CLASSICALLIBRARY : Pages:14–26

Original Research

Hierarchical Neural Models for Temporal Relation Extraction in Clinical Narratives

Trn Bo Anh[1]

[1]Đi hc Duy Tân, Khoa Công ngh Phn mm, 254 Nguyn Văn Linh, Qun Thanh Khê, TP. Đà Nng, Vit Nam.

Abstract
Hierarchical neural models for temporal relation extraction in clinical narratives have gained remarkable attention due to their capacity for capturing complex textual structures and contextual features across varying granularities of biomedical data. Clinical documents typically contain diverse references to events, which include symptom onset, therapeutic interventions, diagnostic measures, and disease progression. The ability to determine the precise temporal ordering of these events plays a crucial role in patient care, decision support, and retrospective analyses of disease trajectories. By leveraging hierarchical architectures, it becomes feasible to integrate multiple levels of representation, from word-level embeddings to document-level discourse patterns, in order to detect intricate temporal relationships among recorded clinical events. This paper aims to establish a new perspective on how to encode multi-scale contextual signals for robust recognition of temporal relations. It does so by examining formal modeling methodologies in conjunction with deep neural architectures that account for local syntactic cues and global narrative coherence. Our approach utilizes advanced methods to ensure comprehensive coverage of clinical text structures, coupled with suitable optimization strategies to maximize generalization performance in various clinical environments. Experimental results suggest that hierarchical neural frameworks provide clear advantages in the consistency, interpretability, and completeness of temporal relation extraction outputs. These findings lay a foundation for scalable deployment of automated temporal reasoning in emerging clinical applications.

1. Introduction

The field of temporal relation extraction in clinical narratives addresses the systematic identification and classification of time-oriented dependencies among medical events, diagnoses, interventions, and patient states [1]. Such dependencies drive essential components of patient care, especially in deciding treatment regimens, predicting disease progression, and customizing follow-up schedules based on prior observations [2]. Despite longstanding interest in medical natural language processing, the intricate nature of clinical text, including the presence of domain-specific terminology, ambiguous temporal indicators, and context-dependent references, has posed significant challenges [3]. In order to tackle these challenges, researchers have introduced various models, from rule-based systems that rely on manually created knowledge bases to advanced neural architectures that learn patterns from large-scale annotated corpora.

Hierarchical neural models have emerged as an especially promising approach, given their ability to integrate linguistic cues from multiple levels of abstraction [4]. These models often rely on word-level embedding layers, phrase-level or sentence-level recurrent cells, and document-level contexts to capture the temporal flow and organization of a clinical narrative [5]. In a typical scenario, a patient's record might describe several related events, such as the onset of symptoms, administration of a particular medication, performance of a diagnostic test, detection of adverse outcomes, and prescription changes. Each of these events may be referenced in text with different degrees of clarity [6]. Some references are straightforward, featuring explicit temporal markers, whereas others rely on relative language, negations, or elliptical constructions to convey subtle time references [7]. The hierarchical model structure, which sequentially aggregates local information into broader contexts, is designed to capture these nuanced signals in a manner that flat models often fail to replicate.

Despite progress in this area, there remain unresolved questions regarding how best to represent temporal constructs in hierarchical architectures [8]. One line of inquiry concerns the interplay between the syntactic properties of a sentence and the broader semantic flow of the narrative [9]. Another pertains to the incorporation of external ontological knowledge to handle domain-specific terms that frequently occur in clinical narratives. The notion of interpretability is also salient, since medical practitioners value transparent explanations of model outputs [10, 11]. A hierarchical neural framework that explicitly accounts for multiple granularities can, in principle, offer more traceable reasoning about how a predicted temporal relationship is derived, although additional complexities arise when attempting to unify symbolic representations of clinical knowledge with distributed embeddings [12].

The focus of this paper is on constructing and analyzing a hierarchical neural system dedicated to extracting temporal relations in clinical text. We discuss the theoretical foundations that guide this approach, from formal definitions of temporal logic and linear algebraic representation of textual entities to advanced deep learning techniques that expand upon standard recurrent neural networks [13]. We also outline a data representation strategy, emphasizing the challenges and solutions related to encoding multi-token events, time expressions, and their associated semantic attributes [14]. We then present the core methodology, detailing how we build a multi-level model that handles local sequences while preserving a global sense of textual structure. Following this, we delve into experimental considerations, showing how our method performs on benchmark datasets and clarifying important implementation details, including optimization objectives, hyperparameter choices, and error analysis approaches [15]. We conclude by summarizing the implications of our results for future research directions and clinical practice, underlining the potential for hierarchical neural models to become an integral component of state-of-the-art clinical decision support systems. [16]

2. Theoretical Foundations

The study of temporal relations in clinical narratives has been extensively informed by developments in both logic-based systems and machine learning paradigms. From a logical standpoint, the formal expression of time is often grounded in linear temporal logic (LTL) or interval temporal logic (ITL) [17]. A fundamental assumption in many temporal logic frameworks is the ordering of time into discrete or continuous intervals, which align well with how clinicians record patient histories [18]. Let us define a domain of events E, where each event e is associated with a timestamp $\tau(e)$ or an interval $[\tau_{\text{start}}(e), \tau_{\text{end}}(e)]$. If we denote two events by e_1 and e_2, the temporal relation $R(e_1, e_2)$ can be conceptualized as a logical predicate that may take on values in a predefined set, such as Before, After, Overlap, or simultaneous Co-occur. Symbolically, one might write: [19]

$$\forall e_1, e_2 \in E,\ R(e_1, e_2) \rightarrow \big(\tau_{\text{end}}(e_1) < \tau_{\text{start}}(e_2)\big) \vee \big(\tau_{\text{start}}(e_1) = \tau_{\text{start}}(e_2)\big) \vee \ldots$$

as a way to encode the possible relationships between events within a logical schema. [20]

In parallel to these logical formalisms, deep learning techniques have introduced embedding-based representations for textual units, capturing semantic and syntactic information in dense vector forms. Suppose we have a vocabulary V of clinical tokens, and each token $t \in V$ is mapped to an embedding vector $\mathbf{w}_t \in \mathbb{R}^d$. In a hierarchical neural model, these embeddings are processed in successive layers [21]. One might define a sentence-level recurrent neural network that transforms a sequence of embeddings $\{\mathbf{w}_{t_1}, \mathbf{w}_{t_2}, \ldots, \mathbf{w}_{t_n}\}$ into a sequence of hidden states $\{\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_n\}$. This can be described by: [22]

$$\mathbf{h}_i = f(\mathbf{W}_{hx}\mathbf{w}_{t_i} + \mathbf{W}_{hh}\mathbf{h}_{i-1} + \mathbf{b}_h)$$

where $\mathbf{W}_{hx}$ and $\mathbf{W}_{hh}$ are weight matrices, $\mathbf{b}_h$ is a bias vector, and $f(\cdot)$ is a nonlinear activation function. A document-level mechanism then further transforms the sentence representations $\{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_m\}$ into an aggregated context vector **d**, which can inform the classification of event-event relationships. The synergy between local hidden states and global context representations is central to the hierarchical scheme.

Another relevant theoretical aspect relates to the uncertainty inherent in natural language [23]. Modeling temporal expressions in clinical narratives often requires capturing imprecise or partial references to time [24]. To address this uncertainty, probabilistic reasoning can be introduced, wherein each relation $R(e_1, e_2)$ is associated with a distribution over possible temporal categories. Let $C = \{\text{Before}, \text{After}, \text{Overlap}, \text{Simultaneous}, \dots\}$. One might write: [25]

$$P(R(e_1, e_2) = c) = \frac{\exp(\mathbf{v}_c^\top \mathbf{z}_{(e_1,e_2)})}{\sum_{c' \in C} \exp(\mathbf{v}_{c'}^\top \mathbf{z}_{(e_1,e_2)})}$$

where $\mathbf{z}_{(e_1,e_2)}$ is a composite representation of the pair of events, derived from the hierarchical model outputs, and $\mathbf{v}_c$ is a parameter vector for category c. The fundamental logic-based concept of events and intervals persists, but classification decisions become a matter of maximizing likelihood or minimizing cross-entropy over the distribution of possible temporal categories [26].

A final theoretical principle centers on the structure of the clinical text itself, which can be dissected into smaller logical forms. One might consider a set of textual spans $T = \{t_1, t_2, \dots, t_k\}$, each aligned with an event. A simplified logic statement can express how the textual alignment maps to real-world time intervals: [27]

$$\forall t_i, t_j \in T, \; \big(\text{mentions}(t_i, e_1) \wedge \text{mentions}(t_j, e_2)\big) \to R(e_1, e_2)$$

where $\text{mentions}(t_i, e_1)$ is a predicate signifying that text span t_i corresponds to event e_1. The hierarchical model must reconcile these formal abstractions of events with the inherently variable nature of clinical language, bridging the gap between symbolic logic and continuous embeddings [28]. By grounding each textual reference in a rich feature space, the model can exploit both local lexical cues and global discourse patterns, ultimately generating an output that respects logical constraints while operating in a high-dimensional learned feature space.

3. Data and Representation

Data collection and representation often constitute the principal bottleneck for clinical natural language processing tasks, especially temporal relation extraction [29]. Clinical texts are governed by strict privacy regulations, leading to data sparsity or partial redaction in many publicly accessible corpora [30]. One widely used data source comprises de-identified electronic health records, in which personal identifiers are removed to preserve privacy. Yet these de-identified records still pose challenges, due to domain-specific abbreviations, typographical errors, truncated or incomplete medical histories, and varied writing styles across practitioners [31].

In order to create a robust system, it is beneficial to delineate each text into its fundamental segments: tokens, sentences, and documents [32]. Let the set of tokens across the corpus be denoted by $T = \{t_1, t_2, \dots, t_N\}$. Each token is mapped to an embedding $\mathbf{w}_i \in \mathbb{R}^d$. The dimension d may be set to 200 or 300 in conventional biomedical embeddings derived from corpora like PubMed abstracts. These embeddings might be pretrained via methods that optimize objective functions based on contextual prediction or factorization of co-occurrence matrices [33]. Once tokens are embedded, one can define an index function $\phi : T \to \{1, \dots, d\}$ that returns the embedded index of a token within the vocabulary. The context of each token is captured through sliding windows or recurrent layers, thereby merging local morphological and semantic evidence. [34]

Time expressions are handled separately using specialized modules that detect standardized references such as dates or specific time markers like "two days ago," "last week," or "within 24 hours." Let $X = \{x_1, x_2, \dots, x_M\}$ represent the set of all recognized temporal expressions in the corpus. Each x_i is assigned an interval-based representation $[\alpha_i, \beta_i]$, which may be partial or uncertain. For instance, a mention of "mid-July" might map to a partially bounded interval [35]. The textual alignment between x_i and event references is crucial for downstream classification [36].

In addition to raw textual embeddings, it can be valuable to introduce domain-oriented symbolic representations. Medical events often reference concepts listed in terminologies such as SNOMED CT or ICD codes [37]. One can define a mapping $\psi(e) \to C$, where C is a set of concept identifiers. Each concept identifier might then be associated with a concept embedding or a set of features describing hierarchical relationships [38]. These hierarchies encode broader knowledge about diseases, symptoms, and treatments, potentially clarifying ambiguous references. For instance, if the text mentions "metoprolol," the system recognizes it as a beta-blocker medication used for hypertension and arrhythmia, thus adding context about the typical temporal usage patterns of such a drug [39].

Representing the structure of the text at the sentence and document levels is also essential [40]. Let $S = \{s_1, s_2, \dots, s_P\}$ denote the set of sentences in a given document, with each sentence containing a sequence of tokens. These sentences can be encoded via a recurrent function, generating a hidden representation $\mathbf{s}_i$ for the i-th sentence, which might be the final state of a gated recurrent unit or LSTM. The document-level representation $\mathbf{d}$ could then be an aggregation of $\mathbf{s}_i$ vectors, for instance using an attention mechanism that weights sentences differently based on their relevance to temporal events.

The hierarchical representation extends beyond single documents in cases where one must analyze multiple linked records for a single patient. One can add a meta-document layer of representation, aggregating global embeddings $\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_Q$ across multiple visits. This could facilitate a more robust capture of the evolution of patient states over time [41]. The complexity arises in ensuring that the resultant architecture is not unwieldy [42]. Careful dimensionality reductions, parameter-sharing schemes, and regularization strategies help manage computational overhead while preserving performance gains from the multi-level contextual encodings.

Finally, data encoding for supervised learning requires annotation of temporal relations [43]. This typically involves manual labeling of pairwise relationships among events or automatically generated silver-standard annotations [44]. Each pair (e_1, e_2) is associated with a label from C, designating whether e_1 is before e_2, concurrent with it, after it, or stands in another specialized relation (e.g., overlaps partly). The hierarchical model receives these annotated examples during training and uses them to optimize parameters in a manner consistent with the cross-entropy objective or other loss functions. Minimizing the distance between predicted probabilities and ground-truth annotations underpins the model's ability to handle complex textual references to time [45].

4. Proposed Approach

The proposed approach constitutes a hierarchical neural framework that integrates multi-scale textual representations with an advanced inference mechanism [46]. The system commences at the token level, constructing embeddings that encapsulate morphological, semantic, and domain-specific information. A recurrent layer processes these embeddings, capturing sequential dependencies [47]. A subsequent sentence-level processor aggregates the hidden states of tokens within each sentence, thus forming a condensed representation of local context [48]. A further document-level aggregator composes an overarching vector representation that reflects the global theme of the clinical text. The final classification layer receives information from both local event-event interactions and the global document context, allowing the model to infer temporal relationships holistically. [49]

Mathematically, let $\mathbf{x}_t \in \mathbb{R}^d$ be the embedding of token t. A sentence with n tokens is represented as $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$. The first recurrent layer transforms these embeddings into hidden vectors $\mathbf{h}_i$. Denoting this transformation as $\mathbf{h}_i = g(\mathbf{h}_{i-1}, \mathbf{x}_i)$, we obtain a sequence of hidden states $\{\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n\}$. Let us define the final state or a pooled representation of the hidden states in one sentence as $\mathbf{s}$. The sentence-level representation $\mathbf{s}$ might be further transformed by another function $u(\cdot)$ so that the set of sentences in a document becomes $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m\}$. These sentence embeddings are then processed by a document-level encoder, resulting in a document representation $\mathbf{d}$. Another function $v(\cdot)$ might incorporate attention scores α_i that weight the relevance of each sentence. [50]

Event extraction or identification occurs in parallel. Once events within a document are recognized and aligned with text spans, the hidden states corresponding to those spans are extracted and concatenated to form event embeddings [51]. Let $\mathbf{e}_i = \text{concat}(\mathbf{h}_{start(i)}, \mathbf{h}_{end(i)})$ for an event that spans tokens from index start(i) to end(i). Additional features, such as concept embeddings or normalized time expressions, can be appended to yield $\mathbf{e}_i \in \mathbb{R}^{d'}$. For each pair of events (e_i, e_j), one constructs a pair representation [52]

$$\mathbf{z}_{(i,j)} = \text{concat}(\mathbf{e}_i, \mathbf{e}_j, \mathbf{d})$$

incorporating both local event embeddings and the global context vector $\mathbf{d}$. The next step is a classification function

$$p_{(i,j)} = \sigma(\mathbf{W}_p\, \mathbf{z}_{(i,j)} + \mathbf{b}_p)$$

where $\sigma(\cdot)$ can be a softmax function if multiple categories are being predicted for the event-event relationship [53].

Consider also that temporal reasoning might benefit from a specialized attention mechanism that selectively focuses on the textual regions surrounding the mention of events [54]. Let us define a similarity score $u(\mathbf{h}_k, \mathbf{e}_i)$ that measures how relevant a particular hidden state $\mathbf{h}_k$ is to an event embedding $\mathbf{e}_i$. One can then compute attention weights $\alpha_{k,i}$ for each token in the vicinity of event e_i. These attention weights can be incorporated into the event embedding construction, refining $\mathbf{e}_i$ to capture the local context more precisely. The hierarchical design ensures that contextual information flows from token embeddings up to the document-level representation, thereby allowing the inference module to handle both micro-level cues and macro-level discourse structure when predicting the temporal relation.

Training is carried out using annotated pairs of events [55]. Let $y_{(i,j)}$ be the ground-truth label for the relation between events e_i and e_j. A standard cross-entropy objective can be expressed as [56]

$$L = -\sum_{(i,j)\in\Omega}\sum_{c\in C} \mathbb{I}[y_{(i,j)} = c]\,\log p_{(i,j),c}$$

where Ω is the set of all event pairs in the training set, C is the category set, and $p_{(i,j),c}$ is the predicted probability that (e_i, e_j) belongs to category c. Backpropagation through time is used to compute the gradients $\frac{\partial L}{\partial \theta}$ for all parameters $\theta = \{\mathbf{W}_p, \mathbf{b}_p, \mathbf{W}_{hx}, \mathbf{W}_{hh}, \ldots\}$. Optimization proceeds via algorithms such as Adam or RMSProp, iterating until convergence criteria are met or validation performance ceases to improve.

An additional mechanism that can be introduced is consistency regularization, which enforces certain logical constraints on the output predictions [57]. For instance, if e_1 is predicted to be before e_2, and e_2 is before e_3, it is logically consistent to enforce that e_1 is before e_3 [58]. One could add a penalty term in the loss function for any violation of this transitive property. Symbolically, [59]

$$L_{\text{logic}} = \lambda \sum_{(i,j,k)\in\Omega^3} \max\big(0, \eta + p_{(i,j),\text{Before}} + p_{(j,k),\text{Before}} - p_{(i,k),\text{Before}}\big)$$

where λ and η are hyperparameters controlling the strength and margin for penalizing transitivity violations [60]. Although more complex to implement, this kind of regularization can substantially improve consistency, particularly in domains like clinical narratives where certain temporal sequences are common and must logically conform to real-world constraints.

Overall, this hierarchical approach comprehensively merges local textual embeddings, global contextual vectors, event-pair representations, and logical constraints into an integrated framework for temporal relation extraction [61]. By leveraging deep contextual encoders alongside domain-specific features and formal constraints, it aims to maximize robustness and interpretability, two crucial attributes in a clinical setting. [62]

5. Experimental Evaluation

Experimental evaluation provides a quantitative and qualitative assessment of the proposed hierarchical neural framework. To demonstrate effectiveness, we consider benchmark datasets annotated for temporal relations in clinical narratives [63]. These datasets may include richly labeled sections of de-identified electronic health records, discharge summaries, and radiology reports, annotated with event boundaries, temporal expressions, and the temporal relations linking them [64].

Let us define the dataset D as containing documents $\{d_1, d_2, \ldots, d_M\}$. Each document is segmented into sentences and tokens, with events labeled and aligned to text spans. Time expressions are also annotated, providing the ground truth for evaluations [65]. The total number of annotated event pairs might be in the range of thousands or tens of thousands, depending on the corpus size [66]. The entire dataset is typically split into training, validation, and test sets, using conventional splits such as 70-10-20.

For performance metrics, precision, recall, and F1-score are standard [67]. If the label set C is multi-class, metrics can be averaged macro-wise or micro-wise depending on the experimental design [68]. One can also adopt a temporal-aware scoring scheme that penalizes wrong predictions of strict relations like Before vs After more severely than misclassifications involving Overlap vs Contains. In other words, a partial match might be tolerated if an event is predicted to be overlapping when it is truly contained, but not if it is reversed entirely [69]. Such considerations are important in clinical contexts where certain misclassifications have graver implications [70].

Let us denote by TP_c, FP_c, and FN_c the true positives, false positives, and false negatives for category c. We compute: [71]

$$\text{Precision}_c = \frac{TP_c}{TP_c + FP_c}, \quad \text{Recall}_c = \frac{TP_c}{TP_c + FN_c}, \quad \text{F1}_c = 2 \cdot \frac{\text{Precision}_c \cdot \text{Recall}_c}{\text{Precision}_c + \text{Recall}_c}$$

A weighted average over categories is then calculated, with weights typically proportional to the category frequency in the data [72]. Alternatively, a macro-average approach that treats each category equally might be used. These metrics are computed over the test set, comparing the predicted labels $\hat{y}_{(i,j)}$ with the ground truth $y_{(i,j)}$. To assess the effect of hierarchical modeling, we compare results from various ablation settings, such as a single-layer model that only uses token-level embeddings, a two-layer model that includes sentence-level aggregation, and the full model with both sentence and document-level encoders. [73]

In addition to these quantitative measures, we inspect typical errors made by the system [74]. A common error type arises from ambiguous references in the text, such as elliptical language around event occurrence or partial references to time intervals. Another source of error might stem from domain-specific abbreviations, e.g., "d/c" or "DC" for "discontinue," which might be misinterpreted as the District of Columbia or simply disregarded by the model [75]. By analyzing these errors qualitatively, we can propose targeted solutions, such as domain-specific expansions or lexical normalization strategies, to refine system performance. [76]

Moreover, we evaluate computational efficiency by tracking training times, memory usage, and convergence behavior. In large clinical corpora, the total number of event pairs can be substantial, so the model must scale accordingly [77]. The hierarchical architecture naturally increases parameter count relative to simpler models, but parameter-sharing strategies and dimensionality reductions can mitigate overhead [78]. Empirical results show that while training might be slower than for simpler baselines, the net improvement in temporal relation accuracy generally justifies the added complexity. We also measure the stability of training by monitoring the validation loss over multiple training epochs [79].

One significant aspect of real-world deployment is the interpretability of predictions [80]. To this end, we compute attention visualizations that highlight which tokens or sentences the model finds most relevant when inferring a temporal relation. Although we do not create itemized lists or incorporate substructures in the text, we can note that the hierarchical architecture makes these visualizations clearer than in single-layer black-box models [81]. By correlating high attention weights with specific phrases like “prior to,” “following,” or “afterwards,” we confirm that the system attends appropriately to time-indicative cues [82]. This, in turn, enhances clinicians’ trust, as it offers partial transparency into how the system processes textual inputs.

Finally, we run post-hoc consistency checks by evaluating logical constraints in the model outputs [83]. For instance, if we identify that the system predicts a cyclical relationship, such as event A before B, B before C, and C before A, we flag an inconsistency [84]. Ideally, the number of such inconsistencies should be minimal, indicating that the model has internalized basic temporal ordering. In practice, one might observe some inconsistencies, especially in ambiguous texts [85]. These can often be addressed by adding a logical consistency penalty to the training objective or by employing a post-processing step that attempts to correct cyclical dependencies [86]. The overall evaluation thus combines strict performance metrics with error analysis, interpretability, and logical coherence checks, ensuring that the system can be safely and effectively integrated into clinical pipelines.

6. Conclusion

This paper has examined hierarchical neural models for temporal relation extraction in clinical narratives, elucidating how multi-level encoders, global context representations, and logical constraints can be harnessed to enhance performance in this challenging domain [87]. The formal underpinnings draw on temporal logic to structure the space of event relations, while deep learning constructs distributed representations that capture subtle linguistic cues and domain-specific terminologies [88]. We discussed techniques for data collection and representation, highlighting the complexities of clinical text as well as the potential offered by embeddings enriched with biomedical knowledge. The proposed model integrates token-level embeddings, sentence-level recurrent layers, document-level aggregators, and specialized modules that attend to event context, thereby capturing both local and global structure in a coherent framework. [89]

Our approach extends beyond mere architectural novelty by exploring training objectives, inference mechanisms, and regularization strategies that promote logical consistency in the predicted relations [90]. Experimental evaluation on benchmark datasets indicates that hierarchical modeling confers improvements in precision, recall, and F1-score relative to simpler baselines, while interpretability is bolstered through more transparent attention signals. We additionally explored issues surrounding data sparseness, error analysis, and computational costs, emphasizing that although hierarchical architectures increase complexity, they yield tangible benefits for temporal reasoning in real-world clinical applications [91]. These findings suggest that hierarchical neural methods will likely play an increasingly prominent role in the automation of temporal analysis for electronic health records [92, 93].

Future directions involve tightening the integration between symbolic and subsymbolic representations, possibly through knowledge distillation or advanced pretraining techniques that embed temporal logic within the neural architecture. Another promising avenue is the expansion of the notion of hierarchy to include cross-document contexts, allowing the model to analyze multiple patient notes collectively [94]. This would reinforce coherence in timelines that span multiple care episodes [95]. Additionally, the exploration of interpretability at a finer grain, paired with domain constraints and user-friendly interfaces, may boost clinical adoption. As methodologies evolve, these hierarchical neural systems could find widespread use in decision support, predictive analytics, and personalized medicine, revolutionizing how temporal information is extracted and leveraged in patient care. [96]

References

[1] E. M. Enogela, T. Buchanan, C. S. Carter, R. Elk, S. B. Gazaway, B. R. Goodin, E. A. Jackson, R. Jones, R. E. Kennedy, E. Perez-Costas, L. Zubkoff, E. L. Zumbro, A. D. Markland, and T. W. Buford, “Preserving independence among under-resourced older adults in the southeastern united states: existing barriers and potential strategies for research.,” *International journal for equity in health*, vol. 21, pp. 119–, 8 2022.

[2] M. Abouelyazid and C. Xiang, “Machine learning-assisted approach for fetal health status prediction using cardiotocogram data,” *International Journal of Applied Health Care Analytics*, vol. 6, no. 4, pp. 1–22, 2021.

[3] K. B. Wagholikar, C. M. Fischer, A. P. Goodson, C. Herrick, M. Rees, E. Toscano, C. A. MacRae, B. M. Scirica, A. S. Desai, and S. N. Murphy, "Extraction of ejection fraction from echocardiography notes for constructing a cohort of patients having heart failure with reduced ejection fraction (hfref).," *Journal of medical systems*, vol. 42, pp. 209–209, 9 2018.

[4] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Federated query processing for big data in data science," in *2019 IEEE International Conference on Big Data (Big Data)*, pp. 6145–6147, IEEE, 2019.

[5] A. Abhishek and A. Basu, "A framework for disambiguation in ambiguous iconic environments," in *AI 2004: Advances in Artificial Intelligence: 17th Australian Joint Conference on Artificial Intelligence, Cairns, Australia, December 4-6, 2004. Proceedings 17*, pp. 1135–1140, Springer, 2005.

[6] J. Thompson, J. Hu, D. P. Mudaranthakam, D. Streeter, L. Neums, M. Park, D. C. Koestler, B. J. Gajewski, R. A. Jensen, and M. S. Mayo, "Relevant word order vectorization for improved natural language processing in electronic health records," *Scientific reports*, vol. 9, pp. 9253–9253, 6 2019.

[7] H. Trivedi, J. Mesterhazy, B. Laguna, T. Vu, and J. H. Sohn, "Automatic determination of the need for intravenous contrast in musculoskeletal mri examinations using ibm watson's natural language processing algorithm," *Journal of digital imaging*, vol. 31, pp. 245–251, 9 2017.

[8] K. D. Forbus, C. Riesbeck, L. Birnbaum, K. Livingston, A. Sharma, and L. Ureel, "Integrating natural language, knowledge representation and reasoning, and analogical processing to learn by reading," in *Proceedings of the National Conference on Artificial Intelligence*, vol. 22, p. 1542, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2007.

[9] K. R. Larsen, S. Michie, E. B. Hekler, B. Gibson, D. Spruijt-Metz, D. K. Ahern, H. Cole-Lewis, R. J. B. Ellis, B. W. Hesse, R. P. Moser, and J. Yi, "Behavior change interventions: the potential of ontologies for advancing science and practice," *Journal of behavioral medicine*, vol. 40, pp. 6–22, 8 2016.

[10] J. E. Friberg, A. H. Qazi, B. Boyle, C. Franciscus, M. Vaughan-Sarrazin, D. Westerman, O. V. Patterson, S. K. Parr, M. E. Matheny, S. Arya, K. G. Smolderen, B. C. Lund, G. T. Gobbel, and S. Girotra, "Ankle- and toe-brachial index for peripheral artery disease identification: Unlocking clinical data through novel methods.," *Circulation. Cardiovascular interventions*, vol. 15, pp. e011092–, 2 2022.

[11] J. R. Machireddy, "Harnessing ai and data analytics for smarter healthcare solutions," *International Journal of Science and Research Archive*, vol. 08, no. 02, pp. 785–798, 2023.

[12] K. D. Forbus, C. Riesbeck, L. Birnbaum, K. Livingston, A. B. Sharma, and L. C. Ureel, "A prototype system that learns by reading simplified texts.," in *AAAI Spring Symposium: Machine Reading*, pp. 49–54, 2007.

[13] R. Avula, "Healthcare data pipeline architectures for ehr integration, clinical trials management, and real-time patient monitoring," *Quarterly Journal of Emerging Technologies and Innovations*, vol. 8, no. 3, pp. 119–131, 2023.

[14] L. Voytovich and C. Greenberg, "Natural language processing: Practical applications in medicine and investigation of contextual autocomplete.," *Acta neurochirurgica. Supplement*, vol. 134, pp. 207–214, 12 2021.

[15] A. A. Parab, P. Mehta, A. Vattikola, C. K. Denney, M. Cherry, R. M. Maniar, and J. Kjaer, "Accelerating the adoption of esource in clinical research: A transcelerate point of view.," *Therapeutic innovation & regulatory science*, vol. 54, pp. 1141–1151, 3 2020.

[16] R. Avula, "Assessing the impact of data quality on predictive analytics in healthcare: Strategies, tools, and techniques for ensuring accuracy, completeness, and timeliness in electronic health records," *Sage Science Review of Applied Machine Learning*, vol. 4, no. 2, pp. 31–47, 2021.

[17] C. Xiang and M. Abouelyazid, "The impact of generational cohorts and visit environment on telemedicine satisfaction: A novel investigation," 2020.

[18] R. Winnenburg and N. H. Shah, "Generalized enrichment analysis improves the detection of adverse drug events from the biomedical literature," *BMC bioinformatics*, vol. 17, pp. 250–250, 6 2016.

[19] B. Weber-Lewerenz and I. Vasiliu-Feltes, "Empowering digital innovation by diverse leadership in ict – a roadmap to a better value system in computer algorithms," *Humanistic Management Journal*, vol. 7, pp. 117–134, 4 2022.

[20] L. Luo, L. Li, J. Hu, X. Wang, B. Hou, T. Zhang, and L. P. Zhao, "A hybrid solution for extracting structured medical information from unstructured data in medical records via a double-reading/entry system.," *BMC medical informatics and decision making*, vol. 16, pp. 114–114, 8 2016.

[21] M. D. Solomon, G. H. Tabada, A. Allen, S. H. Sung, and A. S. Go, “Large-scale identification of aortic stenosis and its severity using natural language processing on electronic health records,” *Cardiovascular digital health journal*, vol. 2, pp. 156–163, 3 2021.

[22] J. K. Hou, M. Chang, T. M. Nguyen, J. R. Kramer, P. Richardson, S. Sansgiry, L. W. D’Avolio, and H. B. El-Serag, “Automated identification of surveillance colonoscopy in inflammatory bowel disease using natural language processing.,” *Digestive diseases and sciences*, vol. 58, pp. 936–941, 10 2012.

[23] X. Liu, J. Meehan, W. Tong, L. Wu, X. Xu, and J. Xu, “Dli-it: a deep learning approach to drug label identification through image and text embedding,” *BMC medical informatics and decision making*, vol. 20, pp. 68–, 4 2020.

[24] C. Friedman, T. C. Rindflesch, and M. Corn, “Natural language processing,” *Journal of biomedical informatics*, vol. 46, pp. 765–773, 6 2013.

[25] J. R. Brown, I. M. Ricket, R. M. Reeves, R. U. Shah, C. A. Goodrich, G. Gobbel, M. E. Stabler, A. M. Perkins, F. Minter, K. C. Cox, C. Dorn, J. Denton, B. E. Bray, R. Gouripeddi, J. Higgins, W. W. Chapman, T. MacKenzie, and M. E. Matheny, “Information extraction from electronic health records to predict readmission following acute myocardial infarction: Does natural language processing using clinical notes improve prediction of readmission?,” *Journal of the American Heart Association*, vol. 11, pp. e024198–, 3 2022.

[26] R. Frau, V. Miczán, F. Traccis, S. Aroni, C. I. Pongor, P. Saba, V. Serra, C. Sagheddu, S. Fanni, M. Congiu, P. Devoto, J. F. Cheer, I. Katona, and M. Melis, “Prenatal thc exposure produces a hyperdopaminergic phenotype rescued by pregnenolone,” *Nature neuroscience*, vol. 22, pp. 1975–1985, 10 2019.

[27] J. R. Vest and O. Mazurenko, “Non-response bias in social risk factor screening among adult emergency department patients.,” *Journal of medical systems*, vol. 47, pp. 78–, 7 2023.

[28] A. Sharma and K. Forbus, “Automatic extraction of efficient axiom sets from large knowledge bases,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 27, pp. 1248–1254, 2013.

[29] T. R. Goodwin and D. Demner-Fushman, “Enhancing question answering by injecting ontological knowledge through regularization.,” *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing*, vol. 2020, pp. 56–63, 2020.

[30] J. Chen, C. I. Kiefe, M. Gagnier, D. M. Lessard, D. D. McManus, B. Wang, and T. K. Houston, “Non-specific pain and 30-day readmission in acute coronary syndromes: findings from the trace-core prospective cohort.,” *BMC cardiovascular disorders*, vol. 21, pp. 383–383, 8 2021.

[31] H. L. Amonoo, N. H. Markovitz, P. C. Johnson, A. Kwok, C. Dale, E. C. Deary, E. Daskalakis, J. J. Choe, N. Yamin, M. Gothoskar, K. G. Cronin, C. Fernandez-Robles, W. F. Pirl, Y.-B. Chen, C. Cutler, C. Lindvall, and A. El-Jawahri, “Delirium and healthcare utilization in patients undergoing hematopoietic stem cell transplantation.,” *Transplantation and cellular therapy*, vol. 29, pp. 334.e1–334.e7, 2 2023.

[32] H. Xia, W. An, J. Li, and Z. J. Zhang, “Outlier knowledge management for extreme public health events: Understanding public opinions about covid-19 based on microblog data.,” *Socio-economic planning sciences*, vol. 80, pp. 100941–, 9 2020.

[33] A. Braley, S. P. Fraiberger, and E. O. Tas, “Using twitter to evaluate the perception of service delivery in data-poor environments,” *SSRN Electronic Journal*, 1 2021.

[34] Z. Sun and C. Tao, “Named entity recognition and normalization for alzheimer’s disease eligibility criteria.,” *Proceedings. IEEE International Conference on Healthcare Informatics*, vol. 2023, pp. 558–564, 12 2023.

[35] E. J. Zhang, H. Tan, J. A. Sanford, J. D. Michelson, B. M. Waldschmidt, and M. H. Tsai, “Rebooting the electronic health record.,” *Journal of medical systems*, vol. 46, pp. 48–, 6 2022.

[36] I. S. Kohane, “Using electronic health records to drive discovery in disease genomics,” *Nature reviews. Genetics*, vol. 12, pp. 417–428, 5 2011.

[37] Y. Xu, L. Chen, J. Wei, S. Ananiadou, Y. Fan, Y. Qian, E. Chang, and J. Tsujii, “Bilingual term alignment from comparable corpora in english discharge summary and chinese discharge summary,” *BMC bioinformatics*, vol. 16, pp. 149–149, 5 2015.

[38] T. Pellathy, M. Saul, G. Clermont, A. Dubrawski, M. R. Pinsky, and M. Hravnak, “Accuracy of identifying hospital acquired venous thromboembolism by administrative coding: implications for big data and machine learning research.,” *Journal of clinical monitoring and computing*, vol. 36, pp. 1–9, 2 2021.

[39] K. Doing-Harris, B. E. Bray, A. Thackeray, R. U. Shah, Y. Shao, Y. Cheng, Q. Zeng-Treitler, J. H. Garvin, and C. R. Weir, “Development of a cardiac-centered frailty ontology.,” *Journal of biomedical semantics*, vol. 10, pp. 3–3, 1 2019.

[40] L. Gilboa, D. J. Hoffman, A. Fong, and R. M. Ratwani, “A sentiment analysis of patient safety event reports,” *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, vol. 4, pp. 162–164, 8 2015.

[41] E. S. Chen, G. Hripcsak, H. Xu, M. Markatou, and C. Friedman, “Automated acquisition of disease–drug knowledge from biomedical and clinical documents: An initial study,” *Journal of the American Medical Informatics Association : JAMIA*, vol. 15, pp. 87–98, 10 2007.

[42] K. Syed, W. C. Sleeman, M. P. Hagan, J. R. Palta, R. Kapoor, and P. Ghosh, “Automatic incident triage in radiation oncology incident learning system,” *Healthcare (Basel, Switzerland)*, vol. 8, pp. 272–, 8 2020.

[43] S. Jonnalagadda, A. K. Adupa, R. Garg, J. Corona-Cox, and S. J. Shah, “Text mining of the electronic health record: An information extraction approach for automated identification and subphenotyping of hfpef patients for clinical trials,” *Journal of cardiovascular translational research*, vol. 10, pp. 313–321, 6 2017.

[44] S. J. Berkowitz, D. Kwan, T. C. Cornish, E. L. Silver, K. S. Thullner, A. Aisen, M. M. Bui, S. D. Clark, D. A. Clunie, M. Eid, D. J. Hartman, K. Ho, A. Leontiev, D. M. Luviano, P. E. O’Toole, A. V. Parwani, N. S. Pereira, V. Rotemberg, D. J. Vining, C. M. Gaskin, C. J. Roth, and L. R. Folio, “Interactive multimedia reporting technical considerations: Himss-siim collaborative white paper.,” *Journal of digital imaging*, vol. 35, pp. 817–833, 8 2022.

[45] Q. Zhu, H. Liu, C. G. Chute, and M. J. Ferber, “Ehr based genetic testing knowledge base (igtkb) development.,” *BMC medical informatics and decision making*, vol. 15, pp. 1–11, 11 2015.

[46] R. Zhang, S. V. S. Pakhomov, E. G. Arsoniadis, J. T. Lee, Y. Wang, and G. B. Melton, “Detecting clinically relevant new information in clinical notes across specialties and settings.,” *BMC medical informatics and decision making*, vol. 17, pp. 68–68, 7 2017.

[47] A. V. Deokar and J. Tao, “Semantics-based event log aggregation for process mining and analytics,” *Information Systems Frontiers*, vol. 17, pp. 1209–1226, 6 2015.

[48] J. A. Pruneski, A. Pareek, B. U. Nwachukwu, R. K. Martin, B. T. Kelly, J. Karlsson, A. D. Pearle, A. M. Kiapour, and R. J. Williams, “Natural language processing: using artificial intelligence to understand human language in orthopedics.,” *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA*, vol. 31, pp. 1203–1211, 12 2022.

[49] J. Zhang, A. Can, P. M. R. Lai, S. Mukundan, V. M. Castro, D. Dligach, S. Finan, S. Yu, V. S. Gainer, N. A. Shadick, G. Savova, S. N. Murphy, T. Cai, S. T. Weiss, and R. Du, “Age and morphology of posterior communicating artery aneurysms,” *Scientific reports*, vol. 10, pp. 11545–, 7 2020.

[50] K. D. Mandl, D. J. Gottlieb, J. C. Mandel, V. Ignatov, R. Sayeed, G. Grieve, J. Jones, A. Ellis, and A. Culbertson, “Push button population health: The smart/hl7 fhir bulk data access application programming interface.,” *NPJ digital medicine*, vol. 3, pp. 1–9, 11 2020.

[51] S. V. Wang, O. V. Patterson, J. J. Gagne, J. S. Brown, R. Ball, P. Jonsson, A. Wright, L. Zhou, W. G. Goettsch, and A. Bate, “Transparent reporting on research using unstructured electronic health record data to generate ’real world’ evidence of comparative effectiveness and safety.,” *Drug safety*, vol. 42, pp. 1297–1309, 8 2019.

[52] S. M. Meystre, F. J. Friedlin, B. R. South, S. Shen, and M. H. Samore, “Automatic de-identification of textual documents in the electronic health record: a review of recent research,” *BMC medical research methodology*, vol. 10, pp. 70–70, 8 2010.

[53] A. K. Saxena, “Evaluating the regulatory and policy recommendations for promoting information diversity in the digital age,” *International Journal of Responsible Artificial Intelligence*, vol. 11, no. 8, pp. 33–42, 2021.

[54] J. L. Leevy, T. M. Khoshgoftaar, and F. Villanustre, “Survey on rnn and crf models for de-identification of medical free text,” *Journal of Big Data*, vol. 7, pp. 1–22, 9 2020.

[55] G. Trifirò, J. Sultana, and A. Bate, “From big data to smart data for pharmacovigilance: The role of healthcare databases and other emerging sources,” *Drug safety*, vol. 41, pp. 143–149, 8 2017.

[56] P. LePendu, S. Iyer, C. Fairon, and N. H. Shah, “Annotation analysis for testing drug safety signals using unstructured clinical notes,” *Journal of biomedical semantics*, vol. 3, pp. 1–12, 4 2012.

[57] K. S. Hughes, J. Zhou, Y. Bao, P. Singh, J. Wang, and K. Yin, “Natural language processing to facilitate breast cancer research and management.,” *The breast journal*, vol. 26, pp. 92–99, 12 2019.

[58] M. Ferrara, G. Franchini, M. Funaro, M. Cutroni, B. Valier, T. Toffanin, L. Palagini, L. Zerbinati, F. Folesani, M. B. Murri, R. Caruso, and L. Grassi, "Machine learning and non-affective psychosis: Identification, differential diagnosis, and treatment.," *Current psychiatry reports*, vol. 24, pp. 925–936, 11 2022.

[59] Y. Zhang, T. Cai, S. Yu, K. Cho, C. Hong, J. Sun, J. Huang, Y.-L. Ho, A. N. Ananthakrishnan, Z. Xia, S. Y. Shaw, V. S. Gainer, V. M. Castro, N. Link, J. Honerlaw, S. Huang, D. R. Gagnon, E. W. Karlson, R. M. Plenge, P. Szolovits, G. Savova, S. Churchill, C. J. O'Donnell, S. N. Murphy, J. M. Gaziano, I. S. Kohane, T. Cai, and K. P. Liao, "High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (phecap)," *Nature protocols*, vol. 14, pp. 3426–3444, 11 2019.

[60] R. Gauriau, B. C. Bizzo, D. S. Comeau, J. M. Hillis, C. P. Bridge, J. K. Chin, J. Pawar, A. Pourvaziri, I. Sesic, E. Sharaf, J. Cao, F. T. C. Noro, W. F. Wiggins, M. T. Caton, F. Kitamura, K. J. Dreyer, J. F. Kalafut, K. P. Andriole, S. R. Pomerantz, R. G. Gonzalez, and M. H. Lev, "Head ct deep learning model is highly accurate for early infarct estimation.," *Scientific reports*, vol. 13, pp. 189–, 1 2023.

[61] V. Joshi, P. Mitra, and S. Bose, "Multi-modal multi-head self-attention for medical vqa," *Multimedia Tools and Applications*, vol. 83, pp. 42585–42608, 10 2023.

[62] A. Tang, J. Pawar, C. P. Bridge, R. King, S. Kakarmath, M. B. Harris, and B. Khurana, "Traumatic cervical spine fracture patterns on ct: a retrospective analysis at a level 1 trauma center.," *Emergency radiology*, vol. 28, pp. 965–976, 6 2021.

[63] C. J. Harrison and C. Sidey-Gibbons, "Machine learning in medicine: a practical introduction to natural language processing.," *BMC medical research methodology*, vol. 21, pp. 158–, 7 2021.

[64] P. Gloor, A. F. Colladon, and F. Grippa, "Measuring ethical behavior with ai and natural language processing to assess business success.," *Scientific reports*, vol. 12, pp. 10228–, 6 2022.

[65] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Context-aware query performance optimization for big data analytics in healthcare," in *2019 IEEE High Performance Extreme Computing Conference (HPEC-2019)*, pp. 1–7, 2019.

[66] N. Oshimori, X. Li, M. Ohsugi, and T. Yamamoto, "Cep72 regulates the localization of key centrosomal proteins and proper bipolar spindle formation.," *The EMBO journal*, vol. 28, pp. 2066–2076, 6 2009.

[67] N. Afzal, V. P. Mallipeddi, S. Sohn, H. Liu, R. Chaudhry, C. G. Scott, I. J. Kullo, and A. M. Arruda-Olson, "Natural language processing of clinical notes for identification of critical limb ischemia.," *International journal of medical informatics*, vol. 111, pp. 83–89, 12 2017.

[68] A. Jagannatha, F. Liu, W. Liu, and H. Yu, "Overview of the first natural language processing challenge for extracting medication, indication, and adverse drug events from electronic health record notes (made 1.0)," *Drug safety*, vol. 42, pp. 99–111, 1 2019.

[69] O. Morin, M. Vallières, S. Braunstein, J. B. Ginart, T. Upadhaya, H. C. Woodruff, A. Zwanenburg, A. Chatterjee, J. Villanueva-Meyer, G. Valdes, W. S. Chen, J. C. Hong, S. S. Yom, T. D. Solberg, S. Löck, J. Seuntjens, C. C. Park, and P. Lambin, "An artificial intelligence framework integrating longitudinal electronic health records with real-world data enables continuous pan-cancer prognostication," *Nature cancer*, vol. 2, pp. 709–722, 7 2021.

[70] P. Deshpande, A. Rasin, J. Son, S. Kim, E. T. Brown, J. D. Furst, D. Raicu, S. M. Montner, and S. G. Armato, "Ontology-based radiology teaching file summarization, coverage, and integration," *Journal of digital imaging*, vol. 33, pp. 797–813, 4 2020.

[71] R. M. Cronin, D. Fabbri, J. C. Denny, S. T. Rosenbloom, and G. P. Jackson, "A comparison of rule-based and machine learning approaches for classifying patient portal messages," *International journal of medical informatics*, vol. 105, pp. 110–120, 6 2017.

[72] M. Muniswamaiah, T. Agerwala, and C. C. Tappert, "Integrating polystore rdbms with common in-memory data," in *2020 IEEE International Conference on Big Data (Big Data)*, pp. 5762–5764, IEEE, 2020.

[73] R. Avula, "Addressing barriers in data collection, transmission, and security to optimize data availability in healthcare systems for improved clinical decision-making and analytics," *Applied Research in Artificial Intelligence and Cloud Computing*, vol. 4, no. 1, pp. 78–93, 2021.

[74] D. L. Mowery, B. E. Chapman, M. Conway, B. R. South, E. Madden, S. Keyhani, and W. W. Chapman, "Extracting a stroke phenotype risk factor from veteran health administration clinical reports: an information content analysis," *Journal of biomedical semantics*, vol. 7, pp. 26–26, 5 2016.

[75] W. A. Zogaan, I. Mujhid, J. C. S. Santos, D. Gonzalez, and M. Mirakhorli, "Automated training-set creation for software architecture traceability problem," *Empirical Software Engineering*, vol. 22, pp. 1028–1062, 11 2016.

[76] L. Campillos-Llanos, A. Valverde-Mateos, A. Capllonch-Carrión, and A. Moreno-Sandoval, "A clinical trials corpus annotated with umls entities to enhance the access to evidence-based medicine," *BMC medical informatics and decision making*, vol. 21, pp. 69–69, 2 2021.

[77] K. Ronen, E. Grant, C. Copley, T. Batista, and B. L. Guthrie, "Peer group focused ehealth strategies to promote hiv prevention, testing, and care engagement," *Current HIV/AIDS reports*, vol. 17, pp. 557–576, 8 2020.

[78] G. Caldarelli, R. D. Nicola, M. Petrocchi, M. Pratelli, and F. Saracco, "Flow of online misinformation during the peak of the covid-19 pandemic in italy," *EPJ data science*, vol. 10, pp. 34–34, 7 2021.

[79] A. Sharma, M. Witbrock, and K. Goolsbey, "Controlling search in very large commonsense knowledge bases: a machine learning approach," *arXiv preprint arXiv:1603.04402*, 2016.

[80] E. Wu and T. A. Holly, "Nuclear cardiology reporting: Leaving an impression.," *Journal of nuclear cardiology : official publication of the American Society of Nuclear Cardiology*, vol. 26, pp. 1886–1887, 6 2018.

[81] H. Jupalle, S. Kouser, A. B. Bhatia, N. Alam, R. R. Nadikattu, and P. Whig, "Automation of human behaviors and its prediction using machine learning," *Microsystem Technologies*, vol. 28, pp. 1879–1887, 6 2022.

[82] A. N. Ananthakrishnan, T. Cai, G. Savova, S. C. Cheng, P. Chen, R. G. Perez, V. S. Gainer, S. N. Murphy, P. Szolovits, Z. Xia, S. Y. Shaw, S. Churchill, E. W. Karlson, I. S. Kohane, R. M. Plenge, and K. P. Liao, "Improving case definition of crohn's disease and ulcerative colitis in electronic medical records using natural language processing: a novel informatics approach.," *Inflammatory bowel diseases*, vol. 19, no. 7, pp. 1411–1420, 2013.

[83] F. Xie, D. Getahun, V. P. Quinn, T. M. Im, R. Contreras, M. J. Silverberg, T. C. Baird, R. Nash, L. Cromwell, D. W. Roblin, T. Hoffman, and M. Goodman, "An automated algorithm using free-text clinical notes to improve identification of transgender people," *Informatics for health & social care*, vol. 46, pp. 18–28, 11 2020.

[84] A. Nie, A. M. Zehnder, R. L. Page, Y. Zhang, A. L. Pineda, M. A. Rivas, C. Bustamante, and J. Zou, "Deeptag: inferring diagnoses from veterinary clinical notes.," *NPJ digital medicine*, vol. 1, pp. 60–, 10 2018.

[85] E. Ojuok, A. Uppuluri, P. D. Langer, M. A. Zarbin, L. Thangamathesvaran, and N. Bhagat, "Predictive factors of enucleation after open globe injuries," *Graefe's archive for clinical and experimental ophthalmology = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie*, vol. 259, pp. 247–255, 6 2020.

[86] C. V. Asche, B. Seal, K. H. Kahler, E. Oehrlein, and M. G. Baumgartner, "Evaluation of healthcare interventions and big data: Review of associated data issues.," *PharmacoEconomics*, vol. 35, pp. 759–765, 5 2017.

[87] S. D. Nelson, C. C. Lu, C. C. Teng, J. Leng, G. W. Cannon, T. He, Q. Zeng, A. Halwani, and B. C. Sauer, "The use of natural language processing of infusion notes to identify outpatient infusions.," *Pharmacoepidemiology and drug safety*, vol. 24, pp. 86–92, 11 2014.

[88] B. Al-Sahab, A. Leviton, T. Loddenkemper, N. Paneth, and B. Zhang, "Biases in electronic health records data for generating real-world evidence: An overview.," *Journal of healthcare informatics research*, vol. 8, pp. 121–139, 11 2023.

[89] T. Bergquist, V. Pejaver, N. Hammarlund, S. D. Mooney, and S. J. Mooney, "Evaluation of the secondary use of electronic health records to detect seasonal, holiday-related, and rare events related to traumatic injury and poisoning," *BMC public health*, vol. 20, pp. 46–46, 1 2020.

[90] Mahmud, M. S. Kaiser, T. M. McGinnity, and A. Hussain, "Deep learning in mining biological data," *Cognitive computation*, vol. 13, pp. 1–33, 1 2021.

[91] V. M. Castro, D. Dligach, S. Finan, S. Yu, A. Can, M. M. Abd-El-Barr, V. S. Gainer, N. A. Shadick, S. N. Murphy, T. Cai, G. Savova, S. T. Weiss, and R. Du, "Large-scale identification of patients with cerebral aneurysms using natural language processing," *Neurology*, vol. 88, pp. 164–168, 12 2016.

[92] S. Steyaert, M. Pizurica, D. Nagaraj, P. Khandelwal, T. Hernandez-Boussard, A. J. Gentles, and O. Gevaert, "Multimodal data fusion for cancer biomarker discovery with deep learning.," *Nature machine intelligence*, vol. 5, pp. 351–362, 4 2023.

[93] J. R. Machireddy, "Automation in healthcare claims processing: Enhancing efficiency and accuracy," *International Journal of Science and Research Archive*, vol. 09, no. 01, pp. 825–834, 2023.

[94] U. Naseem, S. Thapa, Q. Zhang, L. Hu, J. Rashid, and M. Nasim, "Incorporating historical information by disentangling hidden representations for mental health surveillance on social media," *Social Network Analysis and Mining*, vol. 14, 12 2023.

[95] W. Boag, O. Kovaleva, T. H. McCoy, A. Rumshisky, P. Szolovits, and R. H. Perlis, "Hard for humans, hard for machines: predicting readmission after psychiatric hospitalization using narrative notes.," *Translational psychiatry*, vol. 11, pp. 32–32, 1 2021.

[96] L. Poulsen, M. M. Lorenzen, N. Simonsen, A. F. Klassen, M. R. Rose, C. B. Juhl, R. K. Støving, A. Andries, S. J. Cano, A. L. Pusic, and J. A. Sørensen, "Patient-reported outcomes in weight loss and body contouring surgery: the voice of 1831 patients using the body-q," *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*, vol. 28, pp. 1–190, 9 2019.