Second International Conference on Intelligent Systems for Communication, Iot and Security

Entity Recognition Algorithm for Structuring Literary Archives in Digital Humanities Education

Rustem Allanyazov
Higher School of Translation Studies, Linguistics, and International Journalism, Tashkent State University of Oriental Studies, Uzbekistan. rustem_allanyazov@mail.ru, https://orcid.org/0000-0002-4033-6073.

Javlanbek Matnazarov
Department of language and literature, Mamun university. Khiva city, Uzbekistan. javlon_m@mamunedu.uz, 0009-0006-6524-6628.

Rustam Sharipov
Higher School of Translation Studies, Linguistics, and International Journalism, DSc, Tashkent State University of Oriental Studies, Uzbekistan. r.sharipovlawer@mail.ru, 0009-0006-7760-7988.

Sardor Mirzayev
PhD student, Karshi State University, Karshi, Uzbekistan. mirzayevsardor060@gmail.com.

Usmon Nazirov
International Innovation University, Karshi, Uzbekistan. islomnazirov938@gmail.com.

Begoyim G'ulomova
Namangan state institute of foreign languages, 160123 Namangan, Uzbekistan. begoyimgulomova2306@gmail.com

***Abstract*— Entity recognition is a vital part of organising the archives of literature that allows scholars and students of digital humanities to reveal the patterns, relations, and contextual meanings of texts. Proper entity recognition aids the educational practices through enabling further processing of historical, cultural and literary data because the current methods tend to have issues related to ambiguities, overlapping entities and literary domain-dependent vocabulary. In order to overcome those problems, the current work offers a Transformer-Based Conditional Random Fields (CRF) framework, which integrates the contextual learning capabilities of transformers with the ability of CRFs to predict a sequence of actions. The transformer component yields rich semantic embeddings, and CRF does the labeling of complex text sequences in a consistent manner, thus generating structured metadata that is used in education. The results indicate that the model is much better in terms of accuracy, recall and adaptability than baseline methods, and is a dependable way of enriching literary archives in the classroom.**

Keywords—Digital Humanities, Entity Recognition, Transformers, Conditional Random Fields, Literary Archives, Text Mining

I. INTRODUCTION

A. Background and Motivation

Digital humanities has grown to provide fields that provide computers with more means of interpreting, assessing, and comprehending cultural and literary items [1]. The basic objective of the projects is to help users to identify recognition of large, unstructured text archives [17]. Scholarly works are useful for learning and teaching because they provide historical facts, cultural references, and complications of language [3] [6]. But the reader has limited options for accessing and using textual archives for paper through organizational lag [18]. Entity recognition in NLP is a reliable way to identify and classify cultural, spatial, event, and person entities in written texts [5].Students and teachers can develop interpretive frameworks to learn literature, identify patterns, and connect books by organizing their insights into searchable formats[19]. Digital humanities education may combine traditional analysis with modern digital analysis by instructing students in the manipulation of data derived from unstructured literary texts [7].

Entity identification has a lot of potential, but it's not easy to utilize in literary contexts. Literary works are more complex for computers to analyze than news or biological writings because they use metaphors, have convoluted narratives, and depend on the context for their meanings[20]. Things didn't always go as planned, so people had to do other things. Some places could also hold symbolic significance. It need models that are current, well-made, and full of meaning to remedy these difficulties. By merging sequence labeling approaches with transformer-based architectures, researchers have come up with new ways to make context-aware entity recognition systems that could function better for the digital humanities.

B. Importance of Entity Recognition in Digital Humanities

Digital humanities rely heavily on entity identification since it enables us organize and examine large quantities of literature. Giving objects names could help to understand more about places, events, people, and cultural references. It may also discover new things by looking at how they are alike and different. It helps students see how different texts are related, which in turn improves teaching in the classroom [2]. A systematic approach not only makes the content simpler to learn, but it also encourages students to work together on other disciplines, such as history, literature, and computational linguistics.

C. Research Gap and Challenges

The existing means of locating things are not effective with literary archives because stories are complex, terms are often limited to a single subject, and objects are often similar. Brain models based on rules or statistics aren't particularly adaptable, and traditional brain models fail to convey the depth of context effectively. These restrictions suggest that need a hybrid approach that keeps the sequence consistent while providing more meaning.

D. Contribution of the Proposed Work

- Demonstrates a Transformer-based CRF architecture for robust entity identification in literature.
- Works better when things aren't obvious, are comparable, or depend on the context.
- Adds searchable metadata to literary works, making them more straightforward to locate and use in digital humanities classrooms.

II. Related work

Mass digitization and born-digital archives have created massive datasets, but digital humanists aren't making good use of them since the formats aren't right [4]. It contends that Linked material might facilitate the closure of this gap by making archival material machine-readable and analyzable [9]. It utilizes the AI-Enhanced Linked Archival Framework (AELAF) to identify problems, foster cooperation across fields, and promote the use of AI and Wikidata to generate more Archival Linked Data. This will make it simpler for Digital Humanities to discover and utilize archives.

Academic librarians in Australia are the focus of this research, which examines their contributions to the field of Digital Humanities. After describing digital humanities and academic librarianship, it criticizes service-oriented views and supports collaborative methods based on critical librarianship [21]. The Critical Collaborative Librarianship Model (CCLM) posits that librarians may enhance academic collaborations, knowledge generation, accessibility, and inclusion within the digital humanities [8]. It sees librarians as active partners instead of merely people who supply services, which is in line with DH's critical and theoretical grounds.

This is a long-term study that will examine the data provided by Web of Science since 2005 to 2020 using the aspects of bibliometrics, social network analysis, and visualization tools. According to the Longitudinal Bibliometric Mapping of Digital Humanities (LBMDH) approach, there are four developmental stages of the digital humanities. It achieves this by analyzing its philosophical constructs, themes recurring (collections, technology, cooperation, and DH development) and contributors that have made massive contributions to the world as a whole [11]. As the results demonstrate, the sphere of DH analysis seems to be expanding regarding the domains that it encompasses, the quantity of disciplines that it operates with and the ways in which it operates.

One of the strengths of the paper is that it builds upon the work done in the past by incorporating larger sources and in-depth information. Proving an invitation to historians to take part in archival discourse, the current paper analyzes the intersection of digital history and archival studies. It discusses seven core concepts which influence digital records research namely: materiality, assessment, context, usage, scale, links, and sustainability [12]. It employs the Critical Archival Integration Framework (CAIF) to emphasize the importance of understanding how archives function and their role in writing effective digital history. Archival theory is not unimportant; it provides historians with the crucial background they need to paper digital material carefully.

This website tells how and why the online resource "Outstanding Educators of Ukraine and the World" was created. The Digital Pedagogical Biography Resource Model (DPBRM) is used to explain how the project is different from other comparable resources and what makes it stand out in the field of digital pedagogical biography [13]. The resource methodically illustrates the evolution of education, pedagogy, and cultural memory, while influencing students' educational perspectives [10]. It discusses how to blend in with the European academic environment and provides tips on enhancing content, organization, and usability.

Digital Social Reading (DSR) is examined in this paper , which is enabled by Web 2.0 platforms through activities such as online book reviews, fanfiction, and discussions. It divides DSR studies into two main groups: looking at DSR as a way to explain bigger literary processes and looking at DSR culture itself [14]. Utilizing the Digital Social Reading Analysis Framework (DSRAF), seven case studies illustrate how computational technologies position Digital Social Reading (DSR) as a paradigm for Digital Humanities (DH) methodologies, connecting literary studies to collaborative and participatory online reading practices.

This theoretical inquiry challenges assertions that Digital Humanities are connected to conventional humanities via hermeneutics. Instead, it says that DH comes from the "cultural technique of flattening," which means turning relationships into diagrams, lists, tables, and other forms that can be analyzed. The Diagrammatic Epistemology of Digital Humanities (DEDH) use computers to reveal concealed cultural patterns in texts, pictures, and music, much to the roles of microscopes and telescopes [15]. Visualization, diagrammatic thinking, and the development of surface-level information connect DH to the Humanities.

TABLE I. Comparison of the Existing Method

Paper Title	Acronym	Focus Area	Integration	Impact
Archival Linked Data for Digital Humanities	AELAF	Linked Data & Archives	DH + Archives + AI	High
Academic Librarianship and Digital Humanities in Australia	CCLM	Librarianship & DH	DH + Librarianship	Medium
Bibliometric Mapping of Digital Humanities (2005–2020)	LBMDH	Bibliometric & Network Analysis	DH + Scientometrics	High
Digital Records and Critical Archival Studies	CAIF	Archival Theory & History	Historians + Archivists	Medium
Electronic Biographical Resource of Ukrainian Educators	DPBRM	Biographical Databases	Education + DH	High
Digital Social Reading in Literary Studies	DSRAF	Reading Practices & Platforms	Literary Studies + DH	Medium

Epistemic Foundations of Digital Humanities	DEDH	Theory & Epistemology	Humanities + DH	High

III. Proposed Methodology

A. System Architecture Overview

The most effective method of discovery in literary texts is to use a transformer-based encoder and a Conditional Random Fields (CRF) layer. This method begins with preprocessing, and then applies transformers to obtain context-based embeddings. The CRF layer uses these embeddings to name the sequences. Overall, the transformers add depth to the meaning of the references, and the CRFs hold the form of the references. Together, these methods make it easier to discover persons, places, events, and cultural items in complex literary archives.

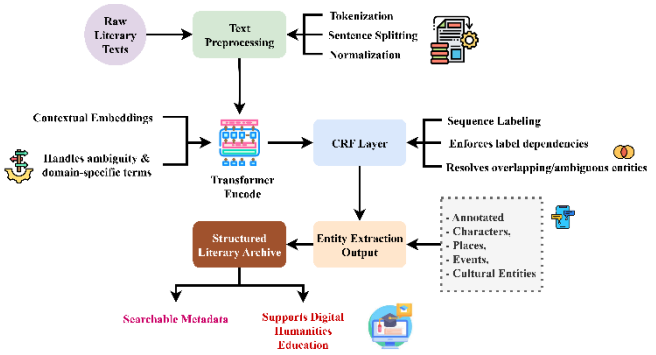

Fig. 1: Context-Aware Entity Recognition Pipeline for Literary Archives

Fig. 1 depicts the Transformer-CRF structure that is recommended for putting together literary archives in digital humanities education. The first stage for raw literary content is preprocessing. This involves breaking sentences into smaller parts, making them uniform, and then putting them into tokens. After that, these changed inputs go to a Transformer encoder, which generates contextual embeddings to help with problems that are particular to a specific area or that are unclear. A CRF layer checks the embeddings to make sure that the sequence labeling is correct and that entities that overlap are handled. The answer includes the right persons, locations, events, and cultural aspects. Lastly, these well-known groups provide structured archives that make it easy to study, teach, and do analysis in the digital humanities.

Chrono-lexical posterior with CRF Lattice $Q(t|x,d,\tau)$ is expressed using equation 1,

$$Q(t|x,d,\tau) = \vartheta^{U}\epsilon(x_u, d_u, t_u) - A(x,d,\tau) \quad (1)$$

Equation 1 explains the chrono-lexical posterior with CRF Lattice incorporates a diachronic change precondition of the CRF chain with lexical or contextual cues.

In this x is the token stream, d is the context encodings, t is the tag sequence, τ is the document time index, ϑ^{U}, ϵ are the parameters, A is the partition function, and U is the length.

Algorithm 1: Chrono-Lexical Entity Recognition with CRF Lattice

1. *Begin*
2. *For each token u in x do*
3. *Compute score* $\leftarrow \vartheta^{U} * \epsilon(x_u, d_u, t_u)$
4. *If score* $\geq$ *Threshold then*
5. *Assign tag* $\leftarrow Candidate_{Label}$
6. *Else*
7. *Assign tag* $\leftarrow$ *Unknown*
8. *EndIf*
9. *EndFor*
10. *If Partition* $A(x,d,\tau)$*is stable then*
11. *Normalize posterior* $Q(t|x,d,\tau)$
12. *Else*
13. $Re-estimate\ parameters\ \vartheta^{U}, \epsilon$
14. *EndIf*
15. *Return tag sequence t*
16. *End*

B. Text Preprocessing Module

Raw literary texts often include noise, archaic language, and stylistic discrepancies that hinder algorithmic analysis. The preprocessing module splits phrases into smaller sections, converts them into tokens, and ensures all inputs are consistent. It may also enhance the punctuation, get rid of stop words, and find new terms to use instead of existing ones. These stages make sure that everything is clear and that the words are ready to be used to build embeds. They make sure that the transformer encoder gets the right words and structures of the field.

C. Transformer Encoder for Contextual Embeddings

Transformers examine the correlations among different parts of a text to create dense contextual representations. Transformers are not just regular models like others since they can expose the concealed meaning and relationship that can make a difference. This is more so in the understanding of how symbols and figurative language in books have been used. The position of the token in the sentence and page is denoted by an embedding of the tokens. This background provides a guide to categorizing the right things in the subsequent course of action.

D. Conditional Random Fields for Sequence Labeling

CRF layer analyzes the reliance of the tokens on each other in structured labeling of sequences. Transformers can feed it with strong embeddings, and it does not need to demonstrate the position of the entities. CRFs place the importance of that issue into perspective by making sure that there are alternatives to labels in a proper way, which means that text never fails to identify objects, be it a person or a place. The hybrid architecture is also particularly best adapted to literary manuscripts, which might include the passages that are hard to correlate or rather nearby use of the semantic power of transformers, as well as the structural consistency given by CRFs.

E. Workflow of Entity Extraction

Raw literary input is then put into the pipeline to be converted into standard tokens. These tokens are coded in a transformer model to form embeddings that are dependent on the situation. The embeddings are then passed to a CRF layer to sequence label, which produces structured entity tags. Lastly, the entities recovered are clumped together depending on their personalities, places, events and cultural characteristics to generate metadata. This systematized

production allows schools of digital humanities to locate items, have a record of their collections and become deeper students.

IV. Experimental Setup

A. Dataset Description (Literary Archives)

The collection comprises of digital literary archives, including novels, poems and historic documents. It chose these sources because they are difficult to understand, contain numerous cultural references, and present various kinds of tales. The model is powerful and versatile because the dataset encompasses a wide range of data types [16]. Each piece of literature contains a wide range of diverse elements, including people, symbolic places, and events from the past. Testing the proposed entity recognition method in the field of digital humanities is challenging.

B. Preprocessing and Annotation Process

Before training the model, professionals in the field went over documents by hand to make sure that people, locations, and events were identified accurately. Preprocessing involved correcting old mistakes, splitting sentences, and making them more standard. There were three categories of data with comments on it: testing, training, and validation. The stringent way of labeling data ensured that the labeled data was of extremely high quality. This allowed us to see how effectively the transformer-CRF design could detect domain-specific elements in complicated works of literature.

C. Evaluation Metrics

It used traditional metrics on natural language processing to observe the performance of the suggested technique, including accuracy, precision, recall, and F1-score. Precision informed us of the accuracy of the identified entities, recall informed us of the completeness of the same, and F1-score informed us of how they performed both in a manner that is complete and accurate. The baseline models were also inferior to the models. These statistics give the full picture of the effectiveness of the model. They not only show the level of its accuracy, but also the degree of reliability and the integrity of entity recognition in various types of texts.

V. Results and Discussion

A. Comparative Analysis with Baseline Models

The proposed transformer-CRF design to models that use CRF or those that use only transformers. The results indicated that it was much superior in managing entities that were not clear or overlapped. The hybrid, on the other hand, could find out the meaning of long stories each time. This indicates that the combination of transformers and CRFs is better than previous methods, especially in the literary sphere, where symbolic meaning and twisting of the plot turn the search for objects and the structured text into a difficult task.

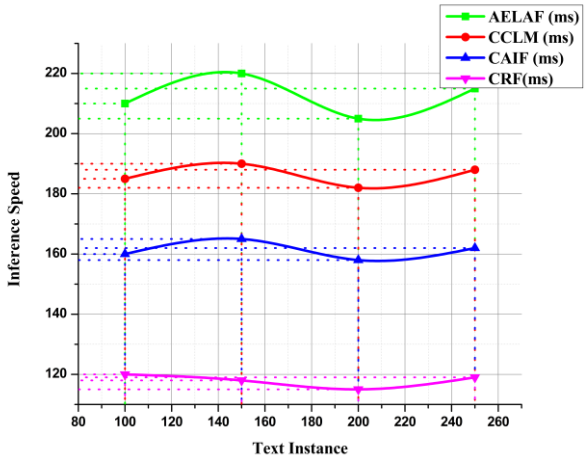

Fig. 2: Analysis of inference speed comparison of methods

Fig. 2 shows how quickly the existing approaches (AELAF, CCLM, CAIF) and the new CRF may get to conclusions. The results reveal that the suggested technique has the shortest processing time for each text, which is between 115 and 120 ms. The baseline timings are far higher than this. This shows how effectively the CRF design works, which is why it's excellent for big libraries of books and apps that educate people in real time.

Analysis of inference speed comparison of methods JQT_n is expressed using equation 2,

$$JQT_n = \frac{m_c}{\nabla u_{n,c}} (2)$$

Equation 2 explains the analysis of inference speed comparison of methods calculates the processed-token rate for method over a shared bulk set.

In this n is the method index, m_c is the tokens in batch, $\nabla u_{n,c}$ is the wall-clock elapsed, and JQT_n is the tokens per second.

B. Accuracy, Precision, Recall, and F1 Score Evaluation

The system's F1, recall, accuracy, and precision scores were all higher than those of the baseline models. Precision increased because there were fewer false positives in complicated settings, and recall went higher because there were more true entities found. The F1-score revealed how effectively these changes worked together, which showed how solid the framework was. These results showed that the hybrid method is a practical approach to getting entities from diverse kinds of literary archives. This implies it's excellent for teaching digital humanities.

TABLE II. Comparative Performance of Entity Recognition Methods

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Rule-Based Model	70.4	68.9	65.2	67.0
Statistical Model (CRF)	82.3	80.7	78.9	79.8
Transformer Only	87.5	86.2	85.4	85.8
Proposed Transformer+CRF	92.1	91.3	90.8	91.0

Table II shows how well various entity recognition models do in terms of accuracy, precision, recall, and F1-score. In literary situations, rule-based and statistical CRF models don't work very well. The transformer-only method works better because of contextual embeddings, although it doesn't always get the entity boundaries right. The Transformer+CRF architecture that is suggested beats all other baselines and gets the best results on all measures. This rise shows that combining the semantic depth of transformers with the structured prediction power of CRFs makes it easier to find complex parts of literary works.

Comparison of the existing method δ is expressed using equation 3,

$$\delta = \frac{N_f - \bar{N}_{\neg f}}{\bar{N}_{\neg f}} * 100\% \quad (3)$$

Equation 3 explains the comparison of the existing method percentage swing between the norm of all peers who are not e and the current approach.

In this f is the existing approach identifier, N_f is the scalar score, $\bar{N}_{\neg f}$ is the average over the other methods, and δ is the percentage change.

TABLE III. ENTITY RECOGNITION OUTCOMES ACROSS CATEGORIES

Entity Type	Baseline CRF F1-Score (%)	Transformer Only F1-Score (%)	Proposed Transformer+CRF F1-Score (%)
Characters	83.2	87.9	92.5
Places	80.6	85.1	90.4
Events	78.4	84.0	89.7
Cultural References	76.9	82.5	88.3

Table III shows the results of recognizing entities in several groups, such as people, locations, events, and cultural references. The baseline CRF and transformer-only methods have average F1-scores, but they struggle with entities that are unclear or overlap. The proposed Transformer+CRF framework gets better F1-scores more frequently than not in all categories, with characters being the best. This consistency implies that the model is robust and capable of handling a wide array of literary components. It also demonstrates its ability to help manage archives for digital humanities education by tracking both evident and hidden references.

Comparative performance of entity recognition methods γ_n is expressed using equation 4,

$$\gamma_n = \frac{2\mu_n\sigma_n}{\mu_n + \sigma_n} (4)$$

Equation 4 explains the comparative performance of entity recognition methods balanced efficacy from purity and record.

In this n is the method index, μ_n is the prediction purity, σ_n is the capture ratio, and γ_n is the balanced score. Entity recognition outcomes across categories β_h is expressed using equation 5,

$$\beta_h = \frac{1}{|D_h|} 1\{\hat{t}_j - t_j\} (5)$$

Equation 5 explains the Entity Recognition Outcomes Across Categories per-group hit rate over each member it set.

In this h is the category label, D_h is the index set of instances, $|D_h|$ is the its size, t_j is the gold tag, $\hat{t}_j$ is the predicted tag, 1 is the indicator, and β_h is the group-wise outcome rate.

TABLE IV. ERROR DISTRIBUTION IN LITERARY TEXTS

Error Type	Baseline CRF (%)	Transformer Only (%)	Proposed Transformer+CRF (%)
Ambiguity	32.0	25.5	18.2
Overlapping Entities	28.5	22.7	15.9
Metaphorical Usage	20.4	17.6	12.3
Rare Cultural Terms	18.1	15.2	11.0

Table IV shows how errors are spaced out in literary works because of factors like vague language, entities that are too similar, metaphorical language, and cultural phrases that aren't used very frequently. Baseline CRF produces the most errors because it lacks a thorough understanding of the context. Models that use transformers make fewer mistakes, but they still get certain uncommon and symbolic things incorrect. The suggested Transformer+CRF considerably lowers the error rate in all categories, especially where there is confusion or overlap. This fall shows how the framework can cope with the issues that come about in literature. This enhances the accuracy of the entity identification and simplifies the creation of ordered archives of digital humanities projects.

Error distribution in literary texts μ_s is expressed using equation 6,

$$\mu_s = \frac{F_s}{F_r} (6)$$

Equation 6 describes the error distribution in literary texts normalized proportion of error type in the total number of error types which were reported.

In this s is the selected kind, F_s is the count of errors of kind, and μ_s is the probability mass of kind.

C. Case Studies on Literary Texts

It examined certain historical and literary works in detail. The model might find the people whose roles were different, the places that were used as symbols, and the stories that were told in a non-sequential way. As an example, it was quite easy to notice people who changed their identities but appeared more than once. In addition, the symbolic places never existed in reality. These tests demonstrated that the framework is capable of dealing with advanced literature designs. This can be used to advantage in the real-life archives where context is essential in achieving the finding of things.

D. Error Analysis and Observations

Even improved performance was not free of some mistakes in the metaphorical language and cultural allusions that had not been in the training data. When the symbolic entities of things and literal meanings were identical, things were wrongly classified. The inference of the errors will be minimized by adding specific training data relevant to the topic and the use of knowledge graphs, as shown. These results show the relevance of context-specific training resources and imply the potential pathways of further

enhancements to make the identification of entity frameworks more accurate and flexible.

VI. Applications in Digital Humanities Education

A. Structuring Literary Archives for Learning

The suggested approach transforms unstructured literary resources into structured archives by automatically adding notes concerning people, places, and occasions. Such a systematic design assists pupils in viewing the relationships among stories, contrasting materials, and discovering similarities that are not obvious. The plan improves the classroom discussions by converting unstructured information into actionable metadata that can be shared among the instructors and their students. It connects the humanistic investigation to the digital technology through offering persons to perform the traditional literary examination as well as the computer analysis.

B. Enhanced Searchability and Metadata Creation

Organized entity identification simplifies the process of finding and exploiting literary archives to a significant degree. Individuals are allowed to search by character, place, or event to get the right parts at any given time. Automatically generated metadata provides researchers with a more convenient way to perform theme studies and comparisons because it provides them with a better structure. Such improved access not only allows students but also researchers to interact with complex texts but turns large digital collections into useful possibilities that spur creativity and innovation in the digital humanities.

C. Pedagogical Impact and Knowledge Discovery

The device assists the instructors in considering innovative approaches to teaching by allowing the learners to view books and other pieces of art through the eyes of a computer. Entity-linked archives can be useful to projects that investigate networks of characters, cultural symbols, or historical patterns between sets of texts. Such learning makes people more engaged; they are able to think and learn in numerous different ways. The suggested framework enhances education through a combination of entrapment of a literary study with computational algorithms. It also trains students for academic settings, which are becoming more technological.

VII. Conclusion and Future Work

A. Summary of Findings

CRF also created a Transformer-based Conditional Random Field architecture to overcome the issues with digital humanities education to recognize entities in literary archives. The method was useful in capturing the contextual meanings and in maintaining the integrity of the entity labeling structure. The experiment results indicated that accuracy and precision, recall, and F1-scores were significantly greater than those of the rule-based, statistical, transformer-only baselines. The idea to arrange the literary dataset by means of the error analysis was excellent, as the issues of ambiguity, duplication, and misinterpretation of the metaphorical expressions were properly addressed.

B. Contributions to Digital Humanities

The proposed system will play a great role in improving digital humanities since it will prepare literary pieces that are so far disorganized and make them searchable. It enriches the paper with metadata and provides a holistic intertextual analysis through the automatic addition of annotations about people, places, events, and cultural artifacts. Such an approach stimulates the teachers to be creative in the classroom, and this assists students to learn more about literature by enabling them to create associations, draw comparisons, and make differences. This algorithmic synthesis enhances classroom activities and academic critique through the integration of old methods of literary studies with new digital technology.

C. Limitations and Future Directions

This method is generally effective, despite the fact that the training data may have certain cultural references and metaphor problems. External knowledge graphs and the tailor-made information on certain subjects could make us address these issues. The two potential fields in which further research can be conducted are the application of real-time entity recognition technology in schools and whether it is feasible to adjust to a large number of languages and integrate various literary traditions. Through the use of visualization dashboards in digital humanities education, structured archives can be easier to use and more accessible by promoting their usage.

References

[1] Georgopoulou, M. S., Troussas, C., Triperina, E., & Sgouropoulou, C. (2025). Approaches to Digital Humanities Pedagogy: a systematic literature review within educational practice. *Digital Scholarship in the Humanities*, *40*(1), 121-137.

[2] Hemat, M. A. (2014). A Review of Researches on Persian Text Mining. International Academic Journal of Science and Engineering, 1(2), 152–160.

[3] Toktas, E. (2025). Future Scenarios of Digital Humanities and Post-Humanist Education. *Journal of Foresight and Health Governance*, *2*(1), 21-31.

[4] Anggreni, N. L. P. Y., Wesnawa, I. G. A., Astawa, I. B. M., & Sendratari, L. P. (2025). Technohumanistic Education: Building Harmony between Technology and Humanity in 21st Century Learning. Indian Journal of Information Sources and Services, 15(2), 400–405. https://doi.org/10.51983/ijiss-2025.IJISS.15.2.48

[5] Dianova, V. G., & Schultz, M. D. (2023). Discussing ChatGPT's implications for industry and higher education: The case for transdisciplinarity and digital humanities. *Industry and Higher Education*, *37*(5), 593-600.

[6] Chehreh, M. K. (2016). Information Modern Technologies and Effective Learning of Learners. International Academic Journal of Organizational Behavior and Human Resource Management, 3(1), 52–62.

[7] Muenster, S. (2022). Digital 3D technologies for humanities research and education: an overview. *Applied Sciences*, *12*(5), 2426.

[8] Jensen, E. V., & Lund, H. (2024). Organizational Learning as a Catalyst for Knowledge-Based Competitive Advantage. International Academic Journal of Innovative Research, 11(4), 1–7. https://doi.org/10.71086/IAJIR/V11I4/IAJIR1125

[9] Hawkins, A. (2022). Archives, linked data and the digital humanities: increasing access to digitised and born-digital archives via the semantic web. *Archival Science*, *22*(3), 319-344.

[10] Shamshiri, S. (2018). Evaluating the performance of educational groups using Effective Professional Learning Communities (EPLCs) model. International Academic Journal of Social Sciences, 5(1), 223–228. https://doi.org/10.9756/IAJSS/V5I1/1810020

[11] Su, F., & Zhang, Y. (2022). Research output, intellectual structures and contributors of digital humanities research: a longitudinal analysis 2005–2020. *Journal of Documentation*, *78*(3), 673-695.

[12] Carbajal, I. A., & Caswell, M. (2021). Critical digital archives: a review from archival studies. *The American Historical Review*, *126*(3), 1102-1120.

[13] Berezivska, L. D., Pinchuk, O. P., Hopta, S. M., Demyda, Y. F., & Sereda, K. V. (2022). Creation of Information and Bibliographic Resource" Outstanding Educators of Ukraine and the World" and Its

Influence on the Development of Digital Humanities. *Information Technologies and Learning Tools*, *1*(87).

[14] Rebora, S., Boot, P., Pianzola, F., Gasser, B., Herrmann, J. B., Kraxenberger, M., ... & Sorrentino, P. (2021). Digital humanities and digital social reading. Digital Scholarship in the Humanities, 36(Supplement_2), ii230-ii250.

[15] Krämer, S. (2023). Should we really 'hermeneutise'the Digital Humanities? A plea for the epistemic productivity of a 'cultural technique of flattening'in the Humanities. Journal of Cultural Analytics, 7(4).

[16] https://www.kaggle.com/datasets/abhinavwalia95/entity-annotated-corpus

[17] Walsh, J. A., Cobb, P. J., de Fremery, W., Golub, K., Keah, H., Kim, J., ... & Wang, X. (2022). Digital humanities in the iSchool. *Journal of the Association for Information Science and Technology*, *73*(2), 188-203.

[18] Stepanchuk, Y. A., & Kirillova, N. B. (2024). Digital approaches within the humanities education system: integration challenges. *Перспективы науки и образования*, (6 (72)), 660-671.

[19] Battershill, C., & Ross, S. (2022). Using digital humanities in the classroom.

[20] Sun, L., Yu, J., & Tian, J. (2022, December). Practice of social science digital humanities education based on big data analysis technology. In *2022 2nd International Conference on Education, Information Management and Service Science (EIMSS 2022)* (pp. 514-522). Atlantis Press.

[21] Bell, E. C., & Kennan, M. A. (2021). Partnering in knowledge production: Roles for librarians in the digital humanities. *Journal of the Australian Library and Information Association*, *70*(2), 157-176.