Annotaterm: A Fragment-Based Approach for Annotating Discontinuous Named Entities and Terms

Giorgio Maria Di Nunzio[1,*], Federica Vezzani[2]

[1]Department of Information Engineering, University of Padova, Via Gradenigo 6/b, 35131 Padova, Italy
[2]Department of Linguistic and Literary Studies, University of Padova, Via Elisabetta Vendramini, 13 35137 Padova, Italy

Abstract

The creation of annotated corpora is a fundamental prerequisite for training and evaluating models in named entity recognition and terminology work, such as term extraction. However, most available annotation environments restrict entities to contiguous text spans, limiting their applicability to domains where concepts are expressed discontinuously. Recent systematic reviews have shown that although numerous corpora contain discontinuous named entities, few tools natively support their annotation, and virtually none address term-level annotation explicitly.

In order to support this kind of activity, we first outline the possible requirements of a system architecture, its functional requirements derived from a comparative review of existing tools. Then, we introduce Annotaterm, a lightweight, web-based annotation environment implemented in R Shiny, designed to support both contiguous and non-contiguous multi-words annotation. Annotaterm adopts a fragment-based data model in which each textual fragment of a discontinuous entity is stored as an individual record linked to a shared identifier, enabling intuitive visualization and straightforward export to different serialization formats.

By providing a modular, open-source interface for the enrichment of domain texts with structured terminology, Annotaterm supports the creation of interoperable linguistic resources within digital libraries and research data repositories.

Keywords

Text annotation, Named Entity Recognition, Term extraction, Discontinuous entities

1. Introduction

Text annotation is a necessary activity in the creation of high-quality datasets for natural language processing (NLP), particularly in tasks such as Named Entity Recognition (NER), term extraction, and relation identification. Annotated corpora provide the empirical basis for training and evaluating supervised learning models, and are therefore essential for knowledge discovery and information extraction in digital libraries and domain-specific repositories.

Beyond their role in training NLP models, annotations have long been recognized as fundamental scholarly instruments in digital libraries and digital humanities. As also discussed by [1], annotations act as structured interpretative layers that connect texts to concepts, terminologies, and external knowledge resources, enabling reuse, interoperability, and long-term preservation. From this perspective, annotation tools are not merely technical utilities but form part of the scholarly infrastructure supporting semantic enrichment, knowledge organization, and research data curation. The ability to express complex phenomena, such as non-contiguous or composite terms, therefore directly impacts the quality and expressiveness of digital library resources.

While annotation practices for contiguous entities or terms (e.g., “waste management”) are well established, the annotation of discontinuous or non-contiguous multi-word terms remains an open challenge. Many domain-relevant expressions are not expressed as continuous text spans but rather as syntactically disjoint fragments. For instance, in the phrase “waste management and disposal,” the two multi-word terms “waste management” and “waste disposal” overlap on shared components and

IRCDL 2026: 22nd Conference on Information and Research Science Connecting to Digital and Library Science, February 19-20, 2026, Modena, Italy

*Corresponding author.
giorgiomaria.dinunzio@unipd.it (G. M. Di Nunzio); federica.vezzani@unipd.it (F. Vezzani)
0000-0001-7116-9338 (G. M. Di Nunzio); 0000-0003-2240-6127 (F. Vezzani)

cannot be represented as single contiguous spans. Existing annotation tools, however, typically assume that an entity corresponds to one continuous sequence of tokens, which limits their ability to represent such complex linguistic structures.

Recent systematic reviews confirm that the majority of available NER tools either do not support discontinuous annotations or require workarounds using multiple spans linked by relations [2, 3, 4]. Only a small subset of systems, such as TextAE [5], explicitly implement non-contiguous annotation modes. This gap poses difficulties for disciplines such as biomedicine, lexicography, terminology, and the digital humanities in general, where complex terminology and nested concepts frequently appear in discontinuous forms.

Furthermore, despite the extensive literature on NER annotation environments, there is little research devoted to term extraction as a distinct annotation task [6]. Term identification often requires additional contextual cues (e.g., domain specificity, morpho-syntactic patterns, or co-occurrence) that are rarely considered in general-purpose NER frameworks. Consequently, researchers and digital library curators face substantial barriers when attempting to build terminological resources and databases of Named Entities semi-automatically.

This paper addresses these limitations by (1) reviewing the current landscape of text-annotation tools with respect to discontinuous and multi-word entity handling (Section 2), (2) identifying functional and usability requirements specific to terminology and NER annotation in digital-library contexts (Section 3), and (3) presenting a prototype implementation developed in R Shiny[1] that enables users to select and visualize non-contiguous spans within a unified annotation interface (Section 4). The aim of Annotaterm is to facilitate the curation of structured linguistic and terminological resources that can enhance semantic indexing, metadata enrichment, and interoperability across research data infrastructures.

2. Background and Related Work

Text annotation remains a foundational process in the creation of high-quality datasets for named-entity recognition (NER) and term extraction. Advances in automatic NER have demonstrated that discontinuous multi-span entities can be effectively detected through deep learning models [7]. These approaches can automatically segment and link disjoint spans with high precision, confirming the representational feasibility of discontinuous mentions. However, no corresponding manual annotation environments exist to create or validate ground-truth datasets with the same level of expressiveness. The absence of user interfaces capable of selecting and managing multiple, non-adjacent spans limits the development of reliable training corpora for these models. The manual annotation separate word fragments remains technically and conceptually challenging.

This phenomenon, that appears across several languages, frequently happen in administrative, legal, and technical texts, where coordination, parenthetical insertions, or relative clauses break otherwise continuous entities. For instance, the sentence “Le président, réélu en 2017, Macron a annoncé ...” (President Macron, re-elected in 2017, announced...) illustrates the challenge of capturing non-adjacent spans in general-domain corpora, where the entity “Président Macron” is expressed discontinuously.

The systematic review by Alhassan et al. [2] provides the most comprehensive analysis to date of discontinuous NER in clinical and biomedical texts. However, their focus remains on corpus characteristics and computational modeling rather than the annotation tools themselves. No systematic comparison of annotation interfaces or user workflows is presented, leaving open the question of how discontinuous entities are best supported in practice.

Complementary to this corpus-based perspective, [3] present a broad comparative analysis of both semantic and non-semantic annotation tools. Their evaluation does not address discontinuous or multi-span entities. The review focuses on usability and interoperability rather than on representational expressiveness, leaving open the specific challenge of annotating non-contiguous terms and entities.

[1] https://shiny.posit.co/

This study underscores that while annotation ecosystems have matured, the integration of expressive linguistic modeling with user-friendly interfaces remains incomplete.

Although these reviews demonstrate the prevalence of discontinuous entities, few annotation tools natively support them. Most corpus-creation projects rely on ad-hoc software or flatten discontinuous spans into continuous segments, losing structural information. Tools like BRAT [8] and INCEpTION [9] permit partial workarounds through relational links, but their interfaces remain cumbersome for curators.

In particular, the BRAT Rapid Annotation Tool (BRAT) introduced a browser-based interface with standoff annotation storage and has been widely adopted for NER and dependency labeling tasks. Although BRAT's default data model only supports contiguous spans, modified forks have been used in biomedical projects to represent discontinuous entities by means of multiple offsets.

MedTator [10] implements multi-span annotation through the BioC format [11], allowing several offset pairs per entity. However, its interface does not yet enable users to interactively select and merge non-contiguous fragments, limiting its usability for general-purpose discontinuous term annotation

Metatron [12] provides a web-based environment for span, relation, and document-level annotation in the biomedical domain. However, like most contemporary tools, it restricts entity selection to contiguous spans. Non-contiguous mentions can only be represented indirectly through linked entities, as the platform currently lacks multi-span annotation support.

TextAE [5], a successor to the BRAT visualization engine, was specifically designed to overcome this limitation. Its 2020 extension introduced a non-contiguous annotation mode, allowing users to activate a multi-span selector and link disjoint tokens under a single entity identifier. Entities such as "waste management and disposal" can thus be captured as two non-adjacent spans associated with one annotation object. TextAE also supports export to the BioC format, which natively represents discontinuous spans via multiple location tags. However, TextAE data model remains entity-centric and offset-based, with limited support for term-oriented workflows and fragment-level aggregation. In contrast, term extraction, the identification of domain-specific lexical units rather than proper names or entities, has received comparatively little attention. Existing annotation tools are optimized for named entities and lack dedicated workflows for capturing terminological variants, morphological patterns, or multi-word expressions that recur across documents. This absence represents a critical gap for the construction of terminological databases and knowledge organization systems within digital libraries.

The proposed Annotaterm system operationalises this convergence by introducing a fragment-based data model capable of linking non-contiguous text segments under a unified entity identifier, supporting both entity recognition and terminology curation within digital-library infrastructures. Therefore, we would like to adopt a fragment-based representation in which each segment is treated as a first-class annotation unit, enabling simpler tabular exports and closer integration with terminology extraction pipelines.

3. Functional Requirements and Design Challenges

As shown in the previous section, handling discontinuous entities introduces substantial complexity at both the interface and data model levels. Annotation interfaces must provide intuitive visual cues for linking disjoint segments without confusing annotators or cluttering the text display. Back-end data structures must support multiple offset pairs per entity, requiring adaptations in export formats (e.g., BIO tagging schemes or standoff representations). Few general-purpose tools have addressed these technical challenges, leaving the task to specialized biomedical frameworks. While tools such as BRAT and INCEpTION offer partial workarounds for discontinuous entities through relation linking, these mechanisms impose a higher cognitive load on annotators, who must manage multiple entity identifiers and explicit relations. Instead, we would like to allow users to construct composite annotations incrementally through direct fragment selection, reducing interaction complexity.

The analysis of current annotation platforms reveals recurring usability and structural barriers that make the annotation of discontinuous and terminological entities difficult or even impossible. These

barriers stem from both interface constraints and data-model limitations, which together restrict the expressiveness of annotations in domain-specific corpora.

In the following section, we try to list a set of functional requirements for bridging this gap in a general-purpose annotation environment.

3.1. Annotation Interface Challenges

Most existing annotation tools assume that each entity corresponds to a single, contiguous span of text. This design assumption simplifies the annotation interface but fails to capture a wide range of linguistic patterns such as coordination, insertion, and nominal modification. For instance, the coordination in "waste management and disposal" contains two separate but conceptually related entities that share a component ("waste") of the multi-word term. In traditional tools, annotators must either select overlapping spans, which leads to ambiguous boundaries, or create two separate entities that cannot be explicitly linked as components of a discontinuous expression.

A second challenge concerns the visual representation of multi-span annotations. When entities consist of disjoint segments, the annotation interface must render the relationship between them in a way that is immediately comprehensible to the user.

TextAE addresses this by highlighting non-contiguous tokens with the same color and connecting them via subtle overlays, whereas INCEpTION and BRAT require external relation links that clutter the workspace. Hence, a more intuitive interaction model is needed for smooth annotation workflows.

3.2. Data Model and Schema Constraints

The internal data representation also imposes limitations on discontinuous annotations. Frameworks based on the BIO or IOB2 token labeling schemes cannot directly encode multiple disjoint spans for a single entity [13]. Similarly, relational standoff models such as those used in BRAT and INCEpTION associate entities with only one offset pair (*start, end*). Supporting discontinuous entities thus requires extending the schema to accept a set of offsets or a list of token indices per entity.

The BioC format [14] natively supports discontinuous annotations through multiple *<location>* elements within a single *<annotation>* block. This design enables multi-span entities to be represented and exchanged consistently across tools. However, while BioC can encode discontinuous mentions, most annotation environments relying on it (e.g., MedTator) lack user interfaces that allow interactive selection and linking of such non-contiguous fragments.

Term extraction introduces further challenges. Unlike named entities, terms are not necessarily proper nouns but may include multi-word expressions, abbreviations, or morphologically derived variants. A functional annotation tool should therefore allow for flexible category hierarchies, annotation of linguistic patterns (e.g., noun phrases or compounds), and the possibility to link term variants across documents.

3.3. Functional Requirements

Based on these observations, the following functional requirements are proposed for a system intended to support both named entity and terminological annotation:

1. Multi-span selection: The interface must allow annotators to select multiple non-contiguous spans and associate them with a single entity label.
2. Unified visualization: Discontinuous spans belonging to the same entity should be visually connected (e.g., color-coded or highlighted consistently) to minimize cognitive load.
3. Flexible data schema: The underlying model must store a list of character offsets or token indices per entity, compatible with export formats such as BioC or JSON-LD.
4. Terminology support: The system should provide dedicated entity categories for terms, including hierarchical labeling and support for variant linking.

5. Interoperability: Export and import mechanisms should preserve discontinuous spans across common formats (CoNLL, BioC) without loss of information.
6. Collaborative editing: Multiple annotators should be able to contribute to the same document set, with change-tracking and conflict-resolution features.
7. Lightweight deployment: Given the needs of digital-library environments, the tool should be web-accessible, easily deployable, and require minimal configuration.

These requirements lead to two broader research questions that guide the prototype design:

- How can a web-based annotation interface support the selection and management of discontinuous entities while maintaining simplicity and usability?
- What data model and export format best balance expressiveness, interoperability, and ease of integration with existing NLP pipelines?

The next section presents the proposed prototype, implemented in the R Shiny framework, which operationalizes these requirements into a minimal, open, and extensible web-based annotation interface.

4. Annotaterm: A Prototype in R Shiny

To demonstrate the feasibility of fragment-based discontinuous annotation, we developed a lightweight web prototype using the R Shiny framework. This prototype has been written with less than 300 lines of code and made available on Github. [2] It is also possible to test it online.[3]

The application is designed around three main components: (1) a text-display and selection interface implemented in HTML and JavaScript, (2) a reactive back-end that records term annotations and offset metadata, and (3) dynamic visualization and export of the resulting annotations.

4.1. System Architecture

Figure 1 illustrates the overall interface of the system. The prototype follows a client–server model in which all user interactions occur through a web browser while the R Shiny server manages the state of annotations and updates the visual display reactively.

Frontend. The user interface is implemented using Shiny's `fluidPage` layout. A central text display area (`div#text_display`) presents the document for annotation. A short JavaScript listener captures mouse selections and transmits the selected text to the R server via Shiny's reactive input channel `current_selection`. The sidebar contains controls for adding *simple* (contiguous) or *composite* (multi-segment) terms, together with tables summarizing the stored annotations.

Backend logic. The server component maintains an in-memory reactive data frame (`terms_rv`) holding all annotations. Each annotation record contains:

- `term_label` – the canonical label assigned to the term or entity;
- `type` – either `simple` or `composite`;
- `segments` – a character vector of the component text fragments; and
- `offsets` – a data frame specifying each occurrence.

The core function `get_offsets()` computes every occurrence of a selected segment within the current text. Instead of storing multiple offsets within one nested list, the system follows a *fragment-based representation*, where each discontinuous fragment is recorded as a separate row.

This design enables straightforward flattening into tabular structures for analysis or export, making the model compatible with data-frame operations and relational storage.

[2] https://github.com/gmdn/IRCDL2026
[3] https://shiny.dei.unipd.it/annotaterm/

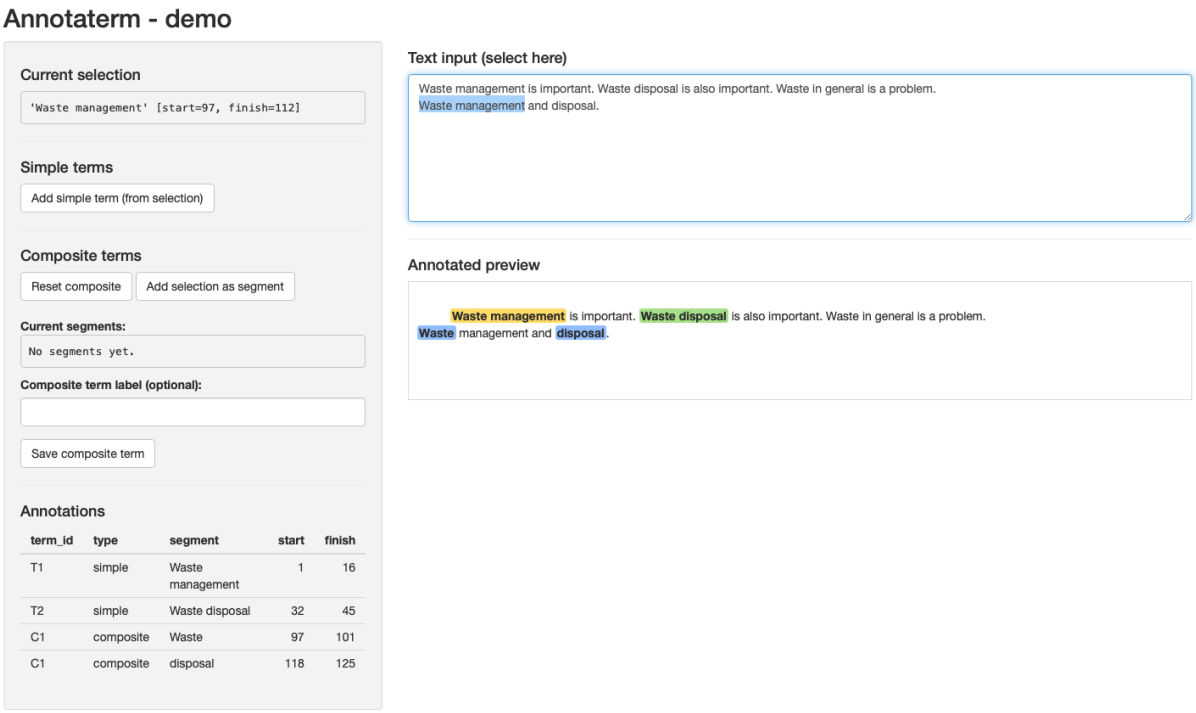

Figure 1: Screenshot of the annotation prototype *Annotaterm*.

Annotation workflow. Users select text directly within the display pane. A single click on *Add simple term* stores a contiguous annotation, while the *composite term* workflow allows multiple selections to be aggregated before saving under one label. Each new annotation triggers a recomputation of highlight overlays via reactive UI rendering. Highlighted fragments are visually linked through color and font weight using custom CSS classes (`.term-highlight`).

Data export and visualization. The `terms_table` output presents a flattened summary of all terms, showing label, type, segments, and offset coordinates. Because every fragment corresponds to one record, exports to CSV or JSON preserve the full discontinuous structure without nested objects or arrays. This approach provides direct compatibility with external tools for corpus analysis and model training.

4.2. Design Advantages

The use of R Shiny offers several benefits for rapid prototyping and integration within digital-library infrastructures:

- Lightweight deployment: the system runs entirely in R without additional web-server dependencies;
- Reactive updates: changes to annotations automatically refresh both the display and summary tables;
- Transparent data model: annotations are represented in tabular form, facilitating downstream processing in the tidyverse ecosystem; and
- Extensibility: the interface can be expanded with user authentication, collaborative features, or export adapters to formats such as BioC or BRAT standoff.

Overall, the prototype operationalizes some of the functional requirements defined in Section 3, demonstrating that discontinuous and terminological spans can be effectively managed using a fragment-based model and a minimal web framework.

4.3. Limitations and Current Work

Several features outlined in the requirements (Section 3) remain under development.

User management and collaboration. The current version operates as a single-user application with in-memory storage. Support for multiple annotators, user authentication, and session persistence will be necessary for collaborative corpus creation in larger projects. Future releases will integrate a lightweight authentication layer and database-backed persistence, enabling version control and inter-annotator agreement analysis. The current prototype does not yet support deletion or modification of existing annotations, nor validation constraints preventing duplicate or degenerate composite annotations. In future versions, composite annotations consisting of a single fragment will either be prevented or automatically normalized to simple annotations to avoid redundancy. These limitations are acknowledged and are being addressed through the introduction of annotation management controls and consistency checks.

Interoperability with external formats. At present, the data model is designed to support straightforward export to CSV or JSON. Planned extensions include export adapters for established annotation standards such as BioC allowing direct interoperability with existing annotation pipelines and automatic NER systems. Particular attention will be given to preserving discontinuous spans in these conversions, ensuring fidelity to fragment-based representations.

Visualization and usability improvements. Although the interface highlights all annotated fragments, composite terms are currently linked visually only by color. Work is ongoing to introduce more expressive visualization elements—such as bracket connectors, hover-based tooltips, or color-coded group identifiers—to improve user comprehension in densely annotated texts. Keyboard shortcuts and undo functionality are also planned to enhance annotation speed.

Automated suggestion and validation. Automatic pre-annotation, as found in tools like INCEpTION or MedTator, is not yet implemented. Future versions will incorporate machine-assisted term and entity suggestions derived from domain-specific lexicons or pretrained NER models. These will support semi-automatic workflows while preserving full manual control for curators.

4.4. Evaluation and Testing Plan

At the current stage, evaluation focuses on design adequacy and requirement coverage; quantitative usability and efficiency measurements are explicitly deferred to future work. In fact, to validate the usability and functional adequacy of the proposed annotation interface, we plan a series of evaluations using existing benchmark datasets and shared tasks devoted to terminology and entity annotation. The goal is twofold: (1) to measure the efficiency and accuracy of discontinuous and composite term annotation, and (2) to assess user experience and learnability in realistic curation scenarios.

Two ongoing community challenges provide suitable testbeds for the system:

- The Automatic Term Extraction – Italian (ATE-IT)[4] challenge focuses on the identification of domain-specific terms in Italian corpora [15]. It includes gold-standard datasets with human-validated multi-word terms, making it ideal for evaluating both contiguous and discontinuous term annotation.
- The Gut–Brain IE challenge[5] addresses complex entity and relation extraction in biomedical literature, including nested and cross-sentential phenomena [16, 17, 18]. This dataset provides a realistic stress test for the fragment-based annotation model, as many biomedical entities appear as discontinuous mentions.

[4] https://nicolacirillo.github.io/ate-it/

[5] https://hereditary.dei.unipd.it/challenges/gutbrainie/2025/

- The DEfinition and Term Extraction Challenge (DETECH),[6] focuses on automatic extraction of domain-specific terminology and generation of natural language definitions. It includes a term extraction subtask and a definition generation subtask within the biomedical domain.

We will recruit domain experts and graduate-level annotators to perform controlled annotation sessions on selected documents from these corpora. Each participant will be provided with the same unannotated text and instructed to label terms or named entities according to the challenge guidelines.

Performance will be assessed along three dimensions:

1. Annotation accuracy: comparison between user annotations and the gold standard evaluation metrics like using precision and recall.
2. Annotation time: mean time per entity and per document.
3. Usability feedback: subjective ratings collected via a short post-task questionnaire baseds.

To specifically test discontinuous annotation handling, participants will annotate a curated subset of examples containing coordination and multi-part terms. Metrics will be reported separately for contiguous and discontinuous entities to identify any differences in cognitive load or task completion time.

4.5. Quantitative and Qualitative Analyses

Quantitative evaluation will use standard corpus-based metrics, while qualitative inspection of error cases will focus on:

- boundary errors caused by partial segment selection;
- missing linkages between fragments belonging to the same conceptual entity; and
- visual or interaction issues reported by annotators.

Feedback will inform refinements to both the interface design (e.g., improved highlighting of linked segments) and the data export logic. Where possible, the annotated outputs will be compared with automatic extraction baselines from existing ATE-IT or GutBrainIE participants to assess compatibility and downstream usefulness.

5. Conclusions and Future Work

This paper addressed the persistent limitations of existing text annotation tools in handling discontinuous and multi-word terms, an issue that constrains the creation of high-quality corpora for named entity recognition and term extraction. Through a structured review of current platforms, we identified that most systems either ignore non-contiguous spans or rely on ad hoc relational workarounds. Term extraction, in particular, remains largely unsupported as an explicit annotation task despite its centrality in domain-specific information management.

To bridge this gap, we proposed a set of functional requirements for a new generation of annotation tools and presented a prototype implementation built with R Shiny. The prototype introduces a fragment-based data model, where each occurrence of a discontinuous term is stored as an independent record linked to a shared entity identifier. This approach provides both conceptual clarity and technical simplicity: it aligns with tabular data processing, facilitates visualization, and supports interoperability with common corpus formats. The interactive interface demonstrates that complex annotation tasks can be implemented within lightweight, open, and easily deployable frameworks suitable for digital-library infrastructures. We intentionally position Annotaterm as a feasibility prototype rather than a production-ready annotation environment; its role is to validate the fragment-based model and interaction paradigm prior to large-scale empirical evaluation.

[6] https://detech2026.dei.unipd.it/

In future work, we plan to perform systematic user evaluations using established community benchmarks such as ATE-IT, GutBrainIE, and DETECH. These studies will help assess not only the correctness and efficiency of discontinuous term annotation but also the usability of the proposed interface in real curation scenarios. Results from these experiments will guide further refinements to the interaction design, particularly in the visualization of linked fragments and in automatic quality-control feedback.

Beyond the immediate prototype, several extensions are envisaged. First, the integration of collaborative editing and version control would enable distributed annotation projects within research infrastructures. Second, adapting the export pipeline to support standardized formats such as BioC and RDF/JSON-LD will improve data interoperability across NLP workflows. Finally, coupling the annotation interface with machine-assisted term suggestions, drawing from domain embeddings or large language models, may substantially reduce manual effort while maintaining expert oversight.

By combining methodological clarity with open-source implementation, this work contributes a step toward more expressive, user-friendly annotation environments for the construction of terminological and named-entity datasets within digital libraries and related knowledge systems.

Acknowledgments

This work is partially supported by the HEREDITARY Project, as part of the European Union's Horizon Europe research and innovation programme under grant agreement No GA 101137074, and it is part of the initiatives of the Center for Studies in Computational Terminology (CENTRICO) of the University of Padua and in the research directions of the Italian Common Language Resources and Technology Infrastructure CLARIN-IT.

Declaration on Generative AI

During the preparation of this work, the author(s) used Chat-GPT-5.2 in order to: Grammar and spelling check. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

[1] E. Giovannetti, D. Albanesi, A. Bellandi, S. Marchi, M. Papini, F. Sciolette, Maia: An Open Collaborative Platform for Text Annotation, E-Lexicography, and Lexical Linking, Umanistica Digitale (2024) 27–52. doi:10.6092/issn.2532-8816/19705.

[2] A. Alhassan, V. Schlegel, M. Aloud, R. Batista-Navarro, G. Nenadic, Discontinuous named entities in clinical text: A systematic literature review, Journal of Biomedical Informatics 162 (2025) 104783. doi:10.1016/j.jbi.2025.104783.

[3] L. Colucci Cante, S. D'Angelo, B. Di Martino, M. Graziano, Text Annotation Tools: A Comprehensive Review and Comparative Analysis, in: L. Barolli (Ed.), Complex, Intelligent and Software Intensive Systems, Springer Nature Switzerland, Cham, 2024, pp. 353–362. doi:10.1007/978-3-031-70011-8_33.

[4] B. Jehangir, S. Radhakrishnan, R. Agarwal, A survey on Named Entity Recognition — datasets, tools, and methodologies, Natural Language Processing Journal 3 (2023) 100017. doi:10.1016/j.nlp.2023.100017.

[5] J. Lever, R. Altman, J.-D. Kim, Extending TextAE for annotation of non-contiguous entities, Genomics & Informatics 18 (2020). doi:10.5808/GI.2020.18.2.e15.

[6] K. Xu, Y. Feng, Q. Li, Z. Dong, J. Wei, Survey on terminology extraction from texts, Journal of Big Data 12 (2025) 29. doi:10.1186/s40537-025-01077-x.

[7] C. Corro, A Fast and Sound Tagging Method for Discontinuous Named-Entity Recognition, in: Y. Al-Onaizan, M. Bansal, Y.-N. Chen (Eds.), Proceedings of the 2024 Conference on Empirical

Methods in Natural Language Processing, Association for Computational Linguistics, Miami, Florida, USA, 2024, pp. 19506–19518. doi:10.18653/v1/2024.emnlp-main.1087.

[8] P. Stenetorp, S. Pyysalo, G. Topić, T. Ohta, S. Ananiadou, J. Tsujii, Brat: A Web-based Tool for NLP-Assisted Text Annotation, in: F. Segond (Ed.), Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, Avignon, France, 2012, pp. 102–107.

[9] J.-C. Klie, M. Bugert, B. Boullosa, R. Eckart de Castilho, I. Gurevych, The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation, in: D. Zhao (Ed.), Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations, Association for Computational Linguistics, Santa Fe, New Mexico, 2018, pp. 5–9.

[10] H. He, S. Fu, L. Wang, S. Liu, A. Wen, H. Liu, MedTator: A serverless annotation tool for corpus development, Bioinformatics 38 (2022) 1776–1778. doi:10.1093/bioinformatics/btab880.

[11] D. C. Comeau, R. T. Batista-Navarro, H.-J. Dai, R. Islamaj Doğan, A. Jimeno Yepes, R. Khare, Z. Lu, H. Marques, C. J. Mattingly, M. Neves, Y. Peng, R. Rak, F. Rinaldi, R. T.-H. Tsai, K. Verspoor, T. C. Wiegers, C. H. Wu, W. J. Wilbur, BioC interoperability track overview, Database 2014 (2014) bau053. doi:10.1093/database/bau053.

[12] O. Irrera, S. Marchesin, G. Silvello, MetaTron: Advancing biomedical annotation empowering relation annotation and collaboration, BMC Bioinformatics 25 (2024) 112. doi:10.1186/s12859-024-05730-9.

[13] E. F. Tjong Kim Sang, J. Veenstra, Representing Text Chunks, in: H. S. Thompson, A. Lascarides (Eds.), Ninth Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, Bergen, Norway, 1999, pp. 173–179.

[14] S.-Y. Shin, S. Kim, W. J. Wilbur, D. Kwon, BioC viewer: A web-based tool for displaying and merging annotations in BioC, Database 2016 (2016) baw106. doi:10.1093/database/baw106.

[15] N. Cirillo, G. M. Di Nunzio, F. Vezzani, ATE-IT at EVALITA 2026: Overview of the automatic term extraction italian testbed task, in: Proceedings of the Ninth Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (EVALITA 2026). In Press., CEUR Workshop Proceedings, Bari, Italy, 2026.

[16] M. Martinelli, G. Silvello, V. Bonato, G. M. Di Nunzio, N. Ferro, O. Irrera, S. Marchesin, L. Menotti, F. Vezzani, Overview of GutBrainIE@CLEF 2025: Gut-Brain Interplay Information Extraction, in: G. Faggioli, N. Ferro, P. Rosso, D. Spina (Eds.), Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2025), volume 4038 of *CEUR Workshop Proceedings*, CEUR, Madrid, Spain, 2025, pp. 65–98.

[17] A. Nentidis, G. Katsimpras, A. Krithara, M. Krallinger, M. R. Ortega, N. Loukachevitch, A. Sakhovskiy, E. Tutubalina, G. Tsoumakas, G. Giannakoulas, A. Bekiaridou, A. Samaras, G. M. Di Nunzio, N. Ferro, S. Marchesin, L. Menotti, G. Silvello, G. Paliouras, BioASQ at CLEF2025: The Thirteenth Edition of the Large-Scale Biomedical Semantic Indexing and Question Answering Challenge, in: C. Hauff, C. Macdonald, D. Jannach, G. Kazai, F. M. Nardini, F. Pinelli, F. Silvestri, N. Tonellotto (Eds.), Advances in Information Retrieval, Springer Nature Switzerland, Cham, 2025, pp. 407–415. doi:10.1007/978-3-031-88720-8_61.

[18] A. Nentidis, G. Katsimpras, A. Krithara, M. Krallinger, M. Rodríguez-Ortega, E. Rodriguez-López, N. Loukachevitch, A. Sakhovskiy, E. Tutubalina, D. Dimitriadis, G. Tsoumakas, G. Giannakoulas, A. Bekiaridou, A. Samaras, G. M. D. Nunzio, N. Ferro, S. Marchesin, M. Martinelli, G. Silvello, G. Paliouras, Overview of BioASQ 2025: The Thirteenth BioASQ Challenge on Large-Scale Biomedical Semantic Indexing and Question Answering, in: J. Carrillo-de-Albornoz, A. García Seco de Herrera, J. Gonzalo, L. Plaza, J. Mothe, F. Piroi, P. Rosso, D. Spina, G. Faggioli, N. Ferro (Eds.), Experimental IR Meets Multilinguality, Multimodality, and Interaction, Springer Nature Switzerland, Cham, 2026, pp. 173–198. doi:10.1007/978-3-032-04354-2_12.