Enhancing Biomedical Semantic Annotations through a Knowledge Graph-Based Approach

Asim Abbas[1], **Mutahira Khalid**[2], **Sebastian Chalarca**[1],
Fazel Keshtkar[1], **Syed Ahmad Chan Bukhari**[*1]

[1]Division of Computer Science, Mathematics and Science, Collins College of Professional Studies, St. John's University, Utopia,Parkway, Queens, 8000, NY, USA

[2]School of Electrical Engineering and Computer Science, NUST, H-12, Islamabad, Pakistan

abbasa@stjohns.edu, mkhalid.msit20seecs@seecs.edu.pk, sebastian.chalarca19@my.stjohns.edu, keshtkaf@stjohns.edu, bukharis@stjohns.edu

Abstract

An abundance of biomedical data is generated in the form of clinical notes, reports, and research articles available online. This data holds valuable information that requires extraction, retrieval, and transformation into actionable knowledge. However, this information has various access challenges due to the need for precise machine-interpretable semantic metadata required by search engines. Despite search engines' efforts to interpret the semantics information, they still struggle to index, search, and retrieve relevant information accurately. To address these challenges, we propose a novel graph-based semantic knowledge-sharing approach to enhance the quality of biomedical semantic annotation by engaging biomedical domain experts. In this approach, entities in the knowledge-sharing environment are interlinked and play critical roles. Authorial queries can be posted on the "Knowledge Cafe," and community experts can provide recommendations for semantic annotations. The community can further validate and evaluate the expert responses through a voting scheme resulting in a transformed "Knowledge Cafe" environment that functions as a knowledge graph with semantically linked entities. We evaluated the proposed approach through a series of scenarios, resulting in precision, recall, F1-score, and accuracy assessment matrices. Our results showed an acceptable level of accuracy at approximately 90%. The source code for "Semantically" is freely available at: https://github.com/bukharilab/Semantically

Keywords: Semantic annotation, Semantic Knowledge graph, Annotation recommendation, Annotation ranking, peer-to-peer recommendations

Introduction

An enormous amount of biomedical information, such as research papers, clinical notes, and biomedical reports, is published each year in biomedical research and clinical practices. Timely information transfer from the scientific research community to peer investigators and other healthcare professionals entails efficient techniques for acquiring biomedical publications. This continued increase in the biomedical field has led to various access-level challenges for practitioners and researchers [Abbas A 2022]. Due to the lack of machine-interpretable meta-data, this indispensable information in the contents on the web is still concealed from information retrieval and knowledge extraction search engines.Search engines, on the other hand, require metadata to properly index the content in a context-aware manner and to support ancillary operations like automated integration for meta-analysis [S. A. C. Bukhari 2017]. It would be ideal to incorporate machine-interpretable semantic annotations into biomedical content during the pre-publication stage (during first drafting), and to maintain them throughout online publication [Warren P 2008]. However, these processes are complicated and require deep technical and/or domain knowledge. Thus, a cutting-edge, publicly available framework for creating biomedical semantic content would be revolutionary. The basic elements of the semantic content authoring process are ontologies, annotators, and user interfaces (UI). Semantic annotators are designed to make it easier to tag/annotate the relevant ontology concepts, either manually, automatically, or in a hybrid way, using pre-defined terminologies. As a result, users create more semantically rich information compared to typical writing methods, such as using a word processor [K. Hasida 2003]. We divide the semantic annotators that are readily available into two groups. There are two types of annotators: 1) Non-biomedical and 2) Biomedical. The biomedical annotators can be further divided into a) general-purpose annotators for biomedicine, which assert to cover all biomedical subdomains, and b) use case-specific annotators for biomedicine, which are created for a specific subdomain or to annotate specific entities like genes and mutations in a given text. In contrast, the general purpose non-biomedical semantic annotators consolidate various technologies like machine learning, Natural Language Processing(NLP), semantic similarity algorithms, ontologies and graph manipulation techniques[Jovanovic 2014]. Biomedical annotators primarily employ term-to-concept matching with or without machine learning-based methods[Tseytlin 2016]. Though Machine learning is applied by biomedical annotators like NOBELE Coder [Tseytlin 2016], ConceptMapper [J.Jovanovi 2017], Neji[Campos 2013], and Open Biomedical Annotator[N.H.Shah 2009] to swiftly annotate the text. However, these annotators lack of strong disambiguation capabilities: the ability to recognize the relevant biomedical concept for a particular piece of text from a list of compet-

Copyright © 2023 by the authors. All rights reserved.

ing concepts. Whereas the NCBO annotator [Jonquet 2009] and MGrep services are somewhat sluggish, the RysannMd annotator attempts to balance annotation speed and accuracy. However, On the other hand, its knowledge base is constrained to certain UMLS (Unified Medical Language System) ontologies and does not fully cover all biomedical subdomains [Cuzzola 2017].

Owing to the aforementioned challenges and problem, We have designed "Semantically," a publicly available, robust, and interactive framework that enables people with varying degrees of expertise in the biomedical domain to create biomedical semantic information collaboratively. A critical research challenge in developing a robust biomedical semantic content editor is balancing speed and accuracy. Finding the appropriate semantic annotations in real-time during content authoring is extremely complicated because one semantic annotation is frequently available in multiple biomedical ontologies with different text or connotations. As a result, we propose a state-of-the-art socio-technical approach to build a biomedical semantic content authoring system that balances the speed and accuracy of the existing biomedical annotators while keeping the original author at every stage of the process. Similarly, we developed the "Semantic Knowledge Café" module to enable authors to receive real-time semantic annotation recommendations for explicit biomedical terminologies from a domain expert.

With the rise of social web applications such as Wikipedia, blogs and online discussion forums have improved information transfer by providing a flexible environment where users can generate and find their favorite content on their terms [Faisal 2019]. However, with the passage of time, online discussion forums accumulate a significant amount of content over time, which may raise questions about the users' reliability and the content's quality. A poor-quality answer in a discussion forum reveals the existence of unprofessional or incompetent members; therefore, a primary goal is to find experts or trustworthy users. The majority of currently used expert-ranking approaches consider fundamental features, such as the total number of responses a certain expert delivers, but ignore the quality and consistency of the expert's answer. Every entity in a social environment typically interacts with other entities and delivers some contribution, which results in dependencies between the entities. To address these issues, we have introduced a semantic knowledge-sharing approach through a knowledge graph in a socio-technical environment. We offer an explainable and interpretive view through a knowledge graph, such as domain expert profile raking, a domain expert confidence score to suggested semantic annotations, and upvotes and downvotes given by community users to the suggested semantic annotation from the domain expert to build authors' or users' decision confidence levels. Additionally, the "Semantically Knowledge Cafe" module enables authors or users to expedite the searching process toward quality semantic annotation for explicit terminology through a knowledge graph. The module also prevents and assists the authors avoid repeating queries or posts about explicit terminology semantic annotation..

The article is further structured into two main categories: Proposed Methodology and Results and Discussion.

Proposed Methodology

We proposed a Knowledge-graph based knowledge-sharing approach to enhance the biomedical content semantic annotation while keeping the original content creator in a loop. Additionally, this approach expedites semantic information searching and indexing through a knowledge graph before publication. Initially, the preliminary semantic annotation for explicit biomedical content is catered from openly online available biomedical ontologies such as UMLS[abbas 2019] and Bioportal[Jonquet 2009]. To further enhance and assign precise semantics to the biomedical content, we developed a "Semantic Knowledge Cafe" module, which provides peer-to-peer communication between authors and domain experts for accurate annotation recommendations. Similarly, other community members can validate the expert-recommended annotation through a voting mechanism. In the meantime, a knowledge graph is created that shows who submits (the author) a request for a recommendation for a valid semantic annotation on biomedical contents, who responds (the expert) to the give request and makes recommendation, and who votes (+ or -) to validate the recommended annotations (community users). As a result, three entities play a cornerstone role in constructing the coherently connected knowledge graph in this situation. Additionally, this knowledge graph clearly represents information regarding recommended annotations and domain expert credibility. Further, it encourages the community's members not to query or post repeatedly for explicit biomedical content semantic annotation. The proposed methodology is further illustrated in three sections: The Initial Level Semantic Annotation and Authoring, Socio-technical Annotations Recommendation, and Semantic Knowledge Sharing Through Knowledge-Graph as shown in Figure. 1.

Initial Level Semantic Annotation and Authoring

A biomedical annotator is an indispensable component and plays a rudimentary role in biomedical content semantic annotation or enrichment [Mbouadeu 2022]. These biomedical annotators leverage publicly accessible biomedical ontology repositories, such as Bioportal [Jonquet 2009] and UMLS [abbas 2019], to assist researchers in the biomedical community in assembling and tagging their data with ontology notions to enhance search engine information retrieval and indexing. However, semantic annotation and augmentation process takes time and requires expert curators. Therefore, we automate the process of assigning semantic annotation using our proposed solutions. To accomplish this, we evaluated the original text and annotated it with the appropriate biomedial ontology terms using NCBO Bioportal webservice resources [Jonquet 2009]. To begin, authors have two options: import pre-existing content from research papers, clinical notes, and biological reports, or start typing in the semantic text editor. Next, when the author clicks the "Annotate" button, the initial level of semantic is generated automatically, without technical expertise or abstraction. Our systems accept the author's free text as input for

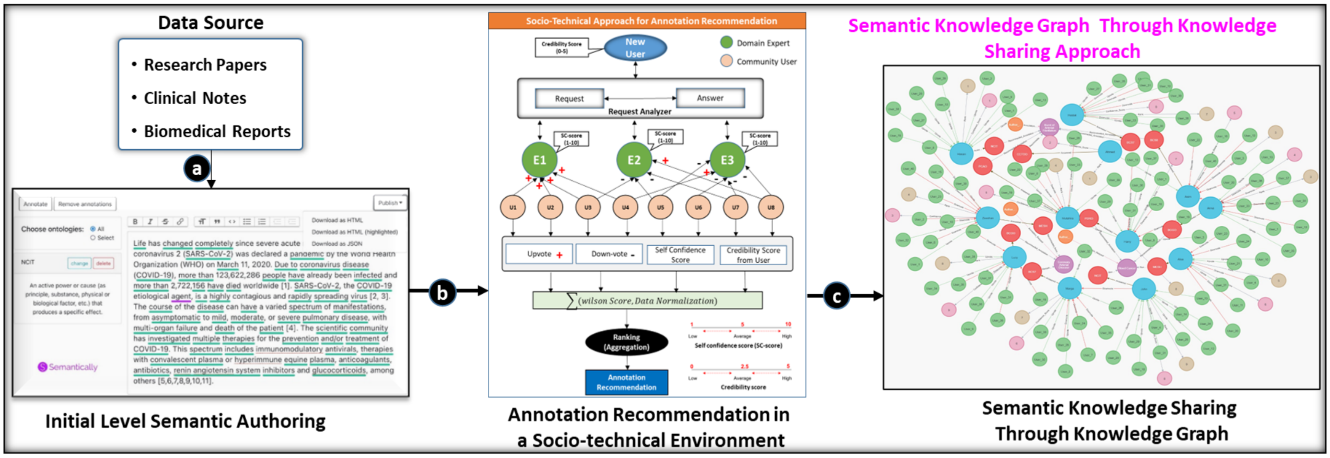

Figure 1: Proposed Methodology for Semantic Knowledge-sharing Through Knowledge-graph

semantic information extraction, which is then fed into a concept recognition engine. A string-matching method is applied by the concept recognition tool to identify the most appropriate acronyms, definitions, ontologies, and hyperlinks for particular biomedical terminologies based on the context. As a result, this semantic information appears in the annotation panel of our system for human interpretation and comprehension. Similarly, our system helps authors amend generated semantic information according to their comprehension, expertise, and experience. For instance, they could select a relevant ontology from the list, grab the initially suggested semantic annotations, and assign new relevant acronyms, vocabulary, and ontology to explicit terminology. While more experienced users may use advanced features to effectively control the semantic annotation and authoring process, individuals with a technical background can easily use a simplified interface.

Annotation Recommendation in a Socio-technical Environment

After obtaining initial semantic annotation, ”Semantically Knowledge Cafe” offers a cutting-edge socio-technical environment for the authors to interact and receive precise and high-quality semantic annotation recommendations from peer review. Assume the author is required to get accurate ontology annotations from experts for the biomedical terminology. The ”Semantically Knowledge Cafe” interface makes it easier for the author to write a query to a domain expert for annotation recommendation. Following the submission of the query, a new thread in the ”Semantically Knowledge Cafe” forum for expert E_i responses appeared, where $E_i = \{e_1, e_2, e_3, ...e_n\}$. Similarly, the ”Semantically Knowledge Cafe” provides the domain E_i an easy-to-use interface to speed up the response process to author queries. The NCBO ontology tree widget tool has been integrated into the interface to make it easier for the author to search for appropriate ontologies and their vocabulary. Similarly, the author can provide an additional explanation for why the following annotation is recommended. Regarding semantic annotation recommendations, the domain expert must provide a self-confidence score between 1 and 10 for their recommended semantic annotation. The expert confidence score renders how confident he/she is about recommended annotations to explicit author queries. When the expert E_i responds to the author's post, other community members $U_i = \{u_1, u_2, u_3, ...u_n\}$ can vote to approve or reject the expert-recommended annotation, as shown in Figure.1(b). We further assess the upvotes (+V) and downvotes (-V) and expert self-confidence score using the Wilson formula and data normalization process as shown in Equ.1 and Equ.5.

$$Wilson_{score} = \Bigg(\hat{p} + \frac{Z^2_{\alpha/2}}{2n} \pm Z_{\alpha/2} \sqrt{\Big[\hat{p}(1-\hat{p}) + Z^2_{\alpha/2}/4n\Big] /n}\Bigg) / \Big(1 + Z^2_{\alpha/2}/n\Big) \quad (1)$$

Where,

$$\widehat{p} = \left(\sum_{n=1}^{N} +V\right) / (n) \quad (2)$$

$$n = \sum_{i=0}^{N} \sum_{j=0}^{M} (+V_i, -V_j) \quad (3)$$

$$and, z_{\frac{\alpha}{2}}\ is\ the\ \left(1 - \frac{\alpha}{2}\right) quantile\ of\ the\ standard\ normal\ distribution \quad (4)$$

In Equ.1. $\hat{p}$ is the sum of the upvotes (+V) of a community members U_i to an expert's E_i response for an author request for correct annotation divided by the total number of votes (+V, -V). Similarly, n is the total of upvotes and downvotes (+V, -V), and α is the statistical confidence level: Pick 0.95 to have a 95% chance that our lower bound is accurate. The z-score in this function, however, remains constant. Ultimately, the author received an optimal annotation recommendation. Whereupon, recommendations are acquired to the author from an expert; the author has the option to accept or reject the recommended annotation with a credibility score between 0 and 5. Whenever an author accepts the recommended annotation, a credibility score between 1 and 5 is added to the expert's profile. Further, a credibility score of 0 is added to the expert's profile if the author rejects the recommendation. All features (upvotes (+V), downvotes (-V), expert self-confidence score, author credibility score) are equally contributed and coherently correlated with each other for the final annotation recommendation. The ultimate output processing (+v,-V) features are between 0 and 1. We normalize and transform the expert self-confidence score and author credibility score between 0 and 1 for the consistent and featured-dependent process applying Equ.5.

$$z_i = (x_i - min(x))/(max(x) - min(x)) * Q \quad (5)$$

Where z_i is the i^{th} normalized value in the dataset. Where x_i is the i^{th} value in the dataset, e.g., the user confidence score. Similarly *min(x)* is the minimum value in the dataset, e.g the minimum value between 1 and 10 is 1 and *max(x)* is the maximum value in the dataset, which is 10. Finally, *Q* is the maximum number wanted for a normalized data value, e.g. we normalized the confidence score between 0 and 1, and the maximum value between 0 and 1 for *Q* is 1.

Finally, all the SR-FS (Semantically Ranking Feature Score) for each expert E_i annotation recommendation is computed and aggregated using Equ.6.

$$Sr - Fs = \sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{k=0}^{p}(F_j, E_i, A_k) \quad (6)$$

$$final - score = argmax[\sum_{i=1}^{N}(Sr - Fs)] \quad (7)$$

Where F_j is the feature score for Expert E_i and Annotation A_k. The final decision or ranking happens based on maximum feature scoring gained by the Expert E_i response to the author's post or query see Equ.7.

Semantic Knowledge Sharing Through Knowledge Graph

Recently, knowledge graphs have gained immense attention because of providing machine-readable details, adding context and depth to data-driven AI techniques. Additionally, many big tech companies like Google have leveraged knowledge-graph to enhance information searching and retrieval. The core of a knowledge graph is its knowledge model, a collection of interconnected descriptions of concepts, entities, events, and relationships known as an ontology. This model offers a framework for taxonomies or statements. Each statement consists of a subject, predicate, and object, known as the "triple model," and each subject or object appears only once in the context of the other subjects and their relationships. In the proposed "Semantic Knowledge Cafe," four core entities, such as the author, domain expert, community users, and biomedical terminologies, are likely to perform coherently dependent and interconnectedly (see Figure). 2. Further, these entities have sub-entities or child nodes, such as the domain expert node, which contains the child node's expert profile rank, expert confidence score, upvotes, downvotes, and suggested semantic annotations for terminologies. These child nodes adequately convey the domain expert's credibility, ability, and trustworthiness to the authors and other community users. For Instance, Figure. 2 shows the knowledge graph from the socio-technical approach for the biomedical terminology "Coronary Arterial Disease." Now the author's or community members desire to thoroughly examine and validate the suggested or recommended ontology for the biomedical term "Coronary Arterial Disease". The domain expert (Zeeshan, Lucy and Marga) suggested relative ontologies("MESH","BCGO" and "BCS7") with some level of confidence score(3,9,6). After visualizing the graph, the author or community users can watch the domain expert profile scoring as Rank and domain expert self-confidence score, how many upvotes and downvotes the community provides to the expert suggested annotation, and who votes for who, how many explicit and implicit votes does a domain expert received. Combining all these features boosts an author or a community members confidence to select or reject the recommended semantic annotation by a domain expert. As shown in the Figure. 2, The domain expert(Lucy) has a high profile ranking, a confidence score, upvotes from community members and upvotes from other domain experts, and zero downvotes, Which automatically boosts the author or community member's confidence level to choose the appropriate ontology "BCGO" for the biomedical terminology "Coronary Arterial Disease".

Results and Discussion

A total of 30 people participated through social media call in the proposed system evaluation. We downloaded 30 biomedical-related research articles from Pubmed.org and assigned them randomly to the participants. After that, the participants execute individual documents on the system, and as a result, sets of initial annotations are obtained. Each participant is encouraged to post queries or questions on the "Semantically Knowledge Cafe" forum who has any concerns about the initial annotations. After which, a total of 645 total posts are generated. All the participants are encouraged to respond to the author's posts with a recommended annotation and a confidence score ranging from 1 to 10. Cumulatively 2845 answers are recorded from the domain expert E_i, and other users provide upvotes and downvotes to validate further the expert E_i recommended annotation. Where an average number of 6056 upvotes and 7942 downvotes are counted. After receiving the recommended annotation, the author can offer credibility scores between 0 and 5 to the recommended annotation and user profile.

A professor-level domain expert from academia is enlisted to manually assess the annotations produced by the socio-technical approach. We also calculated the IAA (Inter Annotator Agreement) among domain experts when determining the kappa (kappa) value. The Inter-Annotator Agreement (IAA), a measure of how well multiple annotators can make the same annotation decision for certain categories. It is an indispensable part of both the validation and reprehensibility of annotation results. We take two measurements into

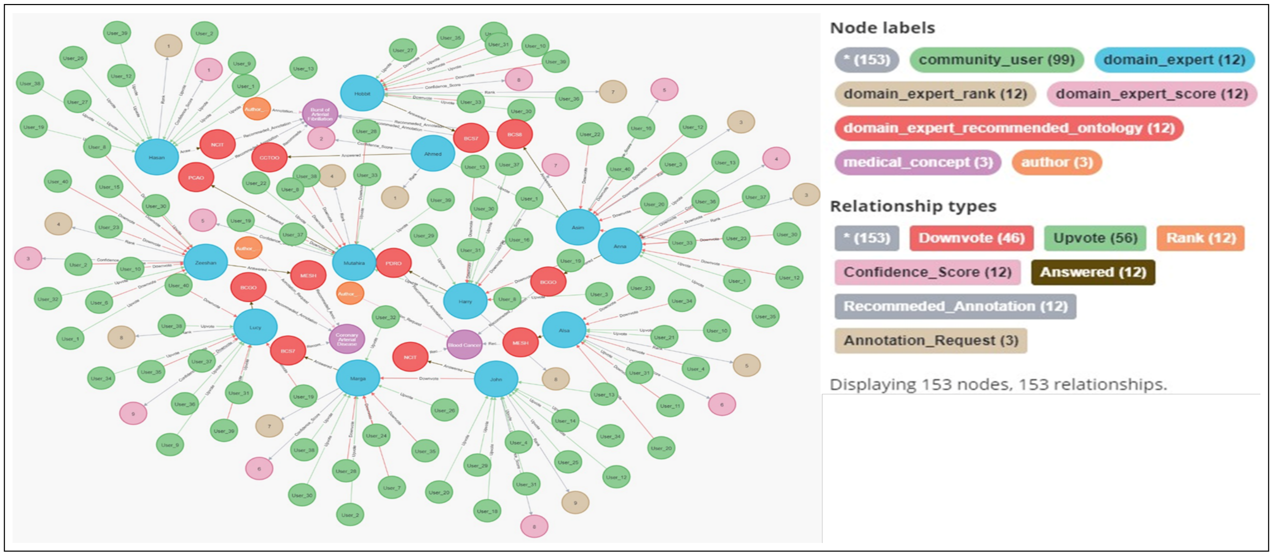

Figure 2: Semantic Annotation Recommendation Representation Through Knowledge Graph.

account for evaluation purposes:i) Cohen's Kappa and ii) Fleiss' Kappa.

$$\text{kappa}(\kappa) = \frac{P_o - P_e}{1 - P_e} \tag{8}$$

In Equ.8 p_o is the relative observed agreement among annotators, and p_e is the hypothetical probability of chance agreement. To interpret Cohen's kappa results, refer to the following study [Landis 1997]. However, if the annotators are in complete agreement then $\kappa = 1$ and perfect agreement. If there is no agreement among the annotators then $\kappa \leq 0$ or slight agreement as shown in Table.1. As per guidlines in the study [Landis 1997], we obtained almost perfect agreeemnt among three domain experts (Annotators) for our proposed socio-technical approach, where all agreement value is placed more than 90%. Similarly, perfect Cohen's and Fleiss kappa value of more than 85% is gained by domain experts for socio-technical annotation recommendation as shown in Table.1.
Similarly, We determined the system-level performance for a socio-technical approach taking into account four well-known evaluation matrices precision, recall, f1-score, and accuracy. To find out the system's effectiveness, we manually compared the socio-technical annotation results to those of domain experts. Consequently, the system has obtained nearly identical performance of 90% in terms of all four evaluation matrices for an annotation recommendation in a socio-technical setting, as illustrated in Figure. 3.

Further, the semantic annotation and the socio-technical related information is stored in the Neo4j database, which is an open-source, NoSQL, and, a native graph database. In the socio-technical environment, every entity(domain expert, community users, authors) is coherently connected with respect to the author query from the domain expert for annotation recommendation, the domain expert responds to the author query, and community users provide upvote and downvote to the expert response. So combining all these features build a social knowledge graph. The knowledge graph is very huge, and its complex to identify and find the correct and exact information. Considering this, the proposed framework supports the users to deeply analyze, seek and get an explainable view through a simplified knowledge graph regarding recommended annotation by the domain expert. Where community users and authors can query in "Semantic Knowledge Cafe" to search for precise and quality semantic annotation for the biomedical term "Blood Cancer". As shown in Figure 4, the three domain experts from community (Aisa, Anna and John), has responded with suggested annotation(MESH, BCGO and NCIT) along with a self-confidence score(6,4,8). The other community members U_i has responded with upvotes(+) and downvotes(-). Additionally, we visualized the community domain expert profile scoring(5,3, and 8). Based on this visualized information, now, the community users or authors confidently made their decision to choose the appropriate ontology annotation as "NCIT" for the biomedical term "Blood Cancer" which is recommended by domain expert "John". Because, "John" has gained the maximum number of upvotes and zero downvotes, high profile ranking, and confidence score related to "Anna" and "Aisa". Similarly, the framework facilitates the author or users to search and visualize the semantic annotation recommended by explicit domain experts such as "John Steve" in the knowledge graph form see Figure. 5. Whereas we also visualized the expert's profile score(5.32), recommended annotations(MS,NCIT, and MESH) for explicit biomedical terminologies(Burst of Arterial Fibrillation, Blood Cancer, and Coronary Arterial Disease), with a confidence score(8,4, and 9) and upvotes(10, 0, and 14) and downvotes(2,7, and 1) from community users. In this way, here too, all these visual features in the form of knowledge graphs provide obvious and deep technical knowledge regarding semantic annotation recommended by explicit ex-

% of Agreement				Cohen's and Fleiss Kappa Value		
	Expert1	**Expert2**	**Expert3**	**IAA between two Expert**	**Cohen's(κ)**	**Fleiss(κ)**
Expert1		**95.56%**	**95.3%**	Expert1, Expert2	**0.88**	**0.88**
Expert2	**95.56%**		**95.04%**	Expert1, Expert3	**0.87**	
Expert3	**95.3%**	**95.04%**		Expert2, Expert3	**0.87**	

Table 1: Inter Annotator Agreement(IAA) Results Among Domain Experts

perts see Figure. 5. In a nutshell, the knowledge and information that appeared in the form of a knowledge graph automatically enhance the author's and community members' decision confidence level to choose the precise and appropriate semantic annotation. Additionally, the knowledge graph has significantly improved the biomedical content semantic annotation searching process and eradicated duplicate requests for semantic annotation.

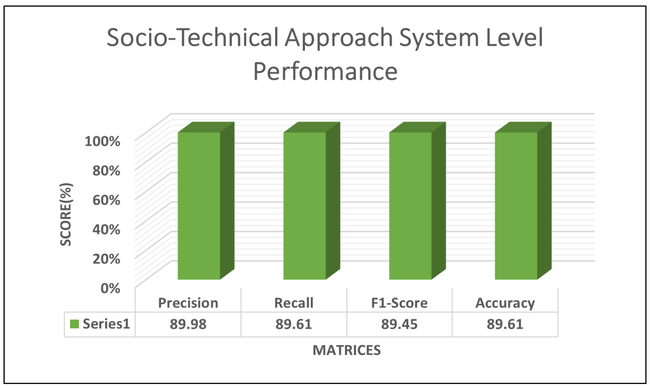

Figure 3: System Level Performance for Annotation Recommendation in Socio-technical Environment.

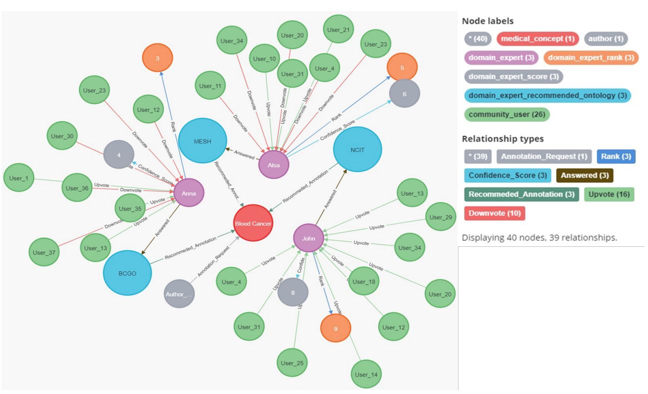

Figure 4: Semantic Annotation Recommendation Representation for Explicit Terminology through Knowledge Graph

Conclusion

The primitive objective of this study is to introduce a freely accessible framework that enables individuals at different levels of expertise in the biomedical domain to author biomedical semantic content collaboratively. To enhance the biomedical content semantic annotation accuracy, we proposed the socio-technical approach while keeping the original content creator(domain experts) in the loop in the entire process. The "Semantically Knowledge Cafe" is built for authors to ask questions from domain experts for explicit biomedical content semantic annotation. The domain expert can respond to the author's query and recommend a semantic annotation with a self-confidence score. Similarly, other community users are allowed to upvote and downvotes to validate further the domain experts suggested semantic annotation. Additionally, we introduce a knowledge graph-based approach to visualize the "Semantically Knowledge Cafe" information to enhance further the author's and "Semantically" users' decision-making level for choosing an appropriate and valid semantic annotation for explicit biomedical terminology. Consequently, the proposed approach dramatically increases the biomedical semantic content annotation accuracy and the searchability and shareability of semantic knowledge through knowledge graphs in a socio-technical environment.

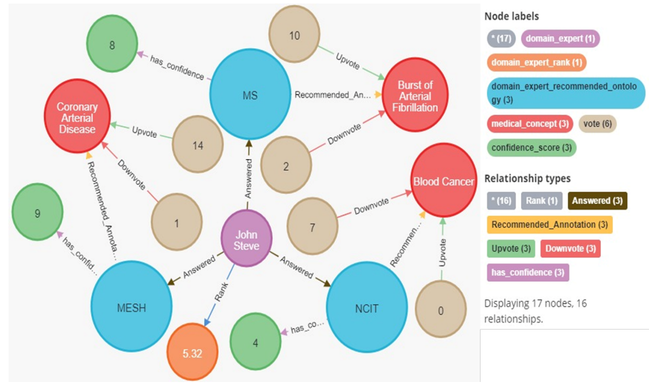

Figure 5: Domain Expert Semantic Annotation Recommendation Representation Through Knowledge Graph for Terminologies

Acknowledgments

This work is supported by the National Science Foundation grant ID: 2101350.

References

Abbas A, Mbouadeu S, Bisram A, Iqbal N, Keshtkar F, Bukhari SA. Proficient Annotation Recommendation in a Biomedical Content Authoring Environment. InKnowledge Graphs and Semantic Web: 4th Iberoamerican Conference and third Indo-American Conference, KGSWC 2022, Madrid, Spain, November 21–23, 2022, Proceedings 2022 Nov 13 (pp. 149-162). Cham: Springer International Publishing.

S. A. C. Bukhari, “Semantic Enrichment and Similarity Approximation for Biomedical Sequence Images,” University of New Brunswick (Canada), 2017.

Warren P, Davies J, Brown D. The Semantic Web-from vision to reality. ICT futures: Delivering pervasive, real-time and secure services. 2008 Apr 30:55-66.

Khalili A, Auer S. User interfaces for semantic authoring of textual content: A systematic literature review. Journal of web semantics. 2013 Oct 1;22:1-8.

K. Hasida, “Semantic Authoring and Semantic Computing,” in New Frontiers in Artificial Intelligence, Springer, Berlin, Heidelberg, 2003, pp. 137–149.

Jovanovic, Jelena, et al. ”Automated semantic tagging of textual content.” IT Professional 16.6 (2014): 38-46.

Tseytlin, Eugene, et al. ”NOBLE–Flexible concept recognition for large-scale biomedical natural language processing.” BMC bioinformatics 17.1 (2016): 1-15.

J. Jovanović and E. Bagheri, “Semantic annotation in biomedicine: the current landscape,” J. Biomed. Semantics, vol. 8, no. 1, p. 44, Sep. 2017.

Campos, David, Sérgio Matos, and José Luís Oliveira. ”A modular framework for biomedical concept recognition.” BMC bioinformatics 14.1 (2013): 1-21.

N. H. Shah, N. Bhatia, C. Jonquet, D. Rubin, A. P. Chiang, and M. A. Musen, “Comparison of concept recognizers for building the Open Biomedical Annotator,” BMC Bioinformatics, vol. 10, no. S9. 2009, doi: 10.1186/1471-2105-10-s9-s14.

Jonquet, Clement, et al. ”NCBO annotator: semantic annotation of biomedical data.” International semantic web conference, poster and demo session. Vol. 110. Washington DC, USA, 2009.

Cuzzola, John, Jelena Jovanović, and Ebrahim Bagheri. ”RysannMD: a biomedical semantic annotator balancing speed and accuracy.” Journal of Biomedical Informatics 71 (2017): 91-109.

Faisal, Muhammad Shahzad, et al. ”Expert ranking techniques for online rated forums.” Computers in Human Behavior 100 (2019): 168-176.

Mbouadeu, Steve Fonin, et al. ”Towards structured biomedical content authoring and publishing.” 2022 IEEE 16th International Conference on Semantic Computing (ICSC). IEEE, 2022.

Abbas, Asim, et al. ”Meaningful information extraction from unstructured clinical documents.” Asia-Pacific Advanced Network–Research Workshop. Vol. 48. 2019.

Landis, J. Richard, and Gary G. Koch. ”The measurement of observer agreement for categorical data.” biometrics (1977): 159-174.