Emerging Knowledge Extraction and Visualization in Medical Document Corpora

Christian Nawroth[1], Marc Herrmann[1], Felix Engel[2], Paul Mc Kevitt[2], and Matthias Hemmje[2]

[1] Chair of Multimedia and Internet Applications, Faculty of Mathematics and Computer Science, University of Hagen
christian.nawroth@fernuni-hagen.de marc.herrmann@studium.fernuni-hagen.de
[2] Research Institute for Communication and Cooperation, Dortmund
{fengel,pmckevitt,mhemmje}@ftk.de

Abstract. In this paper, we demonstrate our concept of emerging Named Entities (eNEs) for the tasks of Emerging Knowledge Extraction and Visualization in medical document corpora. We derive four use cases that utilize eNEs in medical document corpora to support medical expert users accessing emerging knowledge. We design the visual Emerging Named Entity Recognition and Information Retrieval System (visual eNER-IRS), supporting three of these use cases. We demonstrate proof-of-concept emerging knowledge visualizations for the different use cases. Finally, we present a detailed user evaluation of our visualization approach. The evaluation concludes that our approach helps users utilize eNEs on a corpus and a single document level. Overall, this paper demonstrates the benefits of our approach for the related project RecomRatio by providing recent and emerging knowledge for evidence-based medical use cases. Hence, the main contribution is a visualization of new medical concepts and emerging knowledge in literature for medical experts for supporting medical information, decision-making, and reporting in medical research and treatment.

Keywords: Emerging Named Entity Recognition (eNER) · Emerging Knowledge Visualiazation, emerging Named Entities (eNEs), Emerging Named Entity Recognition and Information Retrieval System (visual eNER-IRS)

1 Introduction and Motivation

This article is based on the results of a master thesis [6] and a dissertation project at the Chair of Multimedia and Internet Applications at the University of Hagen. Both works are related to the project *Recommendation Rationalisation* (RecomRatio [2]). RecomRatio is a DFG funded research project that aims to support expert health professionals during informed decision-making processes (e.g., for or against a certain diagnosis/therapy) by providing evidence through textual arguments found in the medical literature and documents. Within RecomRatio, we intend to make emerging Named Entities (eNEs) [12] and emerging

Argument Entities available for Information Retrieval (IR) in medical document corpora supporting medical argumentation engineering. Following our previous work on this topic [9][12], medical eNEs are names for medical entities (e.g., for diseases, drugs) that are in use in a medical document corpus (e.g., PubMed / MEDLINE). Yet, they are not formally acknowledged through the expert community, i.e., by adding them to a medical vocabulary. In addition, to support the underlying medical argumentation use cases within RecomRatio we defined an emerging Argument Entity (eAE) as an argument that contains an eNE in its premise or the conclusion element [14]. We argue that eNEs usually represent the most recent knowledge in a domain. Hence, emerging Named Entity Recognition (eNER) aims to recognize them in the document corpus and make them available for medical IR use cases. To recognize eNEs, we propose a hybrid approach combining textual Natural Language Processing (NLP) with Machine Learning (ML) techniques on temporal features [12]. Whilst our previous publications focused on the recognition of eNEs, here we explain why and how we provide eNEs to the user through visual interfaces that support four use cases for medical expert users. For our tasks presented here, we use two corpora: PubMed MEDLINE Baseline 2020[3] (MEDLINE) and PubMed Open Access (PMC OA) Subset[4]. Whilst the former generally only consists of the title and abstract, the latter also contains the full texts, so we decided to use both for our Document Engineering project. Between 1970 and 2019, the number of citations added to MEDLINE grew from 219.337 entries per year to 1.406.789, based on our corpus index statistic derived from our experimental corpora. So the yearly growth rate increased by a factor 6.4 within 50 years. Furthermore, we outline how our work links knowledge from these two corpora to the ClinicalTrials[5] document corpus.

2 state-of-the-art and Related Work

Our work is related to the task of realtime Emerging Topic Detection in Microblogs as presented by Chen et al. [4], which also utilizes ML techniques on non-textual features to detect emerging topics within microblogs. Our approach differs as it does not focus on realtime detection, but long-term eNEs in a scientific text corpus and therefore, it uses different non-textual temporal features compared to Chen et al. [4]. Furthermore, our approach for eNER aims to recognize eNE in scientific corpora and hence combine Document Engineering techniques from traditional NER and ML. In recent work, Wang et al. [17] apply hot topic detection to the field of academic big data, which they call Academic Hot Topic Detection. Like our approach, they combine a textual NER approach in the first stage with a feature learning approach. Their main features are a co-occurrence graph and word embeddings amongst additional document related features. In contrast, we focus on eNEs in a solely temporal way, not yet analyzing whether these topics are "hot", i.e. setting a trend of popular information

[3] https://www.nlm.nih.gov/databases/download/pubmed_medline.html
[4] https://www.ncbi.nlm.nih.gov/pmc/tools/openftlist/
[5] https://clinicaltrials.gov/

need/interest. The design of the visualization subsystem generally follows the IVIS4BigData Framework introduced by Bornschlegl et al. [3]. The framework describes a method to transform raw data from big data sources into"advanced visual user interfaces for Big Data Analysis" to allow"efficient and effective" Human-Computer Interaction (HCI). A major component of the IVIS4BigData is a pipeline that consists of different steps to provide data insight and effectuation based on raw data. These steps are Raw Data Collection, Data Structures, Visual Structures, and Views. Between Data Collection and Data Structures, the framework includes an analytics layer that comprises an analytical component of the underlying big data use cases.

3 Visual eNER-IRS System Design

To design the visual Emerging Named Entity Recognition and Information Retrieval System (visual eNER-IRS), we apply a user-centered design approach [13]. Therefore, we introduce the four use cases of the visual eNER-IRS, give a brief overview of the general system architecture, and derive the architecture of the visualization subsystem that will become the basis of the later Argument Visualization System. Here, we focus on the visualization subsystem and explain the underlying architecture briefly to ensure general understanding. A detailed description of the eNER pipeline is given in [12, 11].

3.1 Use Cases of the Visual eNER-IRS

The visual eNER-IRS is intended to support four different information retrieval use cases [11]. These are eNE retrieval support, document linking through NEs, emerging Knowledge Discovery, and (later) emerging Argument Entity discovery, as shown in Fig. 1. These four use cases are supported by one or two generic visual use cases provided through the visualization subsystem. In the following, we briefly introduce the four general use cases summarizing [11]. The first, eNE Retrieval Support (see Fig. 1) aims at providing functionality that utilizes eNEs to enhance and support several standard retrieval methods, like query completion, filtering, faceted search, and boosting of ranking results depending on eNEs. The associated visual use case is visual eNE Retrieval Support. This visualization use case is intended to highlight eNEs during several steps of user interaction with the retrieval system. The second general use case supported by one visualization use case is document linking through eNEs. In this use case, eNEs are utilized to provide a link between documents from different corpora. For example, a user finds a clinical trial in the ClinicalTrials (CT) Corpus that contains eNEs that represent new medical knowledge in the respective clinical trial. Then these eNEs can be used to search for documents in another text corpus, e.g., MEDLINE, to retrieve new and emerging knowledge from that text corpus too. The associated visual use case *Visual Linking through eNEs* is intended to provide an interactive graphical representation of that use case, i.e., a network graph showing links between documents from different corpora based

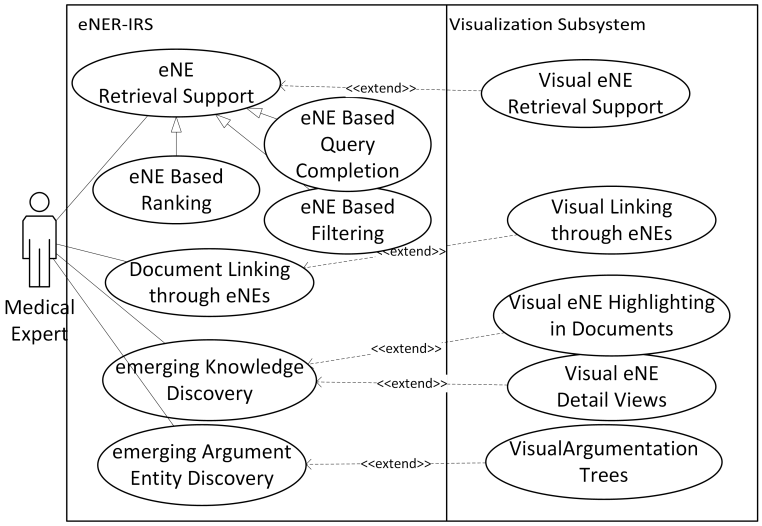

Fig. 1. UML eNER-IRS Use Case Model [11]

on eNEs. The third general use case is visual emerging Knowledge Discovery. This use case has an exploratory characteristic and enables users to explore new knowledge on document level and the single emerging entity level. The associated visual use cases (visual eNE highlighting in documents, visual eNE detailed views) provide views for both exploratory aspects, which means visual highlighting of eNEs in selected papers and providing detailed information on selected eNEs based on the textual and temporal analysis of the eNER-IRS. The fourth use case is emerging Argument Entity Discovery. Based on emerging Argument Entities (e.g., from a survey article) in arguments' premises or conclusions, the expert medical users can retrieve, link, and visualize arguments that cover the most recent medical knowledge. This use case is not covered by this paper but published in [10].

3.2 General Conceptual Architecture

Following the motivation and the three initial use cases, our architectural modelling approach (see Fig. 2) for recognizing eNEs in a medical document and query corpus combines methods from NLP, NER, IR, and ML [12, 11]. Our approach follows the Model View Controller (MVC) paradigm [7].

Here, we focus on the conceptual design of the view layer that contains the visualization components of the eNER-IRS. As the View layer interacts with the controller layer, we introduce the controller layer for a general understanding. A more detailed description and evaluation of the underlying eNER pipeline in the controller layer is published in [12, 11]. The core components of the controller layer that are referenced in the View layer are the medical document corpus, the baseline NLP and NER, the temporal features search engine, and

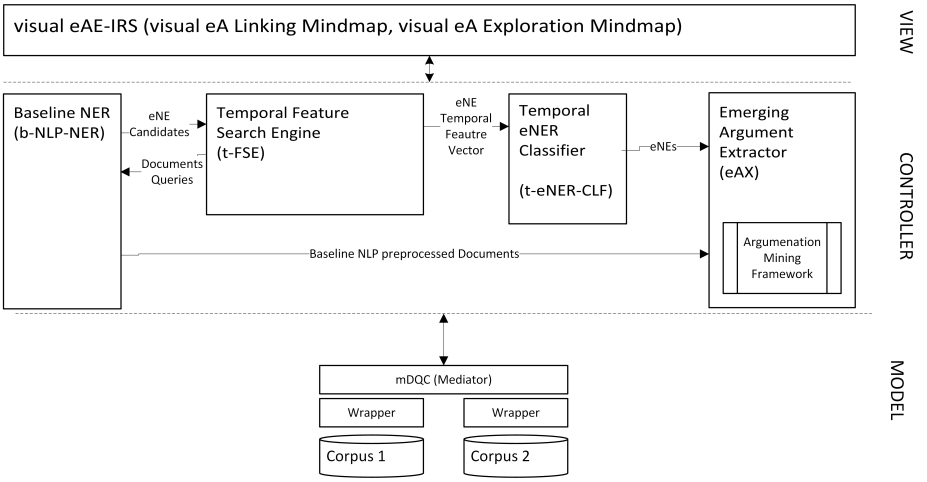

Fig. 2. Conceptual System Architecture adapted from [11]

the temporal eNER Classifier. The medical document corpus is the indexed document corpus containing all medical documents within the system. The baseline NLP and NER provides the extraction of eNER candidates based on textual features. The temporal features search engine extracts temporal features of the eNE-candidates from the medical document corpus (e.g., year of first use). The temporal eNER Classifier is the ML component that finally classifies eNE candidates based on these temporal features. Based on the extracted eNEs and the baseline preprocessed documents, the emerging Argumentation Extractor aims at identifying eAEs in arguments. It will rely on a state-of-the-art argumentation mining framework, e.g., ArgumenText [16]. Following this overview of the core eNER components in the controller layer in the next subsection, we introduce the emerging knowledge visualization subsystem's conceptual design.

3.3 Conceptual Design of the Visual Subsystem

The conceptual design of the visualization subsystem has to consider challenges posed by the big data characteristics of the underlying document corpora and vocabularies. Hence, to address these challenges for the visualization subsystem's conceptual design, we applied the IVIS4BigData Framework. In general, IVIS4BigData aims to make big data resources available and beneficial for users through visualization. Fig. 3 shows how we use the IVIS4BigData Framework to transform raw textual data from medical corpora into views that allow medical expert users' visual data insight and emerging knowledge effectuation. Compared to the full IVIS4BigData Framework in our work, the pipeline part and a feedback channel (user empowerment) are implemented. Furthermore, we focus on the end-user's view, but we do not implement the views on the first three components of the full IVIS4BigData pipeline. Textual Big Data Sources for our

system are the two corpora PubMed and MEDLINE, as introduced above. The raw textual data is collected in the medical document corpus from the system architecture design (see Fig. 2). In the IVIS4BigData, the medical document corpus refers to the "Data Collection" component. The analytics layer of IVIS4BigData in our system design is represented by the two components that perform analytic tasks (eNER) on the raw data and turn the raw data into data structures: The baseline NLP and NER and the temporal eNER classifier in the controller layer as described above. The following IVIS4BigData component "Data Structures" in our work is represented by a JSON structure that encodes the mapping between a corpus document and the automatically recognized eNEs. These mappings are then transformed into a visual structure that is stored in a search engine index whose Entity-Relationship Model (ERM) is shown in Fig. 4. The *Article* entity has several attributes derived directly from the underlying corpus metadata, such as Author, Abstract, or its corpus ID (e.g., for PUBMED / MEDLINE: PMID). In contrast, *Entity* itself and its attributes are not taken from metadata but extracted through the eNER-IRS. For example, those attributes contain the Date Created (time of first use in the corpus), the name, and a possible category of it. Hence, more generally speaking, the mapping introduced above maps already existing corpus knowledge to new knowledge extracted by the eNER-IRS. The final component of the IVIS4BigData Framework is the eNER-IRS-GUI

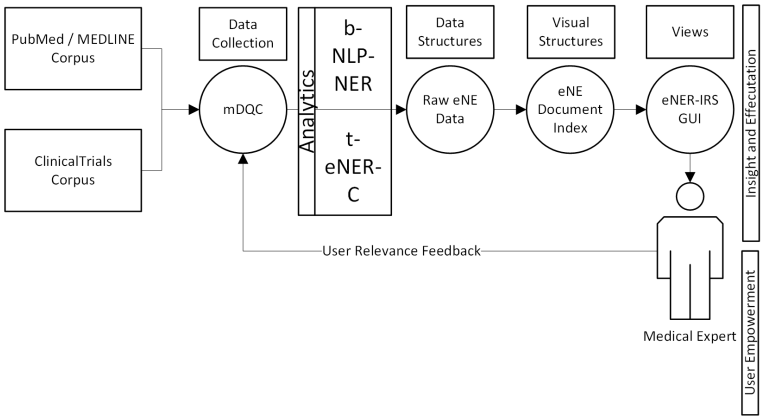

Fig. 3. Visual eNER based on IVIS4BigData [3].

that we discuss in the remainder of this paper. We present a prototypical proof-of-concept implementation following the three initial use cases (eNE retrieval support, document linking through NEs, emerging Knowledge Discovery) in the following section.

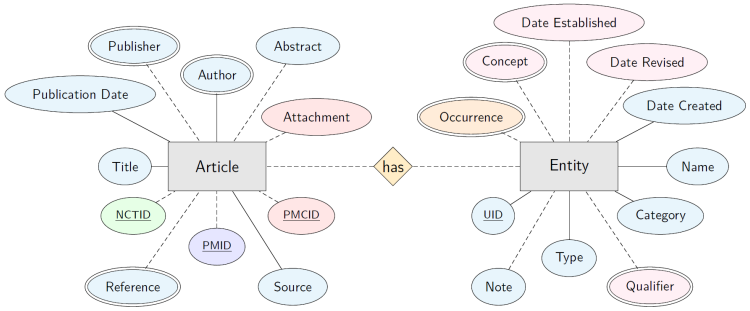

Fig. 4. Entity Relationship Model of Visual Structure [6]

4 Proof-of-Concept Implementation

This section describes the prototypical proof-of-concept implementation of the visual eNER-IRS use cases based on the IVIS4BigData framework using several methods and packages. First, we give a brief overview of the technical implementation followed by descriptions of each of the visual eNER-IRS use cases followed by a description of the graphical user interface (GUI) aspects.

4.1 Technical Implementation

In the following, we describe the prototypical technical implementation of the previously introduced visual eNER-IRS use cases based on eNEs. The underlying concept of the Visualization Subsystem (see Fig. 1) can be considered as two independent software systems based on the IVIS4BigData Framework as follows: To transform data structures into visual structures (see Fig 4) as a first step, we developed a batch application based on the Spring Batch Java-Framework. It converts all the eNER-IRS output JSON files, including data about eNEs such as its name and occurrences in medical documents, into processed and consolidated XML files. These files are then indexed by Apache Solr to create the eNEs visual data structures (IVIS4BigData: Visual Mappings). These files are also used in another batch processing step that reads the raw data of the different medical corpus (PMC, ClinicalTrials), extracts relevant data attributes, and enriches the data by adding information about entities such as MeSH concpets and eNEs. Also, these XML output files are indexed by Apache Solr, and as a result, two visual data structures for medical documents and eNEs are created. The second system is a client-server architecture software based on the Spring Web Model-View-Controller (MVC) Java-Framework. It generally reads the visual structure data from Apache Solr and displays it on different web pages to provide the described visual eNER-IRS uses cases (IVIS4BigData: View Transformations). The web page design and functionality is based on frameworks such as Boot-

strap[6] and jQuery[7]. Additionally, for the specific use case of visual linking and highlighting of eNEs, we rely on the JS-Library D3.js[8] for diagram and PDF.js[9] for document visualization. To transport data from server to client, multiple REST API endpoints are available and tailored for the specific visual eNER-IRS use case. In the following, the prototypical implementations of the use cases are explained.

4.2 Visual eNE Retrieval Support

The document search enables users to browse through the different document collections (PubMed, ClinicalTrials) easily with two filter categories (*eNEs* and *Medical Subject Headings*): The main search field and the result of the document search are displayed in the right view area. Above, next to the search field, there is an additional button to control on which document collection the search should be performed. Below, the matching documents, including title, publication date, unique identifiers, entity categories, and source document collection, are listed separately. The detailed view of a paper shows the authors, the assigned entities. If available, the abstract in which the assigned entities are highlighted, is shown. The left sidebar contains all necessary controls to conveniently browse the dataset of *Emerging Named Entities* and *Medical Subject Headings*. As a result, the documents are filtered based on the selected entities. Furthermore, the sorting of the search result can be influenced by the green Learning To Rank (LTR) button next to the main search field. Generally speaking, the documents are sorted by an individual score in descending order. This score is a measurement for how relevant each document is for the given user query. The gray shaded numerical value reflects the default sorting, whereas the green shaded value also considers the information about eNEs. In detail, this score increases based on the number of assigned eNEs and on whether the user query also matches these entities. To improve user search experience, the main search field is extended by an auto-completion functionality (see Fig. 5) that lists query suggestions from multiple datasets based on the current user input. Depending on the selected document collections, the first suggestions are made by matching document titles and eNEs. Also, further suggestions derive from the entity categories *eNEs* and *Medical Subject Headings* (see Fig. 6). Selecting an entity suggestion results in automatically adding the entity as a filter criterion in the left sidebar.

4.3 Visual Linking through eNEs

The relationships between entities are created once multiple entities appear in a single document. The greater the number of documents in which two related entities appear, the closer their relationship is. Such relationships can be researched interactively with the help of the network graph accessible under the

[6] https://getbootstrap.com/
[7] https://jquery.com/
[8] https://d3js.org/
[9] https://mozilla.github.io/pdf.js/getting_started/

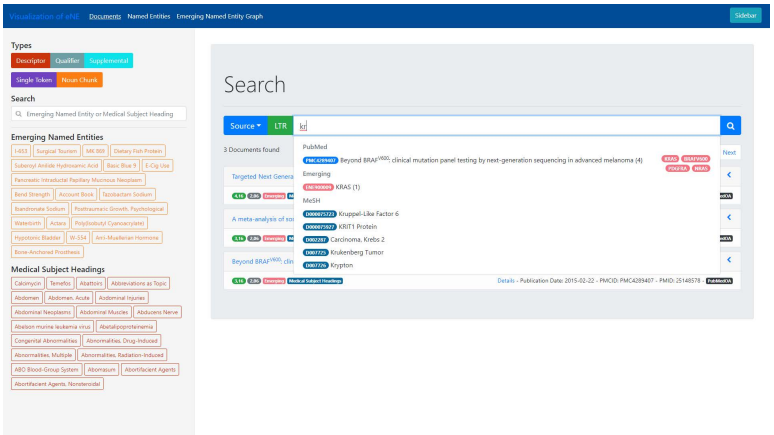

Fig. 5. Visual eNER-IRS: Autocompletion for User Query [6].

navigation item *Emerging Named Entity Graph* (see Fig. 7). In the left sidebar, all entities with at least one relationship are displayed. The list can be filtered regarding type, reference, and name. Additionally, the list automatically updates once an entity is selected by only showing entities in a direct relationship. The network graph itself is shown in the right view area and refreshes automatically once the entities' selection is updated. A node represents each chosen eNE, and its size depends on the total number of documents from the different collections it is assigned to. The links between nodes display the connecting documents, and their amount is represented by edge width. By clicking on the link details of the relationship are revealed.

4.4 Visual eNE Highlighting in Documents

The detailed view of a document contains all mandatory attributes such as title, unique identifiers, document collection and also optional attributes such as authors, assigned entities (eNEs and non-eNEs (MeSH)) and the abstract or original PDF-document. The availability of the optional attributes depends on the collection source of the document. For example, only for documents from the PubMed Central Open Access (PMCOA) collection, the original document in PDF-format can be displayed. In the left sidebar, the assigned eNEs and non-eNEs are listed as interactive buttons. Also, these entities are highlighted in the continuous text of the abstract or the PDF-document, if available (see Fig. 8). In case of viewing a PDF-document, additional buttons to page backwards and forwards and download the document are displayed. For eNEs, the interactive button can be expanded to a drop-down list showing all the document pages on which the term appears.

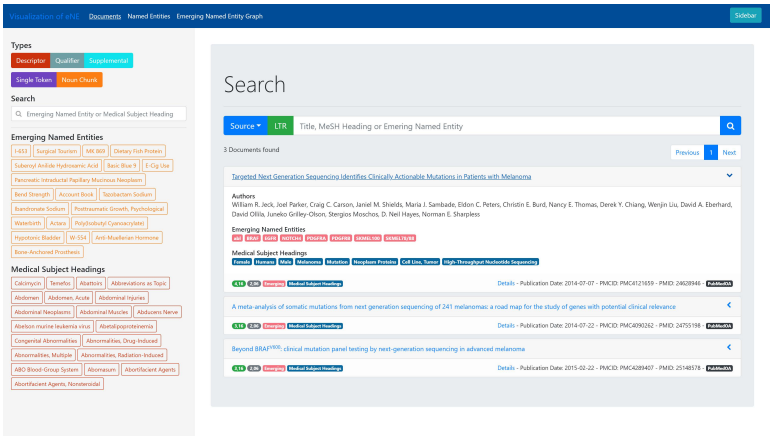

Fig. 6. Visual eNER-IRS: Search for Medical Documents [6].

4.5 Visual eNE Detailed Views

This visual eNER-IRS use case provides the highest interaction between expert medical users and the eNER-IRS processes. Hence, the user can acknowledge or reject an eNE candidate suggested by the eNER-IRS. The requested feedback (and the respective interface) is intentionally binary (ACKNOWLEDGE / REJECT), keeping in mind that expert medical users may lack data science knowledge to give a more differentiated assessment. However, for those expert medical users with data science / ML skills, the visualization provides two metadata parameters from the eNER-IRS for their decision process. The data types of the result set of the ML process for this visual eNER-IRS use case are terms/tokens representing eNEs and temporal (statistical) metadata derived from the big data analysis of the temporal feature search engine. Besides the temporal metadata, the classification threshold from the underlying eNER-IRS component is displayed (see Fig. 9).

The detailed view of an entity contains all mandatory attributes. These are unique identifiers, category, type, name, and also optional attributes such as references to documents from different collections, overall frequency, and multiple dates related to the creation, revision, and establishment (See Fig. 10). The buttons in the left sidebar grouped by document collection reflect the related entities. An additional bar chart showing the frequency of occurrence on an annual basis is displayed depending on data availability. Two drop-down lists can change the plotted range of years on the right side above the diagram. Additionally, related entities' occurrence data can be added interactively by the buttons shown on the right-hand side.

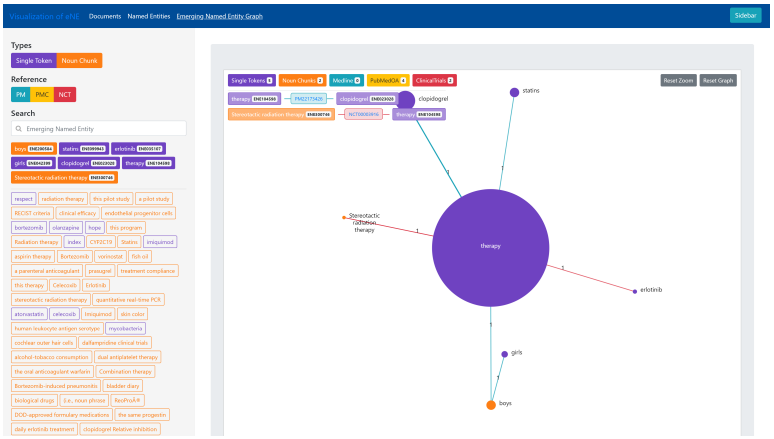

Fig. 7. Visual Linking through eNEs [6].

5 Evaluation

5.1 Evaluation Methodology

The evaluation primarily follows a task-oriented evaluation approach which additionally includes the UMUX methodology to assess the perceived usability [5, 8]. It is based on a 20-page questionnaire which is included as auxiliary material. Nine participants contributed to the evaluation. They belong to the user stereotypes *Medical User*, *Information Retrieval Expert*, *Science and Engineering Expert* and *other*.

5.2 Test Questions

Based on the first preparatory study results, we designed three medical test scenarios described in the questionnaire. The test scenarios aim to figure out to which extent users can use the eNER-IRS visualization to fulfill particular use case scenarios. For each scenario, the questionnaire provides a detailed task description:

1. *Medical Document Search* In this scenario, an exemplary search for medical documents is conducted. The search includes, on the one hand, the filtering of search results and on the other hand the visual highlight of Emerging Named Entities. In particular, the highlighting of eNEs enables the user to perform a context-sensitive and professional evaluation of the terms.
2. *Details of eNEs* This scenario covers the detailed consideration of all the available information about Emerging Named Entities.
3. *Relationships between eNEs* This scenario shows how the use and configure the network graph to identify and investigate the relationships between eNEs.

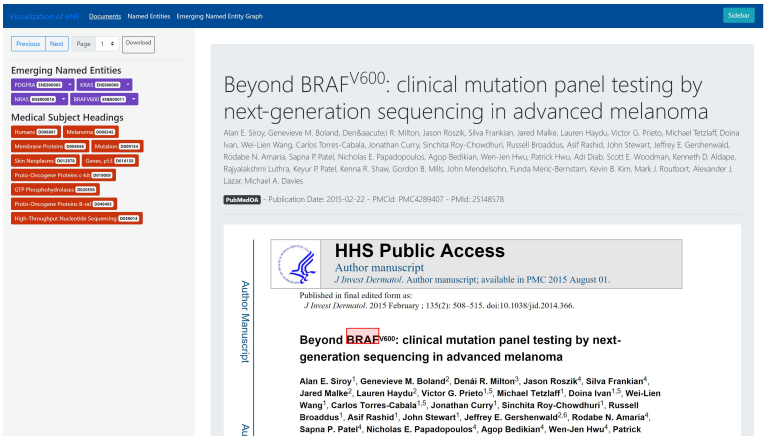

Fig. 8. Visual eNE Highlighting In Documents [6] (Article: [15]).

For each of the scenarios, we defined 2-5 multiple choice test questions (TQs). For each TQ, one or more answer options are correct. Users' answers with all correct answer options are classified as *correct*, with some correct answer options as *partly_correct* and with no correct answer options as *wrong*. Fig. 11 shows the evaluation for the test questions (overall results and results per question). turns out that there is a share of $\geq \frac{2}{3}$ of correct or partially correct answers for all three scenarios. This finding concludes that, in general, users were able to fulfill the three test scenarios defined above. However, within the questions of the particular scenarios, a variance regarding the outcome can be observed. In scenario (1), the TQ04, and scenario (2), the TQ06 has an outcome of correct answers of < 0.5. TQ04 deals with the highlighting of single eNEs in a document, TQ06 is about ranking statistics. This finding concludes that only the visualization details must be improved while the overall system is usable and performs well.users were able

5.3 Usability Metrics

The goal of this step is to assess the perceived usability of the user based on ISO 9241-11. Therefore, the Usability Metric for User Experience (UMUX) with its four-item Likert scale is used [5]. with the following questions:

- Q01 The system's capabilities meet my requirements.
- Q02 Using [this system] is a frustrating experience.
- Q03 [This system] is easy to use.
- Q04 I have to spend too much time correcting things with [this system].

First, Fig. 12 shows the results of Q01 and Q03.

It turns out that a more than 0.5 of the participants gave a positive or neutral assessment for both questions. However, for Q03, turns out that a significant

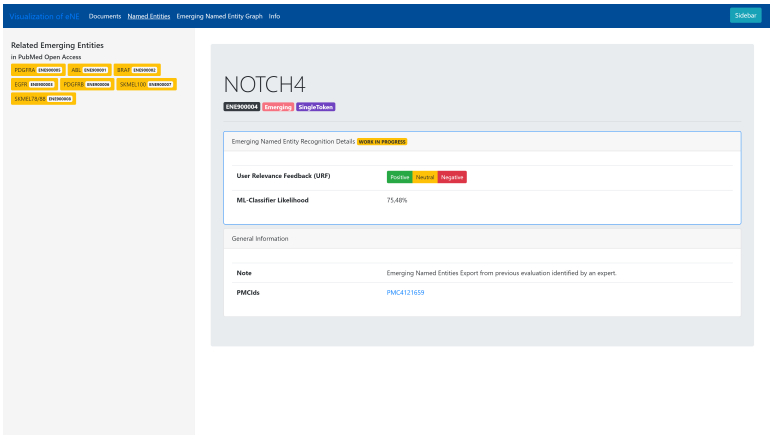

Fig. 9. Acknowledgement of eNEs [6].

portion of participants do not think that the system is easy to use. We argue that this is not surprising but emphasize that the system is an expert system that may require more extensive training to be used beneficially. Secondly, Fig. 13 shows the results of Q02 and Q04, now with a reversed colour scale compared to Fig. 12 as explained before.

Again, turns out that more than half of the participants have a positive or neutral assessment regarding Q02 and Q04.

Thirdly, we plot the mean results and the standard deviations for all questions (See Fig. 14). The plots of the mean and standard deviation show reflect the results introduced earlier. For Questions Q01 and Q03, the mean is greater or equal to the *neutral* assessment, while for Q02 and Q04, it is below. We argue that the relatively strong standard deviations result from the heterogeneous participant group evaluating our system, with different experiences in the medical domain, and using expert retrieval systems.

Overall the usability evaluation showed a positive outcome, leading to the conclusion that the system, in general, has reliable usability whilst there are again improvements in visualization details. Furthermore, it shows the need for sufficient training on the system for users inexperienced in using expert systems or who are new to the medical domain.

5.4 Added Value in Professional Terms

The following questions are intended to assess the added value in professional terms related to certain areas in the prototypical application. In contrast to UMUX, here the 5-point Likert scale is used to express how much the tester agrees (5) or disagrees (1) with a particular statement [1], (Neutral: (3), N/A: (0)):

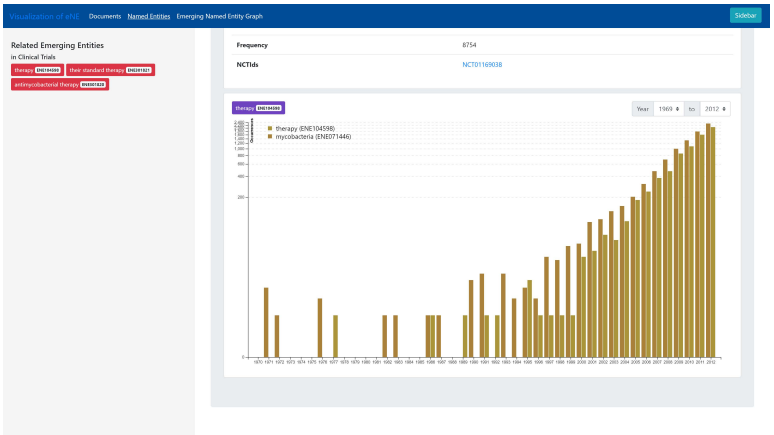

Fig. 10. Visual eNE detailed views [6]

- Q05 The possibility to filter medical documents with the help of eNEs can generate a true added value during the research process.
- Q06 The visually highlighted eNEs in continuous text and original medical documents support the assessment in terms of their quality and professional relevance.
- Q07 The possibility to download the original medical document with visually highlighted eNEs allows distributing the information with expert colleagues simply.
- Q08 The interactive bar chart is a valuable visualization to display the occurrences of eNEs.
- Q09 The interactive network graph is a valuable visualization to display the relationships between eNEs.
- Q10 The visualized relationships between eNEs support the assessment in terms of their quality and professional relevance.

Fig. 15 envsisions the results of the questions above.

Here, turns out that again all questions have a *neutral* or better outcome in more than half of the answers given. Except for Q08, more than 50% have a *better than neutral* outcome. The high ratings for Q05, Q06, and Q09 are promising. These questions reflect the core of our work and our use case. They show that our concept of eNE and its utilization within information retrieval use cases and their visualization are seen as beneficial by most participants. In contrast, the questions Q07 and Q08 focus on detail visual implementations. They emphasize that there is room for improvement regarding the aspects of the visualization. Again, we plotted the mean results per question, including the standard deviation (see Fig. 16).

For all questions, it shows mean values significantly above the neutral value of (3). Question seven has the strongest standard deviation, while the other standard deviations are more moderate. That again reflects a strong variance among

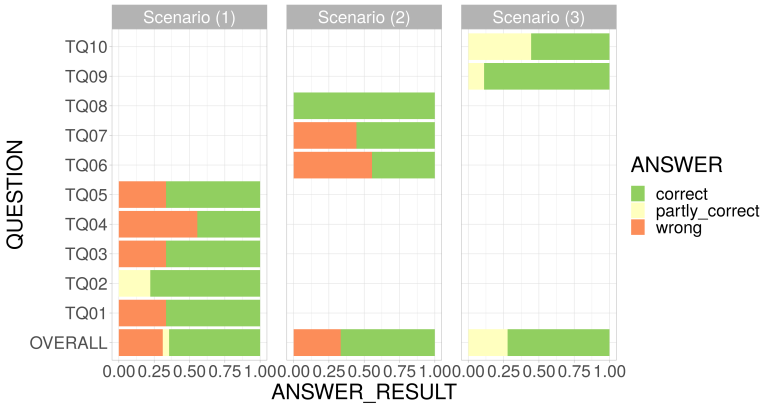

Fig. 11. Test Questions' Results

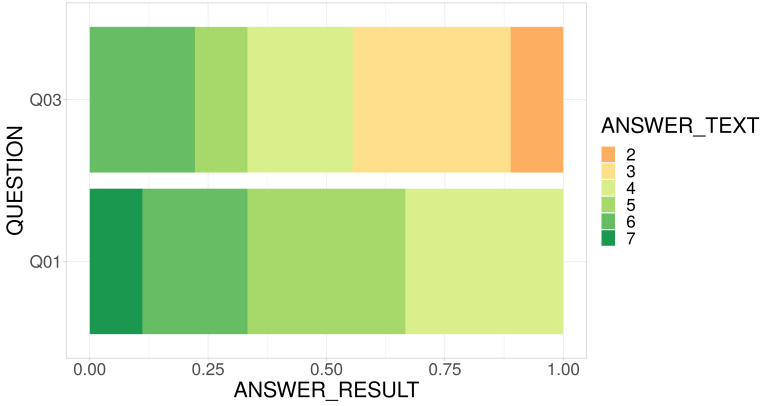

Fig. 12. Usability Assessment I

the participants when it comes to using a detailed implementation feature. The high means for questions Q05, Q06, and Q09 demonstrate that our concept of eNEs and their visualization is useful and beneficial amongst the participants.

6 Conclusion and Discussion

In this paper, we outlined a complete workflow utilizing and visualizing our concept of emerging knowledge represented by eNEs. We showed that eNEs are in use in medical document corpora, and represent the most recent knowledge. We introduced four visual use cases to utilize eNEs by medical experts in document corpora. We derived a system design for recognizing and visualizing them to support medical retrieval and argumentation use cases. We designed a visualization subsystem on the IVIS4BigData Framework, visualizing eNEs in documents and

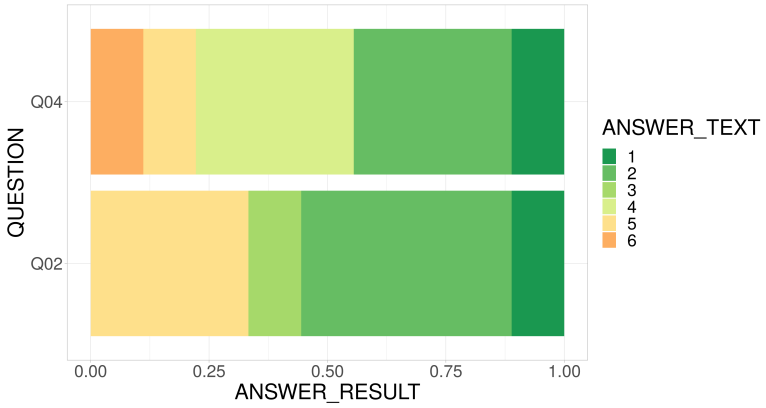

Fig. 13. Usability Assessment II

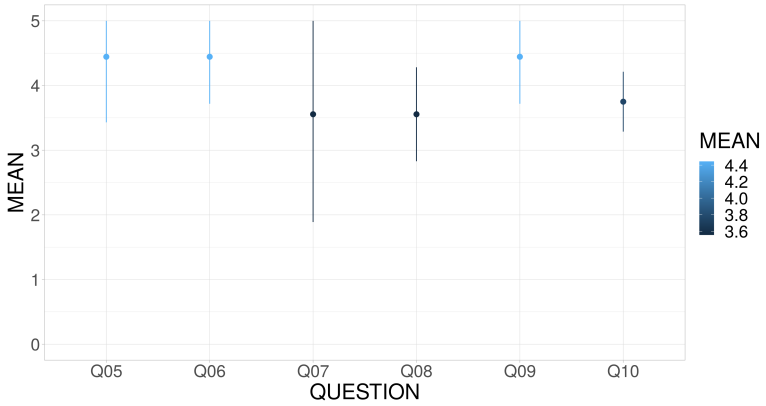

Fig. 14. Overall Usability Assessment (Means)

document corpora and demonstrated a proof-of-concept implementation. Our evaluation indicated that our three use cases and their four visualization could help expert medical users searching for document-based medical evidence. Furthermore, we identified items for improving our concept in visual details. We showed the necessity for sufficient user training with the complex task of eNER in a medical document corpus. Concerning future work, we outlined the emerging Argumentation Extraction system.

References

1. 5-Point Likert Scale. In: Preedy, V.R., Watson, R.R. (eds.) Handbook of Disease Burdens and Quality of Life Measures, pp. 4288–4288. Springer, New York, NY (2010). https://doi.org/10.1007/978-0-387-78665-0_6363, https://doi.org/10.1007/978-0-387-78665-0_6363

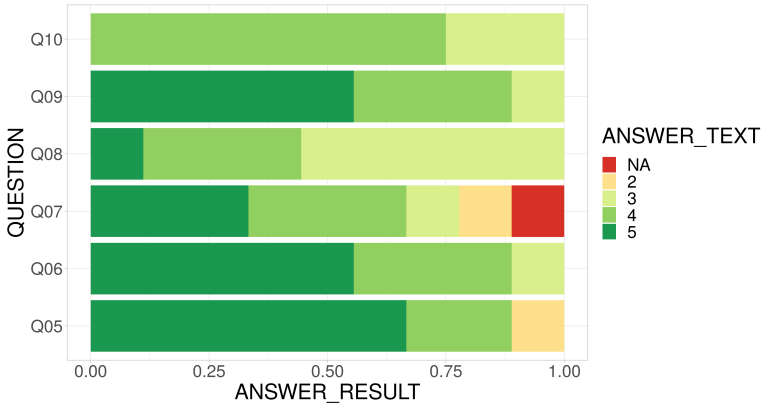

Fig. 15. Added Value in Professional Terms

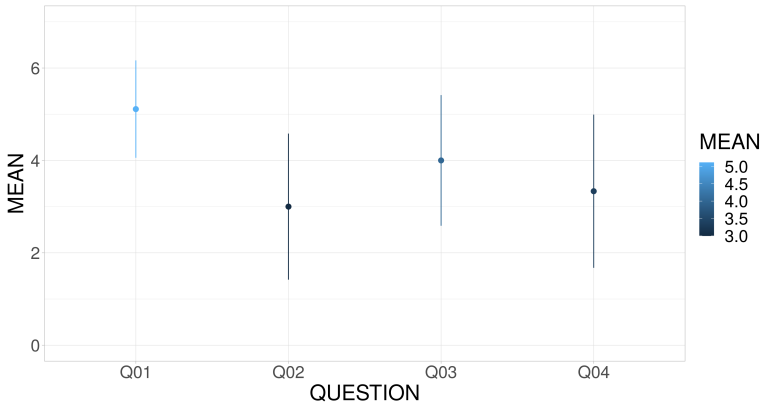

Fig. 16. Added Value in Professional Terms (Means)

2. of Bielefeld, U.: Rationalizing Recommendations (RecomRatio): Project-Homepage. Bielefeld (2017), http://www.spp-ratio.de/de/projekte/ratiorec/
3. Bornschlegl, M.X., Berwind, K., Kaufmann, M., Engel, F.C., Walsh, P., Hemmje, M.L., Riestra, R.: IVIS4BigData: A Reference Model for Advanced Visual Interfaces Supporting Big Data Analysis in Virtual Research Environments. In: Bornschlegl, M.X., Engel, F.C., Bond, R., Hemmje, M.L. (eds.) Advanced Visual Interfaces. Supporting Big Data Applications. pp. 1–18. Lecture Notes in Computer Science, Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-50070-6_1
4. Chen, Y., Amiri, H., Li, Z., Chua, T.S.: Emerging topic detection for organizations from microblogs. In: SIGIR 2013 - Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 43–52 (2013). https://doi.org/10.1145/2484028.2484057
5. Finstad, K.: The usability metric for user experience. Interacting with Computers **22**(5), 323–327 (2010), publisher: Oxford University Press Oxford, UK

6. Herrmann, M.: Visualisierung von Emerging Named Entities im Rahmen von Information Retrieval in medizinischen virtuellen Forschungsumgebungen. Masterthesis, FernUniversität in Hagen, Hagen (2020)
7. Krasner, G.E., Pope, S.T., others: A description of the model-view-controller user interface paradigm in the smalltalk-80 system. Journal of object oriented programming **1**(3), 26–49 (1988)
8. Laubheimer, P.: Beyond the NPS: Measuring Perceived Usability with the SUS, NASA-TLX, and the Single Ease Question After Tasks and Usability Tests (2018), https://www.nngroup.com/articles/measuring-perceived-usability/
9. Nawroth, C., Engel, F., Eljasik-Swoboda, T., Hemmje, M.: Towards enabling emerging named entity recognition as a clinical information and argumentation support. In: DATA 2018 - Proceedings of the 7th International Conference on Data Science, Technology and Applications (2018)
10. Nawroth, C., Duttenhöfer, A., Hemmje, M.: Argumentationsunterstützung durch emergentes Wissen in der Medizin. In: to appear in: Wilhelm Bauer, Joachim Warschat, Innovation durch Natural Language Processing - Mit Künstlicher Intelligenz die Wettbewerbsfähigkeit verbessern (2020)
11. Nawroth, C., Engel, F., Hemmje, M.: Emerging Named Entity Recognition in a Medical Knowledge Management Ecosystem. In: Accepted for: Proceedings of 12th International Conference on Knowledge Engineering and Ontology Development. Budapest (2020)
12. Nawroth, C., Engel, F., Mc Kevitt, P., Hemmje, M.L.: Emerging Named Entity Recognition on Retrieval Features in an Affective Computing Corpus. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 2860–2868 (Nov 2019). https://doi.org/10.1109/BIBM47256.2019.8983247
13. Norman, D.A., Draper, S.W.: User centered system design: New perspectives on human-computer interaction. CRC Press (1986)
14. Palau, R.M., Moens, M.F.: Argumentation mining: the detection, classification and structure of arguments in text. In: Proceedings of the 12th international conference on artificial intelligence and law. pp. 98–107 (2009)
15. Siroy, A.E., Boland, G.M., Milton, D.R., Roszik, J., Frankian, S., Malke, J., Haydu, L., Prieto, V.G., Tetzlaff, M., Ivan, D., Wang, W.L., Torres-Cabala, C., Curry, J., Roy-Chowdhuri, S., Broaddus, R., Rashid, A., Stewart, J., Gershenwald, J.E., Amaria, R.N., Patel, S.P., Papadopoulos, N.E., Bedikian, A., Hwu, W.J., Hwu, P., Diab, A., Woodman, S.E., Aldape, K.D., Luthra, R., Patel, K.P., Shaw, K.R., Mills, G.B., Mendelsohn, J., Meric-Bernstam, F., Kim, K.B., Routbort, M.J., Lazar, A.J., Davies, M.A.: Beyond BRAFV600: Clinical Mutation Panel Testing by Next-Generation Sequencing in Advanced Melanoma. Journal of Investigative Dermatology **135**(2), 508–515 (Feb 2015). https://doi.org/10.1038/jid.2014.366, http://www.sciencedirect.com/science/article/pii/S0022202X15371189
16. Stab, C., Daxenberger, J., Stahlhut, C., Miller, T., Schiller, B., Tauchmann, C., Eger, S., Gurevych, I.: ArgumenText: Searching for Arguments in Heterogeneous Sources. In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations. pp. 21–25. Association for Computational Linguistics, New Orleans, Louisiana (Jun 2018). https://doi.org/10.18653/v1/N18-5005, https://www.aclweb.org/anthology/N18-5005
17. Wang, B., Yang, B., Shan, S., Chen, H.: Detecting Hot Topics From Academic Big Data. IEEE Access **7**, 185916–185927 (2019), publisher: IEEE