

Interactive and Iterative Annotation for Biomedical Entity Recognition

Seid Muhie Yimam¹, Chris Biemann^{1(✉)}, Ljiljana Majnarić²,
Šefket Šabanović², and Andreas Holzinger^{3,4}

¹ CS Department, FG Language Technology, TU Darmstadt,
64289 Darmstadt, Germany
{yimam,biem}@cs.tu-darmstadt.de

² Faculty of Medicine, Josip Juraj Strossmayer University of Osijek, Osijek, Croatia
ljiljana.majnarić@gmail.com

³ Research Unit HCI-KDD, Institute for Medical Informatics, Statistics and
Documentation, Medical University Graz, Auenbruggerplatz 2, 8036 Graz, Austria
a.holzinger@hci-kdd.org

⁴ CBmed - Center for Biomarker Research in Medicine, Stiftingtalstrasse 5,
8010 Graz, Austria

Abstract. In this paper, we demonstrate the impact of interactive machine learning for the development of a biomedical entity recognition dataset using a human-into-the-loop approach: during annotation, a machine learning model is built on previous annotations and used to propose labels for subsequent annotation. To demonstrate that such interactive and iterative annotation speeds up the development of quality dataset annotation, we conduct two experiments. In the first experiment, we carry out an iterative annotation experimental simulation and show that only a handful of medical abstracts need to be annotated to produce suggestions that increase annotation speed. In the second experiment, clinical doctors have conducted a case study in annotating medical terms documents relevant for their research. The experiments validate our method qualitatively and quantitatively, and give rise to a more personalized, responsive information extraction technology.

Keywords: Interactive annotation · Machine learning · Knowledge discovery · Data mining · Human in the loop · Biomedical entity recognition

1 Introduction and Motivation

The biomedical domain is increasingly turning into a data-intensive science, and one challenge with regard to the ever-increasing body of medical literature is not only to extract meaningful information from this data, but to gain knowledge, insight, and to make sense of the data [1]. Text is a very important type of data within the biomedical domain. For example, patient records contain large amounts of text which has been entered in a non-standardized format, consequently posing a lot of challenges to processing of such data and for the clinical

doctor the written text in the medical findings is still the basis for any decision making [2,3]. Further, scientific results are communicated in text form, consequently for the biomedical domain text is an indispensable data type for gaining knowledge [4].

Modern automated information extraction (IE) systems usually are based on machine-learning models, which require large amount of manually annotated data to specify the model according to the task at hand. Unfortunately, particularly in the medical domain, experts have obligations with higher priorities, thus it is very expensive and cumbersome to annotate a large number of training examples. In order to alleviate this problem, there is a need for an approach where human annotators are facilitated to annotate faster than the traditional way, in order to produce required annotations in less time.

In this paper, we tackle the extractions of entity mentions from biomedical texts, specifically from MEDLINE abstracts (www.ncbi.nlm.nih.gov/pubmed), using a recent human-into-the-loop automation strategy that has not been applied in the medical domain before. Unlike named entity recognition (NER) systems on e.g. the news domain, entity recognition on medical domains comprises of extractions of technical terms in the broader medical and biological arena such as name of diseases, proteins, substances and so on, see e.g. [5,6].

Such an automation approach is specifically very important for the medical domain, as a full manual annotation is extremely expensive. Medical professionals in turn, however, are willing to perform this task only diligently if it matches their current field of interest. The human-into-the-loop automation approach enables users to start the automation process without pre-existing annotations, and works by suggesting annotations as soon as the users have annotated a rather small number of documents. This annotate-little and predict-little strategy is deemed adequate for Biomedical domains as it 1) produce quality annotation in a very short period of time, and 2) the approach is adaptive in such a way that newly evolving concepts or entities will not be ignored by an old and static prediction classification model. Note that while models trained on a small number of entity mentions cannot be expected to produce high-quality automatic labels, however their annotation suggestions might still be useful for the task at hand.

We conduct two experiments to exemplify and evaluate our human-into-the-loop approach of entity mention annotation for the medical domain. In the first aspect, we simulate the interactive machine learning approach by incrementally processing the BioNLP/NLPBA 2004 named entity annotated data set [7]. During the simulation, a classifier model is first trained on very few annotations and we measure the number and quality of correctly predicted annotations in the next chunk of the data, which subsequently is added to the training, simulating the annotation process. With this simulation, we can learn whether annotating very few documents already produces reasonable and faithful predictions so that it relieves users from annotating every document in the data set.

In the second experiment, we put our approach to practice and apply it in a use case where medical professionals annotate documents in order to support research on their particular question of interest. Specifically, the task used for

this study is focused towards the investigations of the causes of the B-chronic lymphocytic leukemia (B-CLL) on MEDLINE abstracts.

In the experiment, we compare two setups where annotators are presented, or not presented with suggestions from the classifier in the interactive annotation interface. This experiment sets out to clarify whether medical professionals perceive our human-in-the-loop approach as appropriate and helpful in quantitative terms and in a qualitative assessment.

2 Related Work

Human into the Loop: Automated machine learning algorithms work well in certain environments. However, biomedical data are full of probability, uncertainty, incompleteness, vagueness, noise, etc., which makes the application of automated approaches difficult, yet often impossible. Moreover, the complexity of current machine learning algorithms has discouraged medical professionals from the application of such solutions. However, for increasing the quality of such approaches, the integration of the expert’s domain knowledge is indispensable. The interaction of the domain expert with the data would greatly enhance the whole knowledge discovery process chain. Interactive Machine Learning (iML) puts the human into the loop to enable what neither a human nor a computer could do on their own, cf. [1].

Interactive/Adaptive Learning: Static machine learning assumes that the actual state of the “domain universe” can be sufficiently acquired by listing all available data sets at particular time. In the contrast, adaptive machine learning assumes the possibility that there might exist unrecorded facts at particular time, which can only be appear at some point in the future. Authors of [8] address an industrial case study (tile manufacturing process) and found out that the classical machine learning setup faced difficulties such as 1) feedback is usually obtained after a process is completed, which might help the system, 2) some variables can change through time, and 3) error correction is always done after observation. The research by [9] on clustering a large number of documents using an interactive recommender system shows that users can sort documents into clusters significantly faster with an interactive recommender system than correcting the output of a static automated method. On top of simple user feedback in [10], such as accepting and rejecting suggestions, complex feedback like choosing the best features, suggestions for the re-weighting of features, proposing new features and combining features remarkably improve the system. Moreover, experiments in [11] examine the effect of allowing end users to do feature labeling, instead of annotating instances of training data: Especially for small amounts of training, the feature labeling approach was shown to be effective. In our work, we do not incorporate feature labeling, but we will consider it in our future work.

NER for Medical Domains: Recent years have seen a surge on Biomedical text processing (see [12] for a survey), most of which rely on the GENIA corpus [13], which is a collection of biomedical abstracts. It is mainly annotated for linguistic structures such POS tagging and syntax annotation, semantic annotation

of entities and so on [14, 15]. The work of [16] focuses on the automatic detections of multiple biomedical entities using a single-word classification approach in contrast to earlier works in the area focusing on single entity types such as proteins or genes. In this approach, they use features such as word attributes and contextual information. To alleviate the bottleneck of manual named entity annotation for medical texts, [17] have set up a crowdsourcing project on Amazon Mechanical Turk (www.mturk.com) to annotate three entity types. The research shows that using crowdsourcing is a viable alternative to annotate medical texts at scale for entity types that are understood by laymen like “medication”. However, for a more complex and fine-grained distinction that requires domain knowledge, medical professionals are required.

3 Methodology

Annotation Learning: The development of large amounts of high quality training data at one shot is hard and even undesirable [18]. Instead, an interactive machine learning methodology is more applicable where the machine-learning model is enhanced not using the prevailing train-learn-evaluate technique, but improving the model in a more iterative fashion.

Interactive learning focuses on enhancing an existing machine-learning model based on newly acquired information, which is not possible in a classical machine learning setting. The benefit of interactive learning is many-fold, such as 1) the classifier model gets better and better as new training examples are added to the training data, 2) when there is a sudden change to the underlying data set, what is known as *concept drift*, the machine-learning model gets updated accordingly [19], and 3) it largely reduces the total annotation time required to annotate the whole dataset. Most importantly, such approach will not require a pre-existing annotation dataset so that it is truly responsive and incremental, fully adaptive to the user’s need, and it makes such approach more affordable when integrated into a larger information extraction system.

As the machine-learning model can be enriched incrementally, applications employing this model will not be affected, as the system can still draw suggestions from the old model while building the new model. This approach overcomes the limitations where systems have to wait until full training and prediction cycles are completed, decreasing deployment time.

Medical NER Tagging: Medical named entity mention recognition is a well-researched area with a large number of datasets used in competitions [7, 20–23]. These mainly focus on entity/mention and chunk detections and relation extraction. Unfortunately, Biomedical annotation task is still challenging unlike other language processing tasks due to the fact that most of the annotations require highly experienced professional annotators, as discussed above.

To demonstrate the effect of interactive learning on Biomedical entity tagging, we used the BioNLP/NLPBA 2004 corpus and train a classifier using a rather generic sequence tagging system developed for German named entity recognition [24] based on CRFsuite [25]. The system is highly configurable

regarding features and data formats. For this study, we use basic standard features to characterize the text: Character and word features, which consists of the first and last character ngrams ($n=3$) of the current token as affixes, considered in a time-shifted window of two tokens around the word token in focus. We also incorporated automatically induced part-of-speech (POS) tag clusters as features, which are based on the system by [26] trained on a MEDLINE 2004 dataset. For unseen tokens in the cluster, the Pretree multi-purpose word classifier tool from the ASV toolbox [27] is used to approximate the unsupervised POS tags. Furthermore, word shape features that reflect capitalization and character classes (e.g. numbers vs. letters), were found to be relevant for biomedical mentions, as the shape of such entities often differs from non-entity tokens.

4 Annotation Problem Use Case

In this section, the use case of our medical research professionals is laid out. It focuses on understanding the interplay between risk factors and genetic presuppositions with a leukemia cancer.

B-chronic lymphocytic leukemia (B-CLL), a malignant hematopoietic neoplasm of B-lymphocytes (B cells), is the most common leukemia in the westernized world [28]. Yet, its risk factors and underlying mechanisms are still unknown. Some features of this malignancy, such as the incidence increasing with age and low proliferative capacity combined with impaired apoptosis (homeostatic cell death), categorize this disorder more as a chronic aging disease, than as a “real” leukemia, known to arise from the primary genetic defect and the subsequent block in immune cell differentiation [29]. On the other hand, accumulated evidence indicate that the pathogenesis of some commonly occurring cancers, such as breast, or colon cancer, as well as of some types of lymphomas (malignant neoplasms of the lymphoid tissue), can be explained by the complex interplay of age-related and lifestyle-related mechanisms, operating mainly through chronic inflammation and impaired insulin dependent metabolism, known as insulin resistance condition (decreased insulin action in target tissues followed by chronic hyperglycemia) [30–32].

Biological links towards cancerogenesis and lymphomagenesis go via impaired cell homeostasis mechanisms, including apoptosis and proliferation, as well as inter-cellular and intra-cellular signaling [33, 34]. Medical expert posed a hypothesis that the same risk factors and mechanisms stay also in the background of the pathogenesis of B-CLL. Exact evidence in the literature is absent. Literature search and reasoning could be demanding, because of the need to revealing many complex relationships between the numerous sets of entities and the syntagmatic constructs.

In order to alleviate the efforts of meaningful literature searching, we used the tool of adaptive annotation learning. Firstly, the medical expert prepared a set of selected abstracts, downloaded from the MEDLINE. Then, based on a limited number of specific medical entities, including CELL, CONDITION, DISORDER, GENE, MOLECULE, PROTEIN, MOLECULAR PATHWAY and

SUBSTANCE, she annotated the important structures throughout the entire text body and made them visible.

5 Experiment and Evaluation

Simulating Interactive Learning: In order to prove that interactive machine learning can yield a quality-annotated data set in a short training loop, we conduct our first experiment based on the BioNLP/NLPBA 2004 data set. The data set is divided into an increasing size of documents simulating interactive annotation. As it can be seen from Table 2 and Figure 1, a (simulated) annotation of only 40 sentences already predicted an adequate amount of suggestions where users can quickly accept or modify and proceed to the next iteration. Aiming at maximizing F-score as the harmonic mean of Precision and Recall, we can clearly observe in Tab. 1 that, after simulated annotating of about 500 sentences, the gain in performance decreases, which implies that only annotating small portion of the sentences produces reasonable suggestions that are mostly acceptable by the annotator. Also, we can see that more annotations beyond 5,000-10,000 sentences are subject to diminishing returns, i.e. it takes an increasing number of annotations to achieve the same amount of relative improvements, the more annotations are used for training. In a human-in-the-loop setting, this can be detected during the process, and could be a sign for requiring more advanced features in the machine learning setup.

Qualitative Assessment: In addition to the quantitative experimental simulation done in Section 5, we have conducted practical annotation and automation experiments using a total of 10 MEDLINE abstracts that were chosen in the context of our use case described in Section 4. For this, we have used the annotation and automation component [35] of the WebAnno web-based annotation

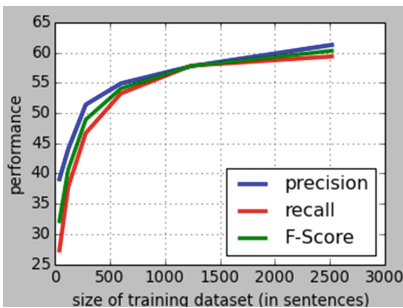


Fig. 1. Learning curve showing the performance of interactive automation for BioNLP/NLPBA 2004 data set using different sizes of training data

Table 2. Evaluation result for the BioNLP/NLPBA 2004 task using an interactive online learning approach with different sizes of training dataset tested on the fixed development dataset.

Sentences	Recall	Precision	F-score
40	27.27	39.05	32.11
120	37.74	44.01	40.63
280	46.68	51.39	48.92
600	53.23	54.89	54.05
1240	57.83	57.74	57.78
2520	59.35	61.26	60.29
5080	62.32	64.03	63.16
10200	66.43	67.50	66.96
18555	69.48	69.16	69.32

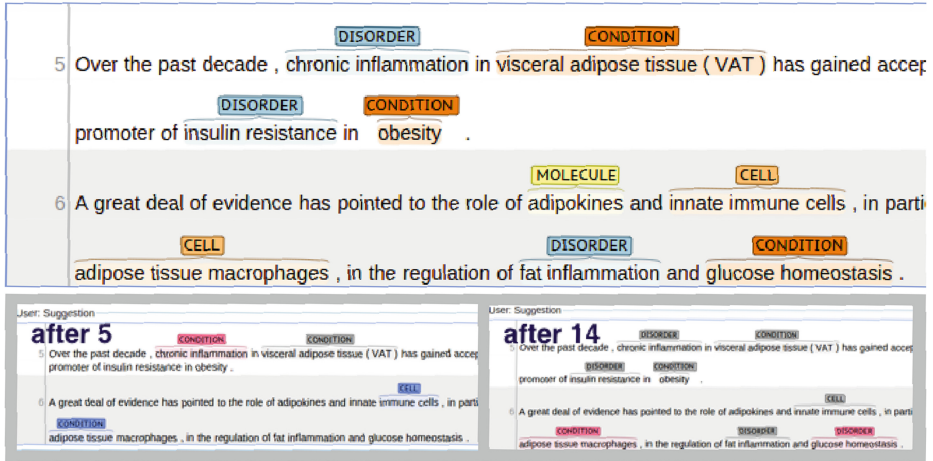


Fig. 2. Automation suggestions using the WebAnno automation component after annotating 5 initial resp. 9 additional abstracts. Correct suggestions are marked in grey, while wrong suggestions are marked in red.

tool. WebAnno [36] features a split-pane visualization, where annotation is performed in the upper pane by selecting text and choosing a label. In the lower pane, suggestions are displayed, which can be accepted and appear as annotations in the upper pane upon clicking on them, cf. Fig. 2. The experiment was conducted in two rounds. In the first round, medical experts have annotated 5 abstracts comprising a total of 86 sentences for specific medical entities as described in Section 4. Once the first round of annotations was completed, the automation was started using WebAnno’s automation component in order to provide initial suggestions. As displayed in Fig. 2, the automation component already suggests some entity annotations immediately after the first round. Using the automation suggestions, the expert continued annotating. After another 9 annotated abstracts that serve as training for the sequence tagging model, the quality and quantity of suggestions have again increased, see Fig. 2.

Qualitatively, annotators found that using the automation component, they perceived a significant increase in annotation speed. This confirms results in [37], where adaptive annotation automation in WebAnno can speed up the annotation process by a factor of 3 to 4 in comparison to a traditional annotation interface without suggestions. On a further note, the WebAnno tool was perceived as adequate and useable by our medical professionals, requiring only very limited usage instructions.

6 Conclusion and Future Outlook

In this work, we investigated the impact of adaptive machine learning for the annotation of quality training data. Specifically, we tackled medical entity recog-

dition on texts from MEDLINE, the largest collection of medical literature on the web. Identifying the need of entity tagging for applications such as information extraction, document summarization, fact exploring and relation extraction, and identifying the annotation acquisition bottleneck which is especially severe in the medical domain, we have carried out two experiments that show the utility of a human-in-the-loop approach for suggesting annotations in order to speed up the process and thus to widen the bottleneck. In the first experimental setup, we have used an existing BioNLP/NLPBA 2004 data set and run experimental simulation by incrementally processing the dataset to simulate the human in the loop. Using a generic sequence tagger, we showed that annotating very few sentences already produces enough correct predictions to be useful, suggesting that interactive annotation is a worthwhile enterprise from the beginning of an annotation project. In the second setup, we have engaged medical professionals in the annotation of medical entities in documents that were deemed relevant for the investigation of the cause of malignant B-CLL. The freely available WebAnno annotation tool (github.com/webanno) has been used for the annotation and automation process and annotators found that the adaptive annotation approach 1) makes it fast and easy to annotate medical entities, and 2) useful entity suggestions were already obtained after the annotation of only 5 MEDLINE abstracts, and suggestions subsequently improved tremendously after having annotated another 9 abstracts, reducing the annotation effort.

On a larger perspective, our results demonstrate that a paradigm change in machine learning is feasible and viable. Whereas the mantra of the past has been 'there is no (annotated) data like more (annotated) data' for supervised machine learning, suggesting large annotation efforts involving many human annotators, it becomes clear from our experiments that these efforts can be sped up tremendously by switching to an approach where the human can continuously improve the model by annotation while using the model to extract information, with the especially good news that the largest model improvements are achieved already very early in the process, as long as the domain is confined.

While such an adaptive approach to machine learning that factors in the user into the equation still calls for new evaluation methodologies to be assessed in all its aspects, it is deemed more adequate, more immediate and quicker deployable. It also fits better the shift towards an interactive, more natural, more adaptive, more contextualized and iterative approach under the umbrella of cognitive computing.

Acknowledgments. Parts of this work have been carried out with the K1 COMET Competence Center CBmed, which is funded by the Austrian Federal Ministry of Transport, Innovation and Technology (BMVIT); the Austrian Federal Ministry of Science, Research and Economy (BMWFV); Land Steiermark (Department 12, Business and Innovation); the Styrian Business Promotion Agency (SFG); and the Vienna Business Agency. The COMET program is executed by the FFG. The development of WebAnno and the research on adaptive machine learning was supported by the German Federal

Ministry of Education and Research (BMBF) as part of the CLARIN-D infrastructure and by German Research Foundation (DFG) as part of the SEMSCH project.

References

1. Holzinger, A.: Human-Computer Interaction and Knowledge Discovery (HCI-KDD): What Is the Benefit of Bringing Those Two Fields to Work Together? In: Cuzzocrea, A., Kittl, C., Simos, D.E., Weippl, E., Xu, L. (eds.) CD-ARES 2013. LNCS, vol. 8127, pp. 319–328. Springer, Heidelberg (2013)
2. Holzinger, A., Schantl, J., Schroettner, M., Seifert, C., Verspoor, K.: Biomedical Text Mining: State-of-the-Art, Open Problems and Future Challenges. In: Holzinger, A., Jurisica, I. (eds.) Interactive Knowledge Discovery and Data Mining in Biomedical Informatics. LNCS, vol. 8401, pp. 271–300. Springer, Heidelberg (2014)
3. Holzinger, A., Geierhofer, R., Modritscher, F., Tatzl, R.: Semantic information in medical information systems: Utilization of text mining techniques to analyze medical diagnoses. *JUCS* **14**, 3781–3795 (2008)
4. Holzinger, A., Yildirim, P., Geier, M., Simonic, K.M.: Quality-based knowledge discovery from medical text on the web. In: Pasi, G., Bordogna, G., Jain, L.C. (eds.) ISRL, vol. 50, pp. 145–158. Springer (2013)
5. Ghiasvand, O., Kate, R.: UWM: Disorder mention extraction from clinical text using CRFs and normalization using learned edit distance patterns. In: Proc. SemEval 2014, Dublin, Ireland (2014)
6. Leser, U., Hakenberg, J.: What makes a gene name? named entity recognition in the biomedical literature. *Briefings in Bioinformatics* **6**, 357–69 (2005)
7. GuoDong, Z., Jian, S.: Exploring deep knowledge resources in biomedical name recognition. In: Proc. NLPBA/BioNLP at COLING 2004, Geneva, Switzerland, pp. 99–102 (2004)
8. Ludl, M.C., Lewandowski, A., Dorffner, G.: Adaptive machine learning in delayed feedback domains by selective relearning. *Appl. Artif. Intell.*, 543–557 (2008)
9. Drucker, S.M., Fisher, D., Basu, S.: Helping Users Sort Faster with Adaptive Machine Learning Recommendations. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part III. LNCS, vol. 6948, pp. 187–203. Springer, Heidelberg (2011)
10. Stumpf, S., Rajaram, V., Li, L., Burnett, M., Dietterich, T., Sullivan, E., Drummond, R., Herlocker, J.: Toward harnessing user feedback for machine learning. In: Proc. 12th IUI, pp. 82–91 (2007)
11. Das, S., Moore, T., Wong, W.K., Stumpf, S., Oberst, I., McIntosh, K., Burnett, M.: End-user feature labeling: Supervised and semi-supervised approaches based on locally-weighted logistic regression. *Artif. Intell.* **204**, 56–74 (2013)
12. Cohen, A.M., Hersh, W.R.: A survey of current work in biomedical text mining. *Briefings in Bioinformatics* **6**, 57–71 (2005)
13. Ohta, T., Tateisi, Y., Kim, J.D.: The GENIA corpus: An annotated research abstract corpus in molecular biology domain. In: Proc. Human Language Technology Research, HLT 2002, San Francisco, CA, USA, pp. 82–86 (2002)
14. Tateisi, Y., Tsujii, J.: Part-of-speech annotation of biology research abstracts. In: Proc. LREC 2004, Lisbon, Portugal, pp. 1267–1270 (2004)
15. Tateisi, Y., Yakushiji, A., Ohta, T., Tsujii, J.: Syntax annotation for the GENIA corpus. In: Proc. IJCNLP 2005, Lisbon, Portugal, pp. 222–227 (2005)

16. Lee, C., Hou, W.J., Chen, H.H.: Annotating multiple types of biomedical entities: A single word classification approach. In: Proc. Int'l Joint Workshop on NLP in Biomedicine and Its Applications, pp. 80–83 (2004)
17. Yetisgen-Yildiz, M., Solti, I., Xia, F., Halgrim, S.R.: Preliminary experience with amazon's mechanical turk for annotating medical named entities. In: Proc. NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk, pp. 180–183 (2010)
18. Vidulin, V., Bohanec, M., Gams, M.: Combining human analysis and machine data mining to obtain credible data relations. *Information Sciences* **288**, 254–278 (2014)
19. Hoens, T.R., Chawla, N.V.: Learning in non-stationary environments with class imbalance. In: Proc. 18th ACM SIGKDD, New York, NY, USA, pp. 168–176 (2012)
20. Uzuner, Ö., Luo, Y., Szolovits, P.: Evaluating the state-of-the-art in automatic de-identification. *J Am. Med. Inform. Ass.* **14**, 550–563 (2007)
21. Uzuner, Ö., Solti, I., Xia, F., Cadag, E.: Community annotation experiment for ground truth generation for the i2b2 medication challenge. *J Am. Med. Inform. Ass.* **17**, 561–570 (2010)
22. Kim, J.D., Ohta, T., Pyysalo, S., Kano, Y., Tsujii, J.: Overview of BioNLP 2009 shared task on event extraction. In: Proc. BioNLP 2009, pp. 1–9 (2009)
23. Kim, J.D., Pyysalo, S., Ohta, T., Bossy, R., Nguyen, N., Tsujii, J.: Overview of bionlp shared task 2011. In: Proc. BioNLP, pp. 1–6 (2011)
24. Benikova, D., Yimam, S.M., Santhanam, P., Biemann, C.: GermaNER: free open german named entity recognition tool. In: Proceedings of the GSCL 2015, Duisburg, Germany (2015)
25. Okazaki, N.: CRFsuite: a fast implementation of Conditional Random Fields (CRFs) (2007)
26. Biemann, C.: Unsupervised Part-of-Speech Tagging in the Large. *Res. Lang. Comput.*, 101–135 (2009)
27. Biemann, C., Quasthoff, U., Heyer, G., Holz, F.: Asv toolbox - a modular collection of language exploration tools. In: Proc. LREC 2008, pp. 1760–1767 (2008)
28. Brown, J.R.: Inherited susceptibility to chronic lymphocytic leukemia: evidence and prospects for the future. *Ther Adv Hematol* **4**, 298–308 (2013)
29. Nieto, W.G., Teodosio, C., López, A., Rodríguez-Caballero, A., Romero, A., Bárcena, P., Gutierrez, M.L., Barez Hernandez, P., Carreño Luengo, M.T., Casado Romo, J.M., Cubino Luis, R., De Vega Parra, J.: Non-cll-like monoclonal b-cell lymphocytosis in the general population: Prevalence and phenotypic/genetic characteristics. *Cytometry Part B* **78B**, 24–34 (2010)
30. Larsson, S.C., Wolk, A.: Obesity and risk of non-Hodgkin's lymphoma: A meta-analysis. *International Journal of Cancer* **121**, 1564–1570 (2007)
31. Tsugane, S., Inoue, M.: Insulin resistance and cancer: Epidemiological evidence. *Cancer Science* **101**, 1073–1079 (2010)
32. Bastard, J.P., Maachi, M., Lagathu, C., Kim, M.J., Caron, M., Vidal, H., Capeau, J., Feve, B.: Recent advances in the relationship between obesity, inflammation, and insulin resistance. *European Cytokine Network* **17**, 4–12 (2006)
33. Ginaldi, L., De Martinis, M., Monti, D., Franceschi, C.: The immune system in the elderly. *Immunologic Research* **30**, 81–94 (2004)
34. Le Marchand-Brustel, Y., Gual, P., Grémeaux, T., Gonzalez, T., Barrès, R., Tanti, J.-F.: Fatty acid-induced insulin resistance: role of insulin receptor substrate 1 serine phosphorylation in the retroregulation of insulin signalling. *Biochem. Soc. Trans.* **31**, 1152–1156 (2003)

35. Yimam, S., Eckart de Castilho, R., Gurevych, I., Biemann, C.: Automatic annotation suggestions and custom annotation layers in WebAnno. In: Proc. ACL 2014 System Demonstrations, Baltimore, MD, USA, pp. 91–96 (2014)
36. Yimam, S., Gurevych, I., Eckart de Castilho, R., Biemann, C.: WebAnno: A flexible, web-based and visually supported system for distributed annotations. In: Proc. ACL 2013 System Demonstrations, Sofia, Bulgaria, pp. 1–6 (2013)
37. Yimam, S.M.: Narrowing the loop: Integration of resources and linguistic dataset development with interactive machine learning. In: Proc. HLT-NAACL: Student Research Workshop, Denver, Colorado, pp. 88–95 (2015)