Extracting Impact Model Narratives from Social Services' Text

Bart Gajderowicz[1], Daniela Rosu[1] and Mark S. Fox[1]

[1]*Department of Mechanical & Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, Ontario, M5S 3G8, Canada*

Abstract

Named entity recognition (NER) is an important task in narration extraction. Narration, as a system of stories, provides insights into how events and characters in the stories develop over time. This paper proposes an architecture for NER on a corpus about social purpose organizations. This is the first NER task specifically targeted at social service entities. We show how this approach can be used for the sequencing of services and impacted clients with information extracted from unstructured text. The methodology outlines steps for extracting ontological representation of entities such as needs and satisfiers and generating hypotheses to answer queries about impact models defined by social purpose organizations. We evaluate the model on a corpus of social service descriptions with empirically calculated score.

Keywords

Named entity recognition, narrative extraction, rule-based reasoning, social services

1. Introduction

In Canada, there are over 65,000 social service providers, each delivering a specialized set of services administered as a variety of programs to different populations in need. These populations include the approximately 4.9 million Canadians living in poverty, along with the additional 401,000 new immigrants and 44,000 refugees that immigrated to Canada in 2021 alone. The Canadian federal government plays an key role in coordinating with private and public agencies to fund the vast majority of social programs active today. However, the task of service provisioning and management is largely the responsibility of individual social purpose organizations (SPO). Navigating the myriad of SPOs to address an individual's (client) needs is a daunting task fraught with errors. Determining the efficacy of the services assigned to a client is also a challenge. As of this writing, no system can manage such a large network of services and clients in a cohesive manner, but such a system would benefit all stakeholders involved.

In our previous work we have developed an ontology that can be used by an SPO to represent its Impact Management models [1]. The term impact "*refers to the intended and unintended (positive or otherwise) changes (outputs, outcomes) that occur across the organization (within and/or across its programs) and with its stakeholders (including users, clients, partners, etc.) over a period*

✉ bartg@mie.utoronto.ca (B. Gajderowicz); drosu@mie.utoronto.ca (D. Rosu); msf@mie.utoronto.ca (M. S. Fox)
🌐 http://bartg.org (B. Gajderowicz); http://eil.utoronto.ca (M. S. Fox)
🆔 0000-0001-6201-8781 (B. Gajderowicz); 0000-0002-5877-9681 (D. Rosu); 0000-0001-7444-6310 (M. S. Fox)

of time (short term, long term) as a result of the organization's activities."[1] Experts have developed numerous Impact Models to help SPOs articulate the change they seek to achieve and how that change is achieved [2, 3, 4, 5, 6]. The Common Impact Data Standard (CIDS) [1] ontology defines classes and relationships that span impact modelling concepts such as Program, Service, Activity, Stakeholder, Outcome, Indicators and Risk. It can be used to define the services an SPO provides and the requirements needed for a client to receive a service. Our most recent ontology research extends CIDS to include client needs (e.g., housing, food) and various ways they can be satisfied (e.g. women shelters, food banks).

In this paper, we describe our efforts in addressing the problem of matching client needs to SPO services. In order to match client needs, we must represent the services an SPO provides and how they satisfy needs, and the characteristics (e.g., age, gender, occupation) and needs of clients for whom they were designed. Although we have the means to represent an SPOs impact model, the information needed to instantiate each SPO's model is buried in a variety of textual sources, such as service descriptions, client success stories, and eligibility criteria.

This paper presents an approach, based on Named Entity Recognition (NER), to extracting an SPO's impact model, i.e., "narrative", from various text sources. NER is a crucial component in understanding the narrative of a given text. By narrative, we mean a "*system of stories structured in such a way as to achieve a rhetorical purpose or vision*" [7]. In the context of SPOs, we define each service as a "story" describing what they offer, to whom, how, and when. The narrative, then, is a system of services that guide clients through various programs towards achieving their goals. The extracted information can then be provided in the same language as the problem domain, allowing for culture-wide, community-based, and individual-level analysis [7].

There are several **challenges** that we face when extracting terms that represent client needs and specific resources that services provide. Consider the example

> "*We welcome clients of all ages and offer services that benefit our core clients (young families, chronically homeless): mental health, education needs; community outreach.*"

that provides required information but is hard to parse. Firstly, there is a lack of language models for identifying social service client characteristics and needs, and how their services satisfy needs. Often, vague descriptors cause confusion about entities: what is a program, what is a service, what is the resource, what is the need. There are no standardized labels across SPO programs, services, clients, and eligibility criteria. Structured data about services is limited, while unstructured text describes various aspects of the service that are hard to infer, such as promoting services versus listing service details, or describing serviced communities versus listing client requirements. Finally, information related to the scheduling or expected sequence of services is often incomplete or unknown, such as quality, service capacity, or availability.

2. Related Work

2.1. Named Entity Recognition

Named Entity Recognition (NER) is a method for identifying the types of terms in unstructured text. Common terms include a person, organization, place, date, currency, and numbers [8]. As

---

[1]https://innoweave.ca/en/modules/impact-measurement

will be described in the following sections, three main approaches used are: 1) a rule-based models for identifying types of terms, 2) a learned model that can infer types found in a training set of text and types, and 3) a hybrid rule-learned model. The rule-based method can identify patterns in text when proper sentence grammar is not followed, or entities are not common enough to be found in a training set (e.g. business names). However, a rule-based method requires manual evaluation of the data and manual rules construction for observed patterns. The trained method can match words in a sentence to its learned vocabulary and assign their type, but is limited to words in the training set. Hybrid models try to take advantage of both rules and learned methods, making best guesses to infer entities not present in the training set.

There is a number of pre-trained language models capable of named entity recognition. Each one is trained on either a specialized or general dataset. Schmitt et al. [9] performed a comprehensive analysis of the performance and applicability of the five most popular packages, including StanfordNLP, NLTK, OpenNLP, SpaCy, and Gate with new ones being developed on an ongoing basis, such as HuggingFace [10]. These packages are trained on varying datasets, the two largest being Common Crawl (http://commoncrawl.org), a database of content crawled on the internet, and English Wikipedia (https://www.wikipedia.org). NER benchmark datasets include MUC-6 [11] and MUC-7 [12] and ACE [13]. The resulting models generally provide support for a standard list of entity types [8]. Unfortunately, existing models are not well suited for social services as they have not been trained on related corpa identify required entities.

2.2. Other Methods

In this section, we highlight several approaches and their methods that identify entities in the text and can assist in building an SPO's impact model narrative. **Linguistic properties** alone provide a great deal of structure to the text being analyzed. For example, Chiarello et al. [14] use linguistic features to identify stakeholders across documents, while Hussain et al. [15] rely on grammar rules to generate narratives, identify keywords and extract important phrases in social media posts. **Query-driven** methods rely on a "seed" query and external vocabulary to guide the search algorithm and find a suitable label in order to perform query answering [16], query modelling [17, 18], and query extensions [19] tasks. **Rule-Based methods** are suitable when a training corpus does not exist, and a set of a *priori* rules provide context for extracting various entities. This includes predefined rules for finding clues for query answering [20] and grammar rules for narrative extraction [15, 21], and to reason about extracted entities [22].

Statistical models rely on data-driven algorithms and encompass both frequency-based and probability-based models [23, 24] for ranking found entities [25], group generalization [17, 19], and calculating similarity scores between documents [20]. **Machine learning** methods include deep learning architectures for NER tasks [26]. Several have characteristics useful for narrative extraction, such as temporal factors, rules, or linguistic properties, and utilize methods such as BiLSTMs [27, 28, 29, 30], ELMo [28, 31, 32, 33], and BERT [34, 35, 36, 37, 38, 39].

3. Methodology

This section summarizes our methodology for performing the NER task in the SPO domain. Our proposed NER architecture is depicted in Figure 1. The input is a corpus of unstructured text
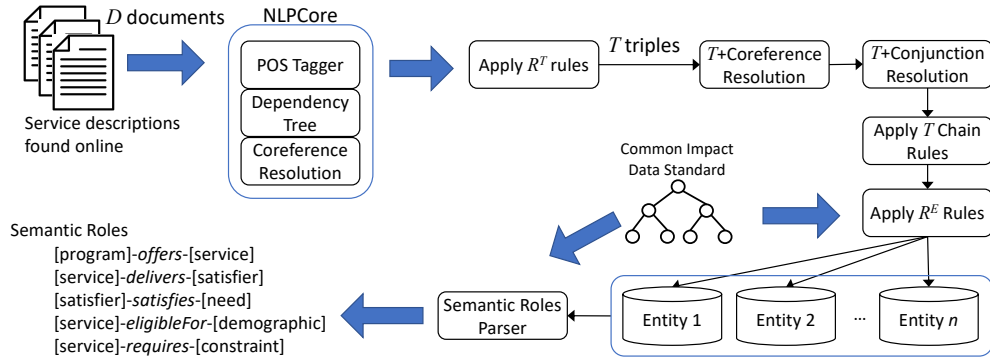
Figure 1: Named Entity Extraction Architecture.

Table 1
SPO-related entity types.

Entity Type	Description
Program Name	Name of the program.
Need Satisfier	Label for the need satisfier that focuses more on the offered resources.
Client Characteristic	A client's characteristics, as defined by the service provider.
Need	The client's need being addressed.
Desired State (outcome)	The desired state of the client, which may more broadly describe the outcome.
Service Description	A broader description of the service being offered.
Client Description	A broader description of the client.
Need Satisfier Description	A broader description of the need satisfier.
Required Criteria	Criteria for using a service as representing a requirement or eligibility.

Table 2
Definitions.

Variables	Description
$d \in D$	Documents d in a corpus D.
$e \in E$	Entities to be extracted from text in D.
$t_x \in T$	Subject-predicate-object triples.
s, p, o	Each triple is comprised of three slots, where $t_x = s(subject), p(predicate), o(object)$
$T_d \subseteq T$	A subset of triples in T extracted from document d.
$r_i^t \in R^T$	Rules identifying triples in $t_x \in T$.
$r_i^E \in R^E$	Rules identifying entities in $e \in E$.
mcc_i	Matthews correlation coefficient (MCC) score for rule r_i^E.
$r_i^E(T_d)$	Classification of entity e by r_i^E from triples in T_d.
$\widetilde{w}_i$	Weight of rule r_i^E in correctly identifying an entity e, as per Equation 2.

describing SPOs. It incorporates the Common Impact Data Standard [1] ontology to identify which entities to extract, then again to generate semantic roles by providing the semantic relationships between the entities. The terms we are interested in are listed in Table 1. They capture key concepts in describing an SPOs logic model, and form the basis of their "narrative" in how services are delivered to clients, what needs they are satisfying, how, and when. We also introduce several definitions used by our NER model in Table 2, and related equations.

$$r_i^E(T_d) = \begin{cases} 1, & \text{if } mcc_i > 0 \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

$$\widetilde{w}_i = r_i^E(T_d) \times mcc_i \tag{2}$$

$$w_e = \frac{\sum \widetilde{w}_i}{|r^E|} \text{ for all rules } r_i^E \in r^E \text{ that extract entity } e. \tag{3}$$

3.1. Annotating with Linguistic Properties

Our method begins by relying on a Stanford NLPCore parser [40] to generate a set of linguistic properties about the service descriptions. We use its part-of-speech (POS) tagger to identify nouns, verbs, adjectives, and so on. Second, the parser creates a dependency tree identifying word modifiers, conjunctions, as well as subjects, predicates, and objects. Third, the parser generates coreference resolutions between terms, associating pronouns like "they" and "our" with the nouns or proper nouns they refer to. Accuracy and further processing is limited by the accuracy of the dependency trees and coreference resolutions generated by the NLPCore parser.

3.2. Semantic Role Triple Extraction

Once the parser has annotated the text with linguistic properties, custom rules $r_x^T \in R^T$ combine key dependencies to form subject-predicate-object triples $t_x \in T$. In some literature, the triple relation is referred to as subject-verb-object (SVO), but T-triples represent a broader structure that does not rely on verbs as predicates alone. Each triple contains three slots: a subject (s), a predicate (p), and an object (o), forming the structure:

$$t_x = \{ \text{s("subject"), p("predicate"), o("object")} \} \qquad .$$

For example, consider the sentence "St. Mary's provides education services." Here we see that "St. Mary's" is the subject, "provides" is the predicate, and "education services" is the object. Consider a rule r_i^T where, given the three terms A, B, and C, and dependencies $(nsubj, obj, obl)$

      If   a $nsubj$ dependency exists between B and A,
               an obj dependency exists between B and C,
               an obl dependency does not exist between B any other term,
    Then $t_x = \{s(A), p(B), o(C)\}$.

By applying this rule to the sentence above, we can infer the T triple:

$$t_x = \{ \text{s("St Mary's"), p(provides), o("education services")} \}.$$

While this example rule is easily inferred from the dependencies alone, 18 rules $r_i^T \in R^T$ have been empirically identified to extract subject-predicate-object relationships.

3.3. Coreference and Conjunction Resolution

Next, each T-slot is extended with their coreference and conjunction terms, if any, using a depth-first search. For example, in the sentences "*St. Mary's provides education services. They also prepare hot meals.*", the pronoun "They" refers to the proper noun "St. Mary's". Hence we infer that in addition to "education services", "St Mary's" also provides "hot meals", giving:

$$t_1 = \{ \text{s("St Mary's"), p("provides"), o("education services")} \}$$
$$t_2 = \{ \text{s("St Mary's"), p("prepares"), o("hot meals")} \}$$

Next, we resolve conjunctions with terms in each T-slot. Conjunctions are lists of terms connected by terms like a comma, "and" and "or". For example, given the sentence "*St. Mary's provides education services, a soup kitchen, and religious counselling.*", we see that all terms following "provides" are of the same type, a "need satisfier."

Like subjects and objects, the predicate can also be a conjunction. For example, in the sentence "*St. Mary's provides education services and prepares hot meals.*", "provides" and "prepares" are both verbs connected as a conjunction in the dependency tree, and hence both have "St Mary's" as their subject. However, they each have their own object, producing two T-triples, namely

$$t_1 = \{ \text{s("St Mary's")}, \ \text{p("provides")}, \ \text{o("education services")} \}$$
$$t_2 = \{ \text{s("St Mary's")}, \ \text{p("prepares")}, \ \text{o("hot meals")} \}$$

To ensure we capture all combinations of T-triples, our model uses a depth-first search to generate hypotheses for all combinations of connected subjects, predicates, and objects.

3.4. Chaining Rules: From Triples To Stories

Given a list of T-triples, and coreferences and conjunctions resolved, we build a chain of T-triples that provide additional structure to the terms in the text. The rules simply connect object slot (X) values in one t_1 triple to subject slot $s(X)$ values in another t_2 triple,

$$\text{If} \quad t_1 = \{s(A), p(B), o(C)\} \text{ and } t_2 = \{s(C), p(D), o(E)\}$$
$$\text{Then} \quad [\{s(A), p(B), o(C)\}; \{s(C), p(D), o(E)\}] \text{ is a } T\text{-chain and}$$
$$t_3 = \{s(A), p(D), o(E)\}.$$

In the sentence "*St. Mary's provides education services to adult learners.*" we see two T-triples:

$$t_1 = \{ \text{s("St. Mary's")}, \ \text{p("provides")}, \ \text{o("education services")} \}$$
$$t_2 = \{ \text{s("education services")}, \ \text{p("to")}, \ \text{o("adult learners")} \}$$

Chaining them together with the rule above using the "education services" term, we can infer that "St Mary's" offers services to adult learners, generating a new T triple:

$$t_3 = \{ \text{s("St. Mary's")}, \ \text{p("to")}, \ \text{o("adult learners")} \}$$

3.5. Named Entity Extraction Rules

From T-triples and T-chains, we can apply additional rules $r_i^E \in R^E$ to extract named entity types. For example, we see that "St Mary's" is the program, "education services" is the need satisfier, and "adult learners" are the clients. The rules utilize all available information about the text, including POS tags, dependencies, and their T-slots. For example, given terms A, B, C:

> If $\quad t_x = \{s(A), p(B), o(C)\}$, where A is a proper noun, B is a synonym for "offers",
> $\quad\quad B$ is a 3rd person singular present verb, and C is a plural noun
> Then $\quad A$ is a program and C is a need satisfier.

Here, synonyms for "offers" have been empirically identified as keywords used by services providers to describe what need satisfiers they offer to clients, and include the terms "provides", "offers", "offer", "provide", "provided", "offered", and "offering". Similar extractions can be performed for additional semantics defined by Common Impact Data Standard [1] such as, [program]-offers-[service description], [service description]-delivers-[need satisfier]. [need satisfier]-satisfies-[need]. [service description]-eligibleFor-[client demographic], and [service description]-requires-[constraint].

4. Evaluation

The evaluation of our model is based on the performance of each rule, aggregated by entity type into a single entity score, namely w_e. The data contains information about SPOs, provided by Help Seeker Technologies (https://helpseeker.co). The testing data consists of 16,048 documents $d \in D$ that contain SPO descriptions. Of those, 7,359 documents had a total of 76,592 unique T-triples extracted. Constructed from the triples, there were 147,299 T-chains found in 6,260 documents. Of those documents that had T-triples, 4,860 descriptions had at least one term extracted. In total 366,588 terms were extracted, and of those 48,729 were assigned an entity type using rules in R^E.

To evaluate the model, a number of documents were selected randomly for each entity, and the extracted terms were manually analyzed. Table 3 lists the results of each rule r_i^E identifying an entity e. The rule's label indicates its number and which slot was used (e.g. Rule = "o-45" means the "o" slot for rule 45). All rules rely on T-triples. Those marked with ($\sharp$) also rely on T-chains. Each entity e extracted from a document's triples T_d was classified as correct (1) or incorrect (0), as per Equation 1.

We note that not all entities have the same number of rules and not all entities are covered equally. For example, the Need Satisfier entity has the largest coverage with 13 rules while Client Characteristics, Desired States, and Required Criteria only have one. We also note that the MCC score is sensitive to large discrepancies between true (TP,TN) and false (FP,FN) values. Rules that identify significantly more true values but are not good at excluding false values can produce a negative MCC score, despite a high F-score, as marked by (§), and include Client Characteristic and Required Criteria.

The model's performance is evaluated by its aggregate score for each entity, namely w_e, as per Equation 3. The evaluation uses the ROC-AUC score to determine whether a high w_e weight correlates with correct classification. The results are listed in the AUC column of Table 3. Any $AUC \geq 0.7$ is considered acceptable, and marked by ($\flat$).

Based on the w_e score, the model has good performance on extracting the "Client Description", "Need Satisfier", "Need Satisfier Description" and "Program Name" entities. The model also performs well with a high F-score on the "Client Characteristic" and "Requirement Criteria" entities but resulted in a low w_e due to a negative MCC score. In cases where the MCC score was negative with a high F-score, we point out that if accuracy metrics (precision and recall) for the rule are high, the rule performed well on NER tasks but is limited to true positives only.

Relying on the entities that were correctly extracted, we can construct a set of semantic roles to build SPO narratives. For example, consider a particular program that delivers language classes (need satisfier) to new immigrants (a client characteristic). We can specify what requirements these clients must meet before receiving these satisfiers, such as language skill assessments. Knowing that another program offers language skill assessments, we can connect the two programs, defining a chain of SPO programs.

Table 3

Rule statistics and score based on MCC, and an aggregate model score w_e evaluated by AUC.

Rule	TN	FP	FN	TP	F-S	MCC/w_e	AUC
Client Characteristic							
o-47♯	28	240	444	2605	**0.88**§	-0.03	
Aggr	268	0	3049	0	0	0	0.52
Client Description							
p-34	32	0	26	16	0.55	0.46	
o-42♯	31	1	42	0	0	-0.13	
o-43♯	31	1	42	0	0	-0.13	
Aggr	32	0	26	16	0.55	0.46	**0.70**♭
Desired State (outcome)							
o-20	1	44	3	40	0.63	-0.11	
Aggr	45	0	43	0	0	0	0.52
Need							
o-8	36	7	69	6	0.14	-0.13	
o-9	10	33	28	47	0.61	-0.14	
Aggr	43	0	75	0	0	0	0.59
Need Satisfier Description							
o-32	5	0	62	21	0.4	0.14	
o-18	3	2	31	52	**0.76**§	0.11	
Aggr	3	2	10	73	0.92	0.31	**0.79**♭
Program Name							
s-35	39	18	52	219	**0.86**§	0.42	
s-44♯	55	2	209	62	0.37	0.19	
s-45♯	24	33	247	24	0.15	-0.49	
Aggr	39	18	52	219	0.86	0.42	**0.79**♭
Required Criteria							
o-28	3	33	41	98	**0.73**§	-0.2	
Aggr	36	0	139	0	0	0	0.61

Rule	TN	FP	FN	TP	F-S	MCC/w_e	AUC
Need Satisfier							
o-44♯	2913	74	14996	2979	0.28	0.14	
s-46♯	2896	91	15042	2933	0.28	0.13	
o-36	2701	286	13846	4129	0.37	0.11	
o-2	2916	71	16862	1113	0.12	0.06	
o-1	2987	0	17950	25	0	0.01	
o-3	2874	113	17301	674	0.07	0	
o-31	2979	8	17931	44	0	0	
o-16	2962	25	17865	110	0.01	-0.01	
o-25	2981	6	17958	17	0	-0.01	
o-11	2919	68	17847	128	0.01	-0.06	
s-12	2856	131	17747	228	0.02	-0.08	
o-13	1738	1249	13800	4175	0.36	-0.15	
o-22	2628	359	17442	533	0.06	-0.16	
Aggr	2485	502	7994	9981	0.7	0.27	**0.73**♭
Service Description							
o-23	20	0	35	0	0	0	
o-24	20	0	35	0	0	0	
o-33	20	0	35	0	0	0	
o-45♯	20	0	35	0	0	0	
o-15	13	7	25	10	0.38	-0.07	
s-41♯	19	1	35	0	0	-0.18	
s-43♯	19	1	35	0	0	-0.18	
o-19	14	6	30	5	0.22	-0.19	
Aggr	20	0	35	0	0	0	0.66

♯ a rule based on T-chains   § a high F-score > 0.7   ♭acceptable $AUC \geq 0.7$ value for a given w_e model score.

5. Conclusion and Future Work

In this paper, we propose a model for extracting SPO-related entities from descriptions. We also present the challenges and state of the NLP field, namely its lack of SPO-related corpora and pre-trained language models. Our model relies on our previous work for representing social services entities and semantics, namely the Common Impact Data Standard ontology [1], needed to capture an SPO's impact model as a "narrative" about their organization, services, and clients. Based on available data, the model relies on linguistic properties as well as empirically derived rules to identify phrases, construct T-triples, and classify phrases as entity types. Without any external data sources to seed the model with annotated text, the model performs well on certain entities, namely need satisfiers, program names, service descriptions, and required criteria.

In future work, the model will be extended with additional features and training data. Negated phrases will generate semantics that negate a relationship, such as "does not offer". A larger corpus with correct annotations will allow for better scoring methods, more rules, the use of statistical methods, and the training of supervised machine learning models. Finally, by incorporating available data associated with specific entities and extracted SPO narratives, we could perform analysis on an SPO's performance and suitability at a given time. For example, we cloud track a client's development as they transition from one program to another, based on the paths they take, the need satisfiers they qualify for, and ultimately use.

References

[1] Common Approach to impact measurement, M. S. Fox, K. Ruff, A. Chowdhury, B. Gajderow-icz, T. Abdulai, J. Zhang, The Common Impact Data Standard: An Ontology for Representing Impact, Technical Report, 2021.

[2] Practical Concepts Incorporated, The logical framework: A Manager's guide to a scientific approach to design & evaluation, Practical Concepts Incorporated, 1979.

[3] C. H. Weiss, Theory-Based Evaluation: Past, Present, and Future, New Directions for Evaluation (1997) 41–55.

[4] S. Earl, F. Carden, T. Smutylo, Outcome mapping: Building learning and reflection into development programs, International Development Research Centre, Ottawa, 2001.

[5] E. Harries, L. Hodgson, J. Noble, Creating your theory of change: NPC's practical guide, Technical Report November, NPC, 2014. URL: http://www.thinknpc.org/publications/creating-your-theory-of-change/.

[6] J. Nicholls, E. Lawlor, E. Neitzert, T. Goodspeed, A guide to social return on investment, Technical Report, Social Value UK, Liverpool, United Kingdom, 2012. URL: https://socialvalueuk.org/resource/a-guide-to-social-return-on-investment-2012/.

[7] S. W. Ruston, More than just a story: Narrative insights into comprehension, ideology, and decision making, in: Modeling Sociocultural Influences on Decision Making: Understanding Conflict, Enabling Stability, 2016, pp. 27–42. doi:10.1201/9781315369587.

[8] S. S. Pradhan, N. Xue, R. Weischedel, M. Palmer, M. Marcus, E. Hovy, S. S. Pradhan, L. Ramshaw, N. Xue, A. Taylor, J. Kaufman, M. Franchini, Others, Ontonotes Release 5.0, Linguistic Data Consortium, Philadelphia, PA 23 (2013). doi:10.3115/1620950.1620956.

[9] X. Schmitt, S. Kubler, J. Robert, M. Papadakis, Y. Letraon, A Replicable Comparison Study of NER Software :, 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS) (2019) 338–343.

[10] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. Le Scao, S. Gugger, M. Drame, Q. Lhoest, A. Rush, Transformers: State-of-the-Art Natural Language Processing (2020) 38–45. doi:10.18653/v1/2020.emnlp-demos.6. arXiv:arXiv:1910.03771v5.

[11] R. Grishman, B. Sundheim, Design of the MUC-6 evaluation, 6th Message Understanding Conference, MUC 1995 - Proceedings (1995) 1–11. doi:10.3115/1119018.1119072.

[12] N. Chinchor, P. Robinson, MUC-7 Named Entity Task Definition, Proceedings of the Sixth Message Understanding Conference MUC6 (1997) 21.

[13] G. Doddington, A. Mitchell, M. Przybocki, L. Ramshaw, S. Strassel, R. Weischedel, The automatic content extraction (ACE) program tasks, data, and evaluation, Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC 2004 (2004) 837–840.

[14] F. Chiarello, L. Trivelli, A. Bonaccorsi, G. Fantoni, Extracting and mapping industry 4.0 technologies using wikipedia, Computers in Industry 100 (2018) 244–257. URL: https://doi.org/10.1016/j.compind.2018.04.006. doi:10.1016/j.compind.2018.04.006.

[15] M. N. Hussain, K. K. Bandeli, H. A. Rubaye, N. Agarwal, Stories from blogs: Computational

extraction and visualization of narratives, in: Fourth International Workshop on Narrative Extraction from Texts, at 43rd European Conference on Information Retrieval, volume 2860, Luca, Itally, 2021, pp. 33–40. URL: https://btracker.host.ualr.edu.

[16] E. Sciore, Query Processing, in: O. Curé, G. Blin (Eds.), RDF Database Systems, Morgan Kaufmann, Boston, 2015, pp. 145–167. URL: https://www.sciencedirect.com/science/article/pii/B9780127999579000067. doi:https://doi.org/10.1016/B978-0-12-799957-9.00006-7.

[17] K. Balog, M. Bron, M. De Rijke, Category-based query modeling for entity search, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 5993 LNCS, 2010, pp. 319–331. doi:10.1007/978-3-642-12275-0_29.

[18] N. Craswell, G. Demartini, J. Gaugaz, T. Iofciu, L3S at INEX 2008: Retrieving entities using structured information, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 5631 LNCS (2009) 253–263. doi:10.1007/978-3-642-03761-0_26.

[19] K. Balog, M. Bron, M. De Rijke, Query modeling for entity search based on terms, categories, and examples, ACM Transactions on Information Systems 29 (2011). doi:10.1145/2037661.2037667.

[20] D. Garigliotti, K. Balog, On type-aware entity retrieval, in: ICTIR 2017 - Proceedings of the 2017 ACM SIGIR International Conference on the Theory of Information Retrieval, Association for Computing Machinery, Inc, 2017, pp. 27–34. doi:10.1145/3121050.3121054. arXiv:1708.08291.

[21] P. Quaresma, V. Beires Nogueira, K. Raiyani, R. Bayot, T. Gonçalves, From Textual Information Sources to Linked Data in the Agatha Project, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 12057 LNAI, 2020, pp. 79–88. URL: http://arxiv.org/abs/1909.05359. doi:10.1007/978-3-030-46714-2_5. arXiv:1909.05359.

[22] M. C. McCord, J. W. Murdock, B. K. Boguraev, Deep parsing in Watson, IBM Journal of Research and Development 56 (2012) 1–15. doi:10.1147/JRD.2012.2185409.

[23] L. Hong, B. D. Davison, Empirical study of topic modeling in twitter, in: Proceedings of the first workshop on social media analytics, 2010, pp. 80–88.

[24] S. Robertson, H. Zaragoza, The probabilistic relevance framework: BM25 and beyond, volume 3, 2009. doi:10.1561/1500000019.

[25] P. H. Oza, L. Dietz, Which entities are relevant for the story?, in: CEUR Workshop Proceedings, volume 2860, 2021, pp. 41–48. URL: http://ceur-ws.org/.

[26] J. Li, A. Sun, J. Han, C. Li, A Survey on Deep Learning for Named Entity Recognition, IEEE Transactions on Knowledge and Data Engineering 34 (2022) 50–70. doi:10.1109/TKDE.2020.2981314. arXiv:1812.09449.

[27] B. Taillé, V. Guigue, P. Gallinari, Contextualized embeddings in named-entity recognition: An empirical study on generalization, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12036 LNCS (2020) 383–391. doi:10.1007/978-3-030-45442-5_48. arXiv:2001.08053.

[28] M. Affi, C. Latiri, BE-BLC: BERT-ELMO-based deep neural network architecture for English named entity recognition task, Procedia Computer Science 192 (2021) 168–181.

URL: https://doi.org/10.1016/j.procs.2021.08.018. doi:10.1016/j.procs.2021.08.018.

[29] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, C. Dyer, Neural architectures for named entity recognition, 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL HLT 2016 - Proceedings of the Conference (2016) 260–270. doi:10.18653/v1/n16-1030. arXiv:1603.01360.

[30] Z. Jie, W. Lu, Dependency-guided LSTM-CRF for named entity recognition, EMNLP-IJCNLP 2019 - 2019 Conference on Empirical Methods in Natural Language Processing and 9th International Joint Conference on Natural Language Processing, Proceedings of the Conference (2019) 3862–3872. doi:10.18653/v1/d19-1399. arXiv:1909.10148.

[31] Y. Peng, S. Yan, Z. Lu, Transfer learning in biomedical natural language processing: An evaluation of BERT and ELMo on ten benchmarking datasets, BioNLP 2019 - SIGBioMed Workshop on Biomedical Natural Language Processing, Proceedings of the 18th BioNLP Workshop and Shared Task (2019) 58–65. doi:10.18653/v1/w19-5006. arXiv:1906.05474.

[32] M. Ulčar, M. Robnik-Šikonja, Cross-lingual alignments of ELMo contextual embeddings (2021) 1–30. URL: http://arxiv.org/abs/2106.15986. arXiv:2106.15986.

[33] C. Dogan, A. Dutra, A. Gara, A. Gemma, L. Shi, M. Sigamani, E. Walters, Fine-Grained Named Entity Recognition using ELMo and Wikidata (2019). URL: http://arxiv.org/abs/1904.10503. arXiv:1904.10503.

[34] T. Moon, P. Awasthy, J. Ni, R. Florian, Towards Lingua Franca Named Entity Recognition with BERT (2019). URL: http://arxiv.org/abs/1912.01389. arXiv:1912.01389.

[35] F. Souza, R. Nogueira, R. Lotufo, Portuguese Named Entity Recognition using BERT-CRF (2019). URL: http://arxiv.org/abs/1909.10649. arXiv:1909.10649.

[36] C. Liang, Y. Yu, H. Jiang, S. Er, R. Wang, T. Zhao, C. Zhang, BOND: BERT-Assisted Open-Domain Named Entity Recognition with Distant Supervision, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2020) 1054–1064. doi:10.1145/3394486.3403149. arXiv:2006.15509.

[37] S. Zhou, J. Liu, X. Zhong, W. Zhao, Named Entity Recognition Using BERT with Whole World Masking in Cybersecurity Domain, 2021 IEEE 6th International Conference on Big Data Analytics, ICBDA 2021 (2021) 316–320. doi:10.1109/ICBDA51983.2021.9403180.

[38] P. Röttger, J. Pierrehumbert, P. Rottger, J. Pierrehumbert, Temporal Adaptation of {BERT} and Performance on Downstream Document Classification: Insights from Social Media, in: Findings of the Association for Computational Linguistics: EMNLP 2021, Association for Computational Linguistics, Punta Cana, Dominican Republic, 2021, pp. 2400–2412. URL: https://aclanthology.org/2021.findings-emnlp.206. doi:10.18653/v1/2021.findings-emnlp.206. arXiv:2104.08116.

[39] K. Vani, S. Mellace, A. Antonucci, Temporal embeddings and transformer models for narrative text understanding, CEUR Workshop Proceedings 2593 (2020) 71–77. arXiv:2003.08811.

[40] C. D. Manning, J. Bauer, J. Finkel, S. J. Bethard, The Stanford CoreNLP Natural Language Processing Toolkit, Aclweb.Org (2014) 55–60. URL: http://macopolo.cn/mkpl/products.asp.