Open Knowledge Extraction Challenge 2018

René Speck[1,2], Michael Röder[2,3], Felix Conrads[3], Hyndavi Rebba[4], Catherine Camilla Romiyo[4], Gurudevi Salakki[4], Rutuja Suryawanshi[4], Danish Ahmed[4], Nikit Srivastava[4], Mohit Mahajan[4], and Axel-Cyrille Ngonga Ngomo[2,3]

[1] Leipzig University, DICE Group, Augustusplatz 10, Leipzig, Germany, D-04109
{speck,conrads}@informatik.uni-leipzig.de
[2] Institute for Applied Informatics (InfAI) e.V., AKSW Group, Hainstraße 11, Leipzig, Germany, D-04109
[3] Paderborn University, DICE Group, Warburger Straße 100, Paderborn, Germany, D-33098
{michael.roeder,axel.ngonga}@uni-paderborn.de
[4] Paderborn University, Warburger Straße 100, Paderborn, Germany, D-33098
{dahmed,mac24079,hyndavi,rcathy,gss,nikit,rutuja}@mail.uni-paderborn.de

Abstract. The fourth edition of the Open Knowledge Extraction Challenge took place at the 15th Extended Semantic Web Conference in 2018. The aim of the challenge was to bring together researchers and practitioners from academia as well as industry to compete of pushing further the state of the art in knowledge extraction from text for the Semantic Web. This year, the challenge reused two tasks from the former challenge and defined two new tasks. Thus, the challenge consisted of tasks such as Named Entity Identification, Named Entity Disambiguation and Linking as well as Relation Extraction. To ensure an objective evaluation of the performance of participating systems, the challenge ran on a version the FAIR benchmarking platform Gerbil integrated in the HOBBIT platform. The performance was measured on manually curated gold standard datasets with Precision, Recall, F1-measure and the runtime of participating systems.

Keywords: Knowledge Extraction, Named Entity Identification, Named Entity Linking, Relation Extraction, Semantic Web

1 Introduction

The vision of the Semantic Web (SW) is an extension of the Document Web with the goal to allow intelligent agents to access, process, share and understand the data in the web. These agents are build upon structured data. Thus, implementing the vision of the SW requires transforming unstructured and semi-structured data from the Document Web into structured machine processable data of the SW using knowledge extraction approaches.

To push the state of the art in knowledge extraction from natural language text, the Open knowledge extraction Challenge (OKE) aims to trigger attention from the knowledge extraction community and foster their broader integration with the SW community. Therefore, the OKE has the ambition to provide a reference framework for research on knowledge extraction from text for the SW by defining a number of tasks (typically from information and knowledge extraction), taking into account specific SW requirements.

The first OKE 2015 [7] and second OKE 2016 [8] were both composed of two tasks, *Entity Recognition, Linking and Typing for Knowledge Base population* and *Class Induction and entity typing for Vocabulary and Knowledge Base enrichment*. In the first version, the challenge had four participants, Adel [9], CETUS [12], FRED [2] and OAK@Sheffield [4]. In the second version the challenge had five participants, a new version of Adel [10], Mannheim [3], WestLab-Task1 [1], WestLab-Task2 [5] and the baseline with CETUS from the former year. In the third version, the OKE 2017 [14] was composed of three tasks, *Focused Named Entity Identification and Linking*, *Broader Named Entity Identification and Linking* and *Focused Musical Named Entity Recognition and Linking*. In this version the challenge had two participants, a new version of Adel [11] and the baseline with FOX [13].

This year, the OKE 2018 reused the first two tasks from the former challenge, *Focused* and *Broader Named Entity Identification and Linking* as well as defined two new tasks, *Relation Extraction* and *Knowledge Extraction*.

The rest of this paper is structured as follows: We begin with defining preliminaries in Section 2 before describing the challenge tasks in Section 3. In Section 4 we give a brief introduction of the participating systems and compare the results achieved by our evaluation on the gold datasets in Section 5. Finally, we discuss the insights provided by the challenge and possible extensions in the last section.

2 Preliminaries and Notations

In this section we define terminologies and notations that are used throughout this paper.

Knowledge Base

Let a knowledge base K consists of a set of entities E_K, an entity type hierarchy T_K with a function that maps each entity to its types $\psi_K : E_K \to 2^{T_K}$, a relation type hierarchy R_K with a function that maps each relation to its domain and range entity types $\phi_K : R_K \to T_K \times T_K$ and relation instances or facts $F_K = \{r(e_1, e_2)\} \subset R_K \times E_K \times E_K$ with $r \in R_K$ and $(e_1, e_2) \in E_K \times E_K$.

Named Entity Identification

Consider each dataset D to be a set of documents and each document d to be sequence of words $d = (w_j)_{j=1,2,...}$ The identification of named entities in a given document d aims to find named entity mentions $M = \{m_i\}_{i=1,2,...}$ that express named entities. A named entity mention m is a sequence of words in d identified by its start and end index $I_M = \{(a, b)_i\}_{i=1}^{|M|}$ where $a, b \in \mathbb{N}$ and $a < b$.

Named Entity Disambiguation and Linking

The aim of named entity disambiguation and linking to a knowledge base K is to assign each named entity mention $m \in M$ to an entity in K if possible, otherwise to generate a new resource for such an emerging entity, i.e. $\varphi : M \to E_K \cup E_{\bar{K}}$ is a function that maps an entity mention to an entity in E_K or to a newly generated entity in $E_{\bar{K}}$ for an emerging entity that does not exist in K.

Closed Binary Relation Extraction

Closed binary relation extraction aims to find relations $r(e_j, e_k)$ expressed in a given text $d \in D$ with $r \in R_K$ and $e_j, e_k \in E_K \cup E_{\bar{K}}$. Often, closed binary relation extraction is limited to a subset of relations $R \subset R_K$ in K.

RDF/Turtle Prefixes

Listing 1.1 depicts the RDF/Turtle prefixes for all in- and output examples we illustrate in this paper.

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dbr: <http://dbpedia.org/resource/> .
@prefix dbo: <http://dbpedia.org/ontology/> .
@prefix aksw: <http://aksw.org/notInWiki/> .
@prefix oa: <http://www.w3.org/ns/oa#> .
```

Listing 1.1: Prefixes for the examples.

3 Open Knowledge Extraction Challenge Tasks

In this section, we describe each of the four challenge tasks and provide examples for a better understanding. All tasks depended on the DBpedia knowledge base. A participating system was not expected to process any preprocessing steps (e.g. pronoun resolution) on the input data. In case a resource for an entity was missing in the knowledge base, a system was expected to generate a URI using a namespace that does not match a known knowledge base (e.g. `http://aksw.org/notInWiki/`) for this emerging entity.

We carried out the evaluation with the HOBBIT benchmarking platform and the benchmark implementation of the HOBBIT project[5] which relies on the Gerbil evaluation framework [15].

[5] `http://project-hobbit.eu`

3.1 Task 1: Focused Named Entity Identification and Linking

The first task compromised a two-step process with a) the identification of named entity mentions in sentences and b) the disambiguation of these mentions by linking to resources in the given knowledge base. A competing system was expected to a) identify named entity mentions $\{m_i\}_{i=1,2,...}$ in a given document d with $m_i \in d$ by the start and end indices $\{(a,b)_i\}_{i=1,2,...}$. Further, b) to find the URIs in K to disambiguate and link each mention if possible. Otherwise, URIs should be generated the emerging entities and link these mentions, $\{\varphi(m_i)\}_{i=1,2,...}$.

This task was limited to a subset T of entity types[6] provided by the DBpedia knowledge base, i.e. $T := \{\text{dbo:Person}, \text{dbo:Place}, \text{dbo:Organisation}\}$.

Example Listing 1.2 is an example request document of task 1 and Listing 1.3 is the expected response document for the given request document.

```
<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-1/sentence-1#char=0,90>
  a nif:RFC5147String , nif:String , nif:Context ;
  nif:beginIndex "0"^^xsd:nonNegativeInteger ;
  nif:endIndex "90"^^xsd:nonNegativeInteger ;
  nif:isString "Leibniz was born in Leipzig in 1646 and attended the University of Leipzig
       from 1661-1666."@en .
```

Listing 1.2: Example request document in task one.

```
<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-1/sentence-1#char=0,7>
  a nif:RFC5147String , nif:String ;
  nif:anchorOf "Leibniz"@en ;
  nif:beginIndex "0"^^xsd:nonNegativeInteger ;
  nif:endIndex "7"^^xsd:nonNegativeInteger ;
  nif:referenceContext <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-1/
       sentence-1#char=0,90> ;
  itsrdf:taIdentRef dbr:Gottfried_Wilhelm_Leibniz .

<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-1/sentence-1#char=20,27>
  a nif:RFC5147String , nif:String ;
  nif:anchorOf "Leipzig"@en ;
  nif:beginIndex "20"^^xsd:nonNegativeInteger ;
  nif:endIndex "27"^^xsd:nonNegativeInteger ;
  nif:referenceContext <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-1/
       sentence-1#char=0,90> ;
  itsrdf:taIdentRef dbr:Leipzig .

<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-1/sentence-1#char=53,74>
  a nif:RFC5147String , nif:String ;
  nif:anchorOf "University of Leipzig"@en ;
  nif:beginIndex "53"^^xsd:nonNegativeInteger ;
  nif:endIndex "74"^^xsd:nonNegativeInteger ;
  nif:referenceContext <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-1/
       sentence-1#char=0,90> ;
  itsrdf:taIdentRef dbr:Leipzig_University .
```

Listing 1.3: Example of the expected response document in task one.

[6] http://mappings.dbpedia.org/server/ontology/classes (full type hierarchy).

3.2 Task 2: Broader Named Entity Identification and Linking

This task extended the former task towards a broader set of entity types. Beside the three types of the first task, a competing system had to identify other types of entities and to link these entities as well.

In the first column in Table 1, a complete list of types that are considered in this task is provided. The middle column contains example subtypes of the corresponding type if any such type is available and the last column contains example instances in the knowledge base for the types.

Table 1: Types, subtype examples and instance examples. All types are defined in the `dbo` namespace.

Type	Subtypes	Instances
`Activity`	`Game`, `Sport`	`Baseball`,`Chess`
`Agent`	`Organisation`, `Person`	`Leipzig_University`
`Award`	`Decoration`, `NobelPrize`	`Humanitas_Prize`
`Disease`		`Diabetes_mellitus`
`EthnicGroup`		`Javanese_people`
`Event`	`Competition`, `PersonalEvent`	`Battle_of_Leipzig`
`Language`	`ProgrammingLanguage`	`English_language`
`MeanOfTransportation`	`Aircraft`, `Train`	`Airbus_A300`
`PersonFunction`	`PoliticalFunction`	`PoliticalFunction`
`Place`	`Monument`, `WineRegion`	`Beaujolais`, `Leipzig`
`Species`	`Animal`, `Bacteria`	`Cat`, `Cucumibacter`
`Work`	`Artwork`, `Film`	`Actrius`, `Debian`

3.3 Task 3: Relation Extraction

Given a dataset D with documents, the DBpedia knowledge base K, a target entity type hierarchy T with $T \subset T_K$ and a target relation type hierarchy R with $R \subset R_K$. Furthermore, annotations of the documents are given, i.e., entity mentions M with the positions I_M, the disambiguation $\varphi : M \rightarrow E_K \cup E_{\bar{K}}$ and the types of entities $\psi_K : E_K \cup E_{\bar{K}} \rightarrow T$.

The aim of this task was to find binary relations $r(e_j, e_k)$ with $r \in R$ and $e_j, e_k \in E_K \cup E_{\bar{K}}$. The domain and range entity types for the applied relations in this task, $\phi_K : R \rightarrow T \times T$ were given and are depicted in Table 2.

For the preparation of an output document, a participating system had to serialize the found binary relations in the input document with RDF statements using the the Web Annotation Vocabulary to connect the extracted statement with the given document.

Table 2: Relation type hierarchy with domain and range entity types. All relations and types are defined in the `dbo` namespace.

Relation	Domain	Range
affiliation	Organisation	Organisation
almaMater	Person	EducationalInstitution
bandMember	Band	Person
birthPlace	Person	Place
ceo	Organisation	Person
child	Person	Person
club	Athlete	SportsTeam
country	Organisation, Person, Place	Country
deathPlace	Person	Place
debutTeam	Athlete	SportsTeam
department	PopulatedPlace	PopulatedPlace
district	Place	PopulatedPlace
doctoralAdvisor	Scientist	Person
doctoralStudent	Scientist	Person
employer	Person	Organisation
formerBandMember	Band	Person
formerTeam	Athlete	SportsTeam
foundationPlace	Organisation	City
headquarter	Organisation	PopulatedPlace
hometown	Organisation, Person	Settlement
leaderName	PopulatedPlace	Person
locatedInArea	Place	Place
location	Organisation, Person, Place	Place
nationality	Person	Country
parent	Person	Person
president	Organisation	Person
relative	Person	Person
spouse	Person	Person
subsidiary	Company	Company
tenant	ArchitecturalStructure	Organisation
trainer	Athlete	Person

Example Listing 1.4 is an example request document of this task and Listing 1.5 shows the triples that have to be added to the request to form the expected response document.

```
<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-3/sentence-1#char=0,78>
  a nif:RFC5147String , nif:String , nif:Context ;
  nif:beginIndex "0"^^xsd:nonNegativeInteger ;
  nif:endIndex "78"^^xsd:nonNegativeInteger ;
  nif:isString "Conor McGregor's longtime trainer, John Kavanagh, is ready to shock the world
       ."^^xsd:string .

<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-3/sentence-1#char=0,22>
  a nif:RFC5147String , nif:String , nif:Phrase ;
  nif:anchorOf "Conor McGregor's"^^xsd:string ;
  nif:beginIndex "0"^^xsd:nonNegativeInteger ;
  nif:endIndex "22"^^xsd:nonNegativeInteger ;
  nif:referenceContext <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-3/
       sentence-1#char=0,78> ;
  its:taClassRef dbo:Person ;
  itsrdf:taIdentRef dbr:Conor_McGregor .

<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-3/sentence-1#char=35,48>
  a nif:RFC5147String , nif:String , nif:Phrase ;
  nif:anchorOf "John Kavanagh"^^xsd:string ;
  nif:beginIndex "35"^^xsd:nonNegativeInteger ;
  nif:endIndex "48"^^xsd:nonNegativeInteger ;
  nif:referenceContext <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-3/
       sentence-1#char=0,78> ;
  its:taClassRef dbo:Person ;
  itsrdf:taIdentRef aksw:John_Kavanagh .
```

Listing 1.4: Example request document in task three.

```
[]
  a rdf:Statement , oa:Annotation ;
  rdf:object dbr:Conor_McGregor ;
  rdf:predicate dbo:trainer ;
  rdf:subject aksw:John_Kavanagh ;
  oa:hasTarget [
    a oa:SpecificResource;
    oa:hasSource <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-3/
      sentence-1#char=0,78> ] .
```

Listing 1.5: Example of the expected response document in task three.

3.4 Task 4: Knowledge Extraction

This task was a combination of task one *Focused Named Entity Identification and Linking* and task three *Relation Extraction*.

In this task, the input documents comprised only natural language text, similar to the input documents of task one. Thus, without annotations of entity mentions, entity types and linkings to the knowledge base. A participating system was expected to provide combined serializations with the information analogous to task one and task three.

Example Listing 1.6 is an example request document of this task and Listing 1.7 shows the triples that have to be added to the request to form the expected response document.

```
<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-4/sentence-1#char=0,78>
  a nif:RFC5147String , nif:String , nif:Context ;
  nif:beginIndex "0"^^xsd:nonNegativeInteger ;
  nif:endIndex "78"^^xsd:nonNegativeInteger ;
  nif:isString "Conor McGregor's longtime trainer, John Kavanagh, is ready to shock the world
       ."^^xsd:string .
```

Listing 1.6: Example request document in task four.

```
<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-4/sentence-1#char=0,22>
  a nif:RFC5147String , nif:String , nif:Phrase ;
  nif:anchorOf "Conor McGregor's"^^xsd:string ;
  nif:beginIndex "0"^^xsd:nonNegativeInteger ;
  nif:endIndex "22"^^xsd:nonNegativeInteger ;
  nif:referenceContext <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-4/
       sentence-1#char=0,78> ;
  itsrdf:taIdentRef dbr:Conor_McGregor .

<http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-4/sentence-1#char=35,48>
  a nif:RFC5147String , nif:String , nif:Phrase ;
  nif:anchorOf "John Kavanagh"^^xsd:string ;
  nif:beginIndex "35"^^xsd:nonNegativeInteger ;
  nif:endIndex "48"^^xsd:nonNegativeInteger ;
  nif:referenceContext <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-4/
       sentence-1#char=0,78> ;
  itsrdf:taIdentRef aksw:John_Kavanagh .

[]
  a rdf:Statement , oa:Annotation ;
  rdf:object dbr:Conor_McGregor ;
  rdf:predicate dbo:trainer ;
  rdf:subject aksw:John_Kavanagh ;
  oa:hasTarget [
    a oa:SpecificResource;
    oa:hasSource <http://www.ontologydesignpatterns.org/data/oke-challenge-2018/task-4/
      sentence-1#char=0,78> ] .
```

Listing 1.7: Example of the expected response document in task four.

4 Participants

The challenge attracted five research groups this year. Four groups from universities and one from industry. Two groups participated in the challenge and competed in task three. Both systems, RelExt and the Baseline, are briefly described in the next subsections.

4.1 RelExt

RelExt is an approach based on a deep learning classifier that uses self attention. The classifier was trained on sentences from Wikipedia pages chosen in a distance supervised fashion with the DBpedia knowledge base. RelExt uses a filtering step to find words in sentences that might express specific relations. These words are manually filtered by the authors and were used to refine the sentences to obtain training data.

RelExt participated in task three of the OKE challenge.

4.2 Baseline

The baseline system for task three simply used the annotated documents in the evaluation phase without a learning or training step on the training dataset. The input documents of task three consisted of annotated entities with entity linkings to the DBpedia knowledge base. Thus, the baseline chose pairwise the given URIs of the linked entities from the input documents to create a SPARQL query to request all predicates that hold between two URIs in DBpedia. In case two entities had a statement in the knowledge base with a predicate included in the task, the baseline chose this statement in the response document.

For instance, let "Leibniz was born in Leipzig." be a sentence in a document together with the entity linkings `dbr:Gottfried_Wilhelm_Leibniz` for the entity mention "Leibniz" and `dbr:Leipzig` for the entity mention "Leipzig" in this example sentence. The baseline took both resource URIs and queried the DBpedia to find all predicates that hold between these resources. In this case, the predicate `dbo:birthPlace` is in DBpedia as well as in the tasks list of predicates and thus chosen to be in the response document.

5 Evaluation

In this section we describe the evaluation of the participating systems. We define the evaluation metrics, describe the datasets and present the evaluation results.

5.1 Platform

The benchmarking platform for named entity recognition and linking implemented within HOBBIT [6][7] reuses some of the concepts developed within the open source project Gerbil. These concepts were migrated and adapted to the HOBBIT architecture. The platform calculates the micro and macro averages of Precision, Recall and F1-measure. Additionally, the time a system needs to answer a request as well as the number of failing requests (i.e. requests that are answered with an error code instead of a valid response) are determined.

5.2 Measures

Equation 1, 2 and 3 formalize Precision p_d, Recall r_d and F1-measure used to evaluate the quality of the systems responses for each document $d \in D$. They consist of the number of true positives TP_d, false positives FP_d and false negatives FN_d. We micro and macro average the performances over the documents.[8]

$$p_d = \frac{TP_d}{TP_d + FP_d} \tag{1}$$

[7] `http://project-hobbit.eu/wp-content/uploads/2017/04/D2.2.1.pdf`

[8] The macro averages for the performance measures can be retrieved from the official HOBBIT SPARQL endpoint at `http://db.project-hobbit.eu/sparql`.

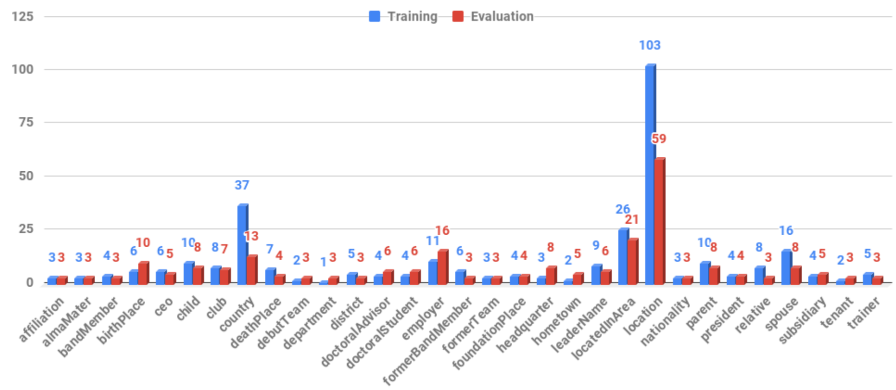

Fig. 1: Predicate distribution of the training and evaluation datasets for challenge task three and four.

$$r_d = \frac{TP_d}{TP_d + FN_d} \tag{2}$$

$$f_d = 2 \cdot \frac{p_d \cdot r_d}{p_d + r_d} \tag{3}$$

5.3 Datasets

The datasets for all tasks were manually curated. For the first two tasks, we reused cleansed and updated versions of the datasets from the former year, i.e. fixed typos and wrong indices. For the two new tasks, we created new datasets with eight annotators. Each annotator reviewed the annotations of another annotator to reduce noise and cognitive bias as well as to ensure a high quality.

In Figure 1, the predicate distributions of the training and evaluation datasets for the challenge task three and four are shown. The training dataset consisted of 319 and the evaluation dataset of 239 annotated relations in the documents (cf. Table 3). In Table 3 the number of annotated documents, entities, entity linkings and relations are shown for the challenge datasets. The datasets for the first two tasks were without relation annotations and the datasets for the last two tasks were the same but differ as input documents for participating systems, cf. the examples in Listings 1.4 and 1.6. For instance, the training datasets for task one consisted of 60 documents with 377 annotated entity mentions as well as linkings of these mentions to 255 unique resources in the DBpedia knowledge base or newly generated resources for emerging entities.

The training datasets for each task are available at the challenge website.[9]

[9] `https://project-hobbit.eu/challenges/oke2018-challenge-eswc-2018`

Table 3: Attributes and their values of the challenge datasets.

	Task 1	Task 2	Task 3 and Task 4
	Training/Evaluation	Training/Evaluation	Training/Evaluation
Documents	60/58	56/53	100/100
Entities	377/381	422/462	308/274
Linkings	255/255	301/327	254/242
Relations	N/A	N/A	319/239

5.4 Results

Table 4 depicts the results of the OKE 2018 on task three. The results are available in Hobbit for both participants, RelExt[10] and the Baseline[11]. RelExt won this task with 54.30% Macro F1-Score and 48.01% Micro F1-Score.

Table 4: RelExt and Baseline.

KPI	RelExt	Baseline
Avg. ms per Doc	836.26	513.42
Error Count	1	0
Macro F1-Score	**54.30**	8.00
Macro Precision	53.98	10.00
Macro Recall	64.17	7.18
Micro F1-Score	**48.01**	8.66
Micro Precision	39.62	68.75
Micro Recall	60.92	4.62

6 Conclusion

The OKE attracted five research groups from academia and industry coming from the knowledge extraction as well as the SW communities. Indeed, the challenge proposal was aimed at attracting groups from these two communities in order to further investigate existing overlaps between both. Additionally, one of the goals of the challenge was to foster the collaboration between the two communities, to the aim of growing further the SW community. To achieve this goal we defined a SW reference evaluation framework, which is composed of a) four tasks, b) a training and evaluation dataset for each task, and c) an evaluation framework to measure the performance of the systems. Although the participation in terms of number of competing systems remained quite limited with two,

[10] `https://master.project-hobbit.eu/experiments/1529075533385`
[11] `https://master.project-hobbit.eu/experiments/1527777181515`

we believe that the challenge is a success in the hybridisation of Semantic Web technologies with knowledge extraction methods.

As a matter of fact, the evaluation framework is available online and can be reused by the community and for next editions of the challenge.

Acknowledgements

This work has been supported by the H2020 project HOBBIT (no. 688227), the BMWI projects GEISER (no. 01MD16014E) and OPAL (no. 19F2028A), the EuroStars projects DIESEL (no. 01QE1512C) and QAMEL (no. 01QE1549C).

References

1. Chabchoub, M., Gagnon, M., Zouaq, A.: Collective disambiguation and semantic annotation for entity linking and typing. In: Sack, H., Dietze, S., Tordai, A., Lange, C. (eds.) Semantic Web Challenges. pp. 33–47. Springer International Publishing, Cham (2016)
2. Consoli, S., Recupero, D.R.: Using fred for named entity resolution, linking and typing for knowledge base population. In: Gandon, F., Cabrio, E., Stankovic, M., Zimmermann, A. (eds.) Semantic Web Evaluation Challenges. pp. 40–50. Springer International Publishing, Cham (2015)
3. Faralli, S., Ponzetto, S.P.: Dws at the 2016 open knowledge extraction challenge: A hearst-like pattern-based approach to hypernym extraction and class induction. In: Sack, H., Dietze, S., Tordai, A., Lange, C. (eds.) Semantic Web Challenges. pp. 48–60. Springer International Publishing, Cham (2016)
4. Gao, J., Mazumdar, S.: Exploiting linked open data to uncover entity types. In: Gandon, F., Cabrio, E., Stankovic, M., Zimmermann, A. (eds.) Semantic Web Evaluation Challenges. pp. 51–62. Springer International Publishing, Cham (2015)
5. Haidar-Ahmad, L., Font, L., Zouaq, A., Gagnon, M.: Entity typing and linking using sparql patterns and dbpedia. In: Sack, H., Dietze, S., Tordai, A., Lange, C. (eds.) Semantic Web Challenges. pp. 61–75. Springer International Publishing, Cham (2016)
6. Ngonga Ngomo, A.C., Röder, M.: HOBBIT: Holistic benchmarking for big linked data. In: ESWC, EU networking session (2016),
7. Nuzzolese, A.G., Gentile, A.L., Presutti, V., Gangemi, A., Garigliotti, D., Navigli, R.: Open knowledge extraction challenge. In: Gandon, F., Cabrio, E., Stankovic, M., Zimmermann, A. (eds.) Semantic Web Evaluation Challenges. pp. 3–15. Springer International Publishing, Cham (2015)
8. Nuzzolese, A.G., Gentile, A.L., Presutti, V., Gangemi, A., Meusel, R., Paulheim, H.: The second open knowledge extraction challenge. In: Sack, H., Dietze, S., Tordai, A., Lange, C. (eds.) Semantic Web Challenges. pp. 3–16. Springer International Publishing, Cham (2016)
9. Plu, J., Rizzo, G., Troncy, R.: A hybrid approach for entity recognition and linking. In: Gandon, F., Cabrio, E., Stankovic, M., Zimmermann, A. (eds.) Semantic Web Evaluation Challenges. pp. 28–39. Springer International Publishing, Cham (2015)
10. Plu, J., Rizzo, G., Troncy, R.: Enhancing entity linking by combining ner models. In: Sack, H., Dietze, S., Tordai, A., Lange, C. (eds.) Semantic Web Challenges. pp. 17–32. Springer International Publishing, Cham (2016)

11. Plu, J., Troncy, R., Rizzo, G.: Adel@oke 2017: A generic method for indexing knowledge bases for entity linking. In: Dragoni, M., Solanki, M., Blomqvist, E. (eds.) Semantic Web Challenges. pp. 49–55. Springer International Publishing, Cham (2017)
12. Röder, M., Usbeck, R., Speck, R., Ngonga Ngomo, A.C.: Cetus – a baseline approach to type extraction. In: 1st Open Knowledge Extraction Challenge @ 12th European Semantic Web Conference (ESWC 2015) (2015),
13. Speck, R., Ngonga Ngomo, A.C.: Ensemble learning for named entity recognition. In: The Semantic Web – ISWC 2014, Lecture Notes in Computer Science, vol. 8796, pp. 519–534. Springer International Publishing (2014),
14. Speck, R., Röder, M., Oramas, S., Espinosa-Anke, L., Ngonga Ngomo, A.C.: Open knowledge extraction challenge 2017. In: Semantic Web Evaluation Challenge. pp. 35–48. Springer International Publishing (2017),
15. Usbeck, R., Röder, M., Ngonga Ngomo, A.C., Baron, C., Both, A., Brümmer, M., Ceccarelli, D., Cornolti, M., Cherix, D., Eickmann, B., Ferragina, P., Lemke, C., Moro, A., Navigli, R., Piccinno, F., Rizzo, G., Sack, H., Speck, R., Troncy, R., Waitelonis, J., Wesemann, L.: GERBIL – general entity annotation benchmark framework. In: 24th WWW conference (2015),