Open Knowledge Extraction Challenge

Andrea Giovanni Nuzzolese[1], Anna Lisa Gentile[2], Valentina Presutti[1],
Aldo Gangemi[1], Darío Garigliotti[3], and Roberto Navigli[3]

[1] Semantic Technology Lab, ISTC-CNR. Italy
[2] Department of Computer Science, University of Sheffield, UK
[3] Department of Computer Science, Sapienza University of Rome, Italy

Abstract. The Open Knowledge Extraction (OKE) challenge is aimed at promoting research in the automatic extraction of structured content from textual data and its representation and publication as Linked Data. We designed two extraction tasks: (1) *Entity Recognition, Linking and Typing* and (2) *Class Induction and entity typing*. The challenge saw the participations of four systems: CETUS-FOX and FRED participating to both tasks, Adel participating to Task 1 and OAK@Sheffield participating to Task 2. In this paper we describe the OKE challenge, the tasks, the datasets used for training and evaluating the systems, the evaluation method, and obtained results.

1 Introduction

The vision of the Semantic Web (SW) is to populate the Web with machine understandable data so as to make intelligent agents able to automatically interpret its content - just like humans do by inspecting Web content - and assist users in performing a significant number of tasks, relieving them of cognitive overload. The Linked Data movement [1] kicked-off the vision by realising a key bootstrap in publishing machine understandable information mainly taken from structured data (typically databases) or semi-structured data (e.g. Wikipedia infoboxes). However, most of the Web content consists of natural language text, e.g., Web sites, news, blogs, micro-posts, etc., hence a main challenge is to extract as much relevant knowledge as possible from this content, and publish it in the form of Semantic Web triples.

There is huge work on knowledge extraction (KE) and knowledge discovery contributing to address this problem, and several contests addressing the evaluation of Information Extraction systems. Hereafter we shortly list some of the most popular initiatives which have contributed to the advancement of research on automatic content extraction:

MUC-6 The Message Understanding Conferences is a series of conferences designed to evaluate research in information extraction. MUC-6 [7] was the first to define the "named entity" task, where the participants had to identify the names of all the people, organizations, and geographic locations in a collection of textual documents in English.

HUB-4 The Hub-4 Broadcast News Evaluation[4] included a MUC-style evaluation for Named Entity Recognition, but with the focus on speech input in the domain of broadcast news.

MUC-7 and MET-2 The main difference between MUC-7/MET-2[5] to previous MUC is the introduction of multilingual NE evaluation, using training and test articles from comparable domains for all languages.

CONLL The CoNLL-2002 and CoNLL-2003 shared task focused on language independent named entity recognition. The evaluation focused on entities of four types: persons (PER), organizations (ORG), locations (LOC) and miscellaneous (MISC) and the task was performed on Spanish and Dutch for CoNLL-2002 and German and English for CoNLL-2003 [12,13].

ACE The Automatic Content Extraction program evaluates methods to extract (i) entities, (iii) relations among these entities and (iii) the events in which these entities participate. In the first edition extraction tasks were available in English, Arabic and Chinese . In the entity detection and tracking (EDT) task, all mentions of an entity, whether a name, a description, or a pronoun, are to be found. ACE defines seven types of entities: Person, Organization, Location, Facility, Weapon, Vehicle and Geo-Political Entity (GPEs). Each type is further divided into subtypes (for instance, Organization subtypes include Government, Commercial, Educational, Non-profit, Other) [4]. ACE started in 2004 with following successful editions[6].

TAC The Text Analysis Conference[7] is a series of evaluation workshops on Natural Language Processing, with several specific tasks (known as "tracks"). The Knowledge Base Population (KBP) task[8] is present since 2009 and has the goal to populate knowledge bases (KBs) from unstructured text. The current KB schema consists of named entities that can be a person (PER), organization (ORG), or geopolitical entity (GPE) and predefined attributes (or slots) to fill for those named entities.

TREC-KBA The Knowledge Base Acceleration (KBA) track[9] ran in TREC 2012, 2013 and 2014. It evaluates systems that filter a time-ordered corpus for documents and slot fills that would change an entity profile in a predefined list of entities. The focus is therefore on spotting novelty and changes for predefined entities.

SemEval-2015 Task 13 The Multilingual All-Words Sense Disambiguation (WSD) and Entity Linking (EL) are tasks that address the lexical ambiguity of language, but they use different meaning inventories: EL uses encyclopedic knowledge, while WSD uses lexicographic information. The main goal of this

---

[4] `http://www.itl.nist.gov/iad/mig/publications/proceedings/darpa99/html/ie5/ie5.htm`

[5] `http://www.itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/muc_7_proceedings/overview.html`

[6] `https://www.ldc.upenn.edu/collaborations/past-projects/ace/annotation-tasks-and-specifications`

[7] `http://www.nist.gov/tac/tracks/index.html`

[8] `http://www.nist.gov/tac/2015/KBP`

[9] `http://trec-kba.org/`

combined task is to treat the two problems holistically using a resource that integrates both kinds of inventories (i.e., BabelNet 2.5.1).

Despite the numerous initiatives for benchmarking KE systems, there is lack of a "genuine" SW reference evaluation framework for helping researchers and the whole community to assess the state of the art in this domain. In fact, results of Knowledge Extraction systems are usually evaluated against tasks that do not focus on specific Semantic Web goals. For example, tasks such as named Entity Recognition, Relation Extraction, Frame Detection, etc. are certainly of importance for the SW, but in most cases such tasks are designed without considering the output design and formalisation in the form of Linked Data and OWL ontologies. This makes results of existing methods often not directly reusable for populating the SW, until a translation from linguistic semantics to formal semantics is performed.

The OKE challenge, inspired by [9], has the ambition to provide a reference framework for research on *Knowledge Extraction from text for the Semantic Web* by re-defining a number of tasks (typically from information and knowledge extraction) by taking into account specific SW requirements.

2 Tasks

The OKE challenge defines two tasks. This section provides their detailed description.

2.1 Task 1: Entity Recognition, Linking and Typing for Knowledge Base population

This task consists of (i) identifying Entities in a sentence and create an OWL individual (`owl:Individual` statement) representing it, (ii) link (`owl:sameAs` statement) such individual, when possible, to a reference Knowledge Base (i.e, DBpedia [2]) and (iii) assigning a type to such individual (`rdf:type` statement) selected from a set of given types. In this task by Entity we mean any discourse referent (the actors and objects around which a story unfolds), either named or anonymous that is an individual of one of the following DOLCE Ultra Lite classes[10] [5], i.e., `dul:Person`[11], `dul:Place`, `dul:Organization`, and `dul:Role`. By entities we also refer to anaphorically related discourse referents. Hence, anaphora resolution is part of the requirements for the identification of entities. As an example, for the sentence:

> *Florence May Harding studied at a school in Sydney, and with Douglas Robert Dundas , but in effect had no formal training in either botany or art.*

we want to recognize the entities reported in Table 1.

[10] http://stlab.istc.cnr.it/stlab/WikipediaOntology/

[11] The prefix `dul:` stands for the namespace `http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#`.

Table 1. Task 1: example.

Recognized Entity	Generated URI	Type	SameAs
Florence May Harding	oke:Florence_May_Harding	dul:Person	dbpedia:Florence_May_Harding
school	oke:School	dul:Organization	
Sydney	oke:Sydney	dul:Place	dbpedia:Sydney
Douglas Robert Dundas	oke:Douglas_Robert_Dundas	dul:Person	

Sentences were provided in input to systems as RDF by using the NIF notation[12] [8]. The following is an example of input for the previous sentence.

```
oke:task-1/sentence-1#char=0,146
        a                     nif:RFC5147String , nif:String , nif:Context ;
        nif:beginIndex        "0"^^xsd:nonNegativeInteger ;
        nif:endIndex          "146"^^xsd:nonNegativeInteger ;
        nif:isString          "Florence May Harding studied at a school in Sydney, and with
                              Douglas Robert Dundas , but in effect had no
                              formal training in either botany or art."@en .
```

System were asked to provide recognised entities by using a NIF-compliant output as shown in the following example.

```
...
oke:Florence_May_Harding
     a                        owl:Individual, dul:Person ;
     rdfs:label               "Florence May Harding"@en ;
     owl:sameAs               dbpedia:Florence_May_Harding .

oke:task-1/sentence-1#char=0,20
        a                     nif:RFC5147String , nif:String ;
        nif:anchorOf          "Florence May Harding"@en ;
        nif:beginIndex        "0"^^xsd:nonNegativeInteger ;
        nif:endIndex          "20"^^xsd:nonNegativeInteger ;
        nif:referenceContext  oke:task-1/sentence-1#char=0,146 ;
        itsrdf:taIdentRef     oke:Florence_May_Harding .
```

The RDF above[13] is an example of possible output for annotating the string that represents the entity *Florence May Harding* in the original sentence. This string is typed as a `nif:RFC5147String` and is related to a reference context (cf., property `nif:referenceContext`), which identifies the input sentence, and to an `owl:Individual` (cf., property `itsrdf:taIdentRef`), which represents the entity within the dataset. This entity is further typed as `dul:Person` and linked to its corresponding entity in DBpedia (cf. property `owl:sameAs`). The namespace prefix `oke:` is used to identify the URIs of recognised entities. There is not a given rule for generating these URI, thus any system can implement its own algorithm for generating URIs. The linking to DBpedia can be omitted in case a system is not able to identify a corresponding entity in such a dataset. This means that it might be possible to have entities that cannot be linked to any DBpedia entities. This is always the case occurring when dealing with anonymous entities. For example, given the sentence:

[12] http://persistence.uni-leipzig.org/nlp2rdf/

[13] The prefixes `nif:`, `itsrdf:`, `dul:`, and `dbpedia:` identify the namespaces `http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#`, `http://www.w3.org/2005/11/its/rdf#`, `http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#`, and `http://dbpedia.org/resource/` respectively.

She was appointed as Senator for Life in Italy by the President Carlo Azeglio Ciampi.

We want to recognise the term *She* an as anonymous entity within our dataset and to type it as `owl:Individual` and `dul:Person`. However, we do not want any linking to DBpedia because it would introduce an error.

2.2 Task 2: Class Induction and entity typing for Vocabulary and Knowledge Base enrichment.

This task was designed for producing `rdf:type` statements for an entity, given its definition as natural language text. The participants were provided with a dataset of sentences, each defining an entity (known a priori). More in detail the task required the participants to (i) identify the type(s) of the given entity as they are expressed in the given definition, (ii) create a `owl:Class` statement for defining each of them as a new class in the target knowledge base, (iii) create a `rdf:type` statement between the given entity and the new created classes, and (iv) align the identified types, if a correct alignment is available, to a set of given types from a subset of DOLCE+DnS Ultra Lite classes. Table 2 shows the complete list of these types[14]

Table 2. The subset of DOLCE+DnS Ultra Lite classes used for typing entities in the Task 2.

Class	Description
dul:Abstract	Anything that cannot be located in space-time.
d0:Activity	Any action or task planned or executed by an agent intentionally causing and participating in it.
dul:Amount	Any quantity, independently from how it is measured, computed, etc.
d0:Characteristic	An aspect or quality of a thing.
dul:Collection	A container or group of things (or agents) that share one or more common properties.
d0:CognitiveEntity hline dul:Description	Attitudes, cognitive abilities, ideologies, psychological phenomena, mind, etc. A descriptive context that creates a relational view on a set of data or observations.
d0:Event	Any natural event, independently of its possible causes.
dul:Goal	The description of a situation that is desired by an agent.
dul:InformationEntity	A piece of information, be it concretely realized or not: linguistic expressions, works of art, knowledge objects.
d0:Location	A location, in a very generic sense e.g. geo-political entities, or physical object that are inherently located.
dul:Organism	A physical object with biological characteristics, typically able to self-reproduce.
dul:Organization	An internally structured, conventionally created social entity such as enterprises, bands, political parties, etc.
dul:Person	Persons in commonsense intuition.
dul:Personification	A social entity with agentive features, invented or conceived through a cultural process.
dul:PhysicalObject	Any object that has a proper space region, and an associated mass: natural bodies, artifacts, substances.
dul:Process	Any natural process, independently of its possible causes.
dul:Process	Any natural process, independently of its possible causes.
dul:Role	A concept that classifies some entity: social positions, roles, statuses.
dul:Situation hline d0:System	A unified view on a set of entities, e.g. physical or social facts or conditions, configurations, etc. Physical, social, political systems.
dul:TimeInterval	A time span.
d0:Topic	Any area, discipline, subject of knowledge.

For example, given the entity `dbpedia:Skara_Cathedral` and its definition

[14] Prefixes `d0:` and `dul:` stand for namespaces `http://ontologydesignpatterns.org/ont/wikipedia/d0.owl#` and `http://www.ontologydesignpatterns.org/ont/dul/DUL.owl#` respectively.

Skara Cathedral is a church in the Swedish city of Skara.

the types that the systems were asked to recognise are reported in Table 3 .

Table 3. Task 2: example.

Recognized string for the type	**Generated Type**	**Subclass of**
fictional villain	oke:FictionalVillain	dul:Personification
villain	oke:Villain	oke:FictionalVillain, dul:Person

Target entities, i.e. the entities to type, along with their definition in natural language were provided as RDF by using the NIF notation. The following is an example of input for the previous example.

```
oke:task-2/sentence-1#char=0,150>
     a                    nif:RFC5147String , nif:String , nif:Context ;
     nif:isString         "Brian Banner is a fictional villain from the Marvel Comics Universe
                          created by Bill Mantlo and Mike Mignola and first
                          appearing in print in late 1985." ;
     nif:beginIndex       "0"^^xsd:int ;
     nif:endIndex         "150"^^xsd:int .

oke:task-2/sentence-1#char=0,12
     a                    nif:RFC5147String , nif:String ;
     nif:anchorOf         "Brian Banner"@en ;
     nif:referenceContext  oke:task-2/sentence-1#char=0,150 ;
     nif:beginIndex       "0"^^xsd:int ;
     nif:endIndex         "12"^^xsd:int ;
     itsrdf:taIdentRef    dbpedia:Brian_Banner .

dbpedia:Brian_Banner
     rdfs:label          "Brian Banner"@en .
```

Participants were asked to complete the RDF snippet above with the following information about typing by using the NIF notation:

```
...
oke:FictionalVillain
     a                    owl:Class ;
     rdfs:label           "fictional villain"@en ;
     rdfs:subClassOf      dul:Personification .

oke:Villain
     a                    owl:Class ;
     rdfs:label           "villain"@en ;
     rdfs:subClassOf      oke:FictionalVillain, dul:Person .

oke:sentence-1#char=18,35
     a                    nif:RFC5147String , nif:String ;
     nif:anchorOf         "fictional villain"@en ;
     nif:referenceContext  oke:task-2/sentence-1#char=0,150 ;
     nif:beginIndex       "18"^^xsd:int ;
     nif:endIndex         "35"^^xsd:int ;
     itsrdf:taIdentRef    oke:FictionalVillain .

oke:sentence-1#char=28,35>
     a                    nif:RFC5147String , nif:String ;
     nif:anchorOf         "villain"@en ;
     nif:referenceContext  oke:task-2/sentence-1#char=0,150 ;
```

```
nif:beginIndex        "28"^^xsd:int ;
nif:endIndex          "35"^^xsd:int ;
itsrdf:taIdentRef     oke:Villain .
```

We designed the task in order to ask participants to report as `rdfs:label` the string recognised within a definition as a valid type for a given entity. Additionally, we asked participants to record span indexes for such strings with respect to the original definition by using `nif:beginIndex` and `nif:endIndex`. Namely, `nif:beginIndex` and `nif:endIndex` were used to identify the initial and final span index respectively.

3 Training and evaluation datasets

We built two separate datasets for each task in order to distinguish between a (i) a dataset to be used for training purposes and (ii) another one to use for evaluating the systems. In next sections we describe how we built such datasets for each task and we provide details about them.

3.1 Task 1

The training and the evaluation datasets for Task 1 were built by manually annotating a set of 196 sentences. These sentences were selected from Wikipedia articles reporting biographies of scholars. This choice comes from the observation that biographies about scholars typically contain entities about people (e.g., the scholar that is subject of the given Wikipedia article, her collegues, her relatives, etc.), locations (e.g., the places the scholar lived in), organisations (e.g., the universities the scholar worked for) and roles (e.g., the academic roles held by the scholar during her career). Hence, we split the 196 into the training and the evaluation datasets by taking care to have:

- no overlap of sentences between the two datasets;
- a comparable number of sentences in both datasets;
- as much as possible an equal distribution of DOLCE entity types (i.e., Person, Place, Organization, and Role) within the datasets.

Therefore, the training dataset for Task 1 is composed of 95 sentences while the evaluation dataset for the same task counts 101 sentences. Table 4 shows the details about the two datasets in terms of the number of sentences, the overall number of annotated entities, the average number of annotated entities per sentence, and the number of entities linked to DBpedia.

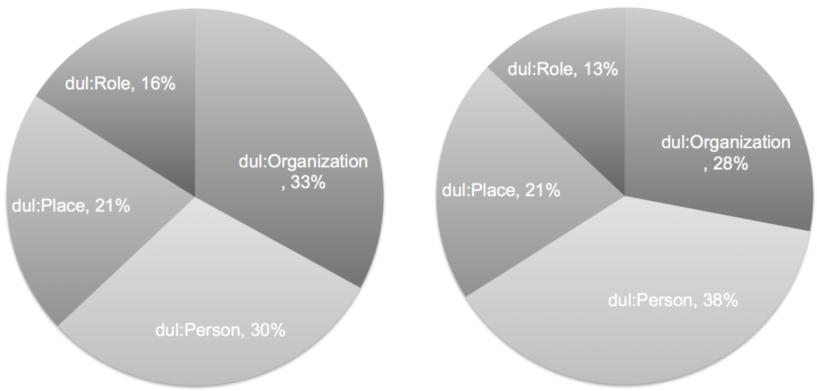
Figure 1 shows the distribution of entities with respect to the four DOLCE types used for typing in Task 1. More in detail, Figure 1(a) and Figure 1(b) show the distribution in the training dataset and the evaluation dataset respectively.

Both datasets are available on-line as TURTLE for download[15].

[15] The training dataset is available at `https://github.com/anuzzolese/oke-challenge/blob/master/GoldStandard_sampleData/task1/dataset_`

Table 4. Figures about the training and the evaluation datasets for Task 1.

Parameter	Training dataset	Evaluation dataset
# of sentences	95	101
# of annotated entities	290	428
Avg # of annotated entities per sentence	3.51	5.37
# of entities linked to DBpedia	255	321



(a) Training dataset. (b) Evaluation dataset.

Fig. 1. Distribution of entities according to their DOLCE type.

3.2 Task 2

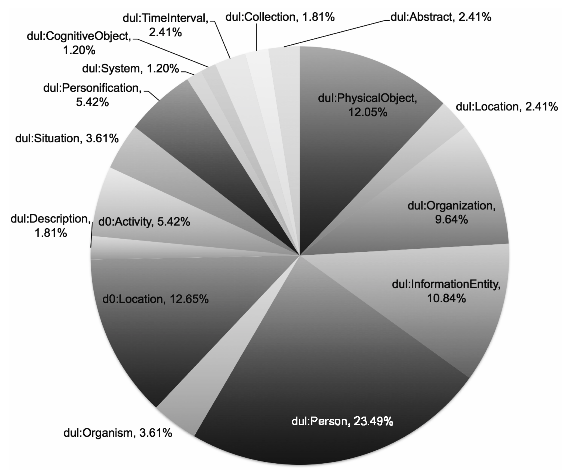
For Task 2, similarly to Task 1, we built a training and an evaluation dataset by manually annotating a set of 198 sentences, using the NIF notation. Each sentence provided a definition of a DBpedia entity expressed as natural language. The number of sentences for the two datasets was split in order to have the training dataset and the evaluation dataset composed of 99 sentences each. Table 5 reports the details about the training and the evaluation datasets for Task 2 in terms of (i) number of sentences annotated, (ii) number of `rdfs:subClassOf` axioms used within the datasets and (iii) number of classes extracted from the natural language and used for typing the DBpedia entities. It is worth remarking that each sentence provided a definition for a single DBpedia entity only, meaning that the number of sentences and the number of DBpedia entities in the datasets were the same.

Figures 2(a) and 2(b) show the distribution of entities over the subset of DOLCE Ultra Lite classes used for Task 2, for the training and the evaluation datasets respectively.

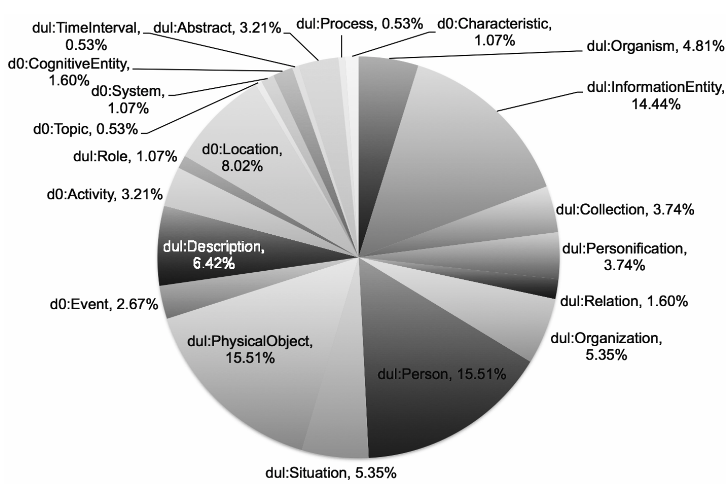
`task_1.ttl`. Similarly, the evaluation dataset is available at `https://github.com/anuzzolese/oke-challenge/blob/master/evaluation-data/task1/evaluation-dataset-task1.ttl`.

Table 5. Figures about the training and the evaluation datasets for Task 2.

Parameter	Training dataset	Evaluation dataset
# of sentences	99	99
# of `rdfs:subClassOf` axioms	166	282
# of annotated classes	165	186



(a) Training dataset.



(b) Evaluation dataset

Fig. 2. Distribution of entities over the subset of DOLCE Ultra Lite classes used for Task 2.

Both datasets are available on-line as TURTLE for download[16].

[16] The training dataset is available at `https://github.com/anuzzolese/oke-challenge/blob/master/GoldStandard_sampleData/task2/dataset_`

4 Results

The evaluation of the challenge was enabled by designing a dedicated version of GERBIL [14], which was used as benchmarking system for evaluating precision, recall and F-measure for both tasks. For Task 1 GERBIL was designed in order to evaluate systems with respect to (i) their ability to recognize entities using the NIF offsets returned by the systems (only full matches were counted as correct, e.g., if the system returned "Art School" instead of "National Art School", this was counted as a miss), (ii) their ability to assign the correct type among the 4 target DOLCE types (cf. Section 2.1), and (iii) their ability to link individuals to DBpedia 2014. Instead, for Task 2 GERBIL was designed in order to evaluate systems with respect to (i) their ability to recognize strings (i.e., linguistic evidences) in the definition that identify the type of a target entity (i) their ability to align identified types to a the subset of DOLCE Ultra Lite classes (cf. Table 2 in Section 2.2).

We received four submissions: two of them participated to one task only and the other two participated to both tasks of the challenge. Table 6 lists the systems participating to the challenge by providing the names and the short descriptions of the systems, and the tasks they are involved in.

Table 6. Systems participating to the challenge.

System	Description	Participating to Task
CETUS-FOX [11]	A Baseline Approach to Type Extraction	1-2
Adel [10]	A Hybrid Approach for Entity Recognition and Linking	1
FRED [3]	Named Entity Resolution, Linking and Typing for Knowledge Base population	1-2
OAK@Sheffield [6]	Exploiting Linked Open Data to Uncover Entity Types	2

The winner for Task 1 was Adel, which obtained micro F1 and macro F1 of 0.6075 and 0.6039 respectively. The exhaustive results for all the system involved in Task 1 is reported in Table 7.

Table 7. Task 1 results

Annotator	Micro F1	Micro Precision	Micro Recall	Macro F1	Macro Precision	Macro Recall
Adel	0.6075	0.6938	0.5403	0.6039	0.685	0.54
FOX	0.4988	0.6639	0.4099	0.4807	0.6329	0.4138
FRED	0.3473	0.4667	0.2766	0.2278	0.3061	0.1814

The winner for Task 2 was CETUS-FOX, which obtained micro F1 and macro F1 of 0.4735 and 0,4478 respectively. The exhaustive results for all the system involved in Task 2 are reported in Table 8.

`task_2.ttl`. Similarly, the evaluation dataset is available at `https://github.com/anuzzolese/oke-challenge/blob/master/evaluation-data/task2/evaluation-dataset-task2.ttl`.

Table 8. Task 2 results

Annotator	Micro F1	Micro Precision	Micro Recall	Macro F1	Macro Precision	Macro Recall
CETUS	0.4735	0.4455	0.5203	0,4478	0.4182	0.5328
OAK@Sheffield	0.4416	0.5155	0.39	0.3939	0.3965	0.3981
FRED	0.3043	0.2893	0.3211	0.2746	0.2569	0.3173

5 Conclusions

The Open Knowledge Extraction challenge attracted four research groups coming from the Knowledge Extraction (KE) and the Semantic Web (SW) communities. Indeed, the challenge proposal was aimed at attracting research groups from these two communities in order to further investigate existing overlaps between KE and the SW. 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 (i) two tasks, (ii) a training and and evaluation dataset for each task, and (iii) an evaluation framework to measure the accuracy of the systems.

Although the participation in terms of number of competing systems remained quite limited, we believe that the challenge is a breakthrough in the hybridisation of Semantic Web technologies with Knowledge Extraction methods. As a matter of fact, the evaluation framework is available on-line[17] and can be reused by the community and for next editions of the challenge.

References

1. C. Bizer, T. Heath, and T. Berners-Lee. Linked Data - The Story So Far. *Int. J. Semantic Web Inf. Syst.*, 5(3):1–22, 2009.
2. C. Bizer, J. Lehmann, G. Kobilarov, S. Auer, C. Becker, R. Cyganiak, and S. Hellmann. DBpedia - A crystallization point for the Web of Data. *J. Web Sem.*, 7(3):154–165, 2009.
3. S. Consoli and D. Reforgiato. Using fred for named entity resolution, linking and typing for knowledge base population. In F. Gandon, M. Sabou, H. Sack, E. Cabrio, M. Stankovic, and A. Zimmermann, editors, *ESWC 2015 Challenges*, Communications in Computer and Information Science, page to appear. Springer International Publishing, 2015.
4. G. R. Doddington, A. Mitchell, M. A. Przybocki, L. A. Ramshaw, S. Strassel, and R. M. Weischedel. The automatic content extraction (ace) program-tasks, data, and evaluation. In *LREC*, 2004.
5. A. Gangemi, N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. Sweetening ontologies with dolce. In A. Gòmez-Pérez and V. Benjamins, editors, *Knowledge Engineering and Knowledge Management: Ontologies and the Semantic Web*, volume 2473 of *Lecture Notes in Computer Science*, pages 166–181. Springer Berlin Heidelberg, 2002.

[17] https://github.com/anuzzolese/oke-challenge

6. J. Gao and S. Mazumdar. Exploiting linked open data to uncover entity types. In F. Gandon, M. Sabou, H. Sack, E. Cabrio, M. Stankovic, and A. Zimmermann, editors, *ESWC 2015 Challenges*, Communications in Computer and Information Science, page to appear. Springer International Publishing, 2015.
7. R. Grishman and B. Sundheim. Message understanding conference-6: A brief history. In *Proceedings of the 16th Conference on Computational Linguistics - Volume 1*, COLING '96, pages 466–471, Stroudsburg, PA, USA, 1996. Association for Computational Linguistics.
8. S. Hellmann, J. Lehmann, S. Auer, and M. Brümmer. Integrating nlp using linked data. In *12th International Semantic Web Conference, 21-25 October 2013, Sydney, Australia*, 2013.
9. G. Petasis, V. Karkaletsis, G. Paliouras, A. Krithara, and E. Zavitsanos. Ontology population and enrichment: State of the art. In *Knowledge-Driven Multimedia Information Extraction and Ontology Evolution*, volume 6050 of *Lecture Notes in Computer Science*, pages 134–166. Springer, 2011.
10. J. Plu, G. Rizzo, and R. Troncy. A hybrid approach for entity recognition and linking. In F. Gandon, M. Sabou, H. Sack, E. Cabrio, M. Stankovic, and A. Zimmermann, editors, *ESWC 2015 Challenges*, Communications in Computer and Information Science, page to appear. Springer International Publishing, 2015.
11. M. Röder, R. Usbeck, and A.-C. Ngonga Ngomo. Cetus – a baseline approach to type extraction. In F. Gandon, M. Sabou, H. Sack, E. Cabrio, M. Stankovic, and A. Zimmermann, editors, *ESWC 2015 Challenges*, Communications in Computer and Information Science, page to appear. Springer International Publishing, 2015.
12. E. F. Tjong Kim Sang. Introduction to the conll-2002 shared task: Language-independent named entity recognition. In *Proceedings of the 6th Conference on Natural Language Learning - Volume 20*, COLING-02, pages 1–4, Stroudsburg, PA, USA, 2002. Association for Computational Linguistics.
13. E. F. Tjong Kim Sang and F. De Meulder. Introduction to the conll-2003 shared task: Language-independent named entity recognition. In *Proceedings of the Seventh Conference on Natural Language Learning at HLT-NAACL 2003 - Volume 4*, CONLL '03, pages 142–147, Stroudsburg, PA, USA, 2003. Association for Computational Linguistics.
14. R. Usbeck, M. Röder, A. N. Ngomo, C. Baron, A. Both, M. Brümmer, D. Ceccarelli, M. Cornolti, D. Cherix, B. Eickmann, P. Ferragina, C. Lemke, A. Moro, R. Navigli, F. Piccinno, G. Rizzo, H. Sack, R. Speck, R. Troncy, J. Waitelonis, and L. Wesemann. GERBIL: General Entity Annotator Benchmarking Framework. In A. Gangemi, S. Leonardi, and A. Panconesi, editors, *Proceedings of the 24th International Conference on World Wide Web, WWW 2015*, pages 1133–1143. ACM, 2015.