Results of the 2016 ENtity Summarization Evaluation Campaign (ENSEC 2016)

Kalpa Gunaratna[1], Gong Cheng[2], Andreas Thalhammer[3], and Qingxia Liu[2]

[1] Kno.e.sis Center, Wright State University, USA
[2] National Key Laboratory for Novel Software Technology, Nanjing University, China
[3] Institute AIFB, Karlsruhe Institute of Technology, Germany
kalpa@knoesis.org,gcheng@nju.edu.cn,andreas.thalhammer@kit.edu,
qxliu.nju@gmail.com

Abstract. Entities and their descriptions are becoming an important part of the datasets and knowledge graphs available on the Web. These descriptions can be used in concise representation (i.e., summaries) to help users understand the Web content (e.g., summaries generated from Google Knowledge Graph in Google Search). In the recent past, several systems emerged to tackle the problem of automatic summary generation for entity descriptions. Even though these proposed systems continuously push the boundaries, the problem is not yet resolved completely. Therefore, there is a need to support and encourage researchers in the community to participate in solving this important problem. ENSEC, the entity summarization evaluation campaign, is the first step taken towards realizing that goal, and we present the results of the systems participating in the campaign.

1 Introduction

The volume of entity-centric data is rapidly increasing on the Web, including RDF-based Linked Data, Schema.org annotations, Facebook's Open Graph, and Google's Knowledge Graph. The interlinked datasets, entity annotations, and knowledge graphs on the Web describe entities (e.g., actors and films) and relations between them (e.g., starring). The description of an entity, consisting of a set of entity-property-value triples, is often too long to present to a user, primarily for (quick) understanding of the entity. As a substitute, a summary generated from the entity description can be used to efficiently support an end-user task (e.g., browsing, searching).

Specifically, an *entity summary* is a subset of entity-property-value triples selected from the description of an entity. *Entity summarization* is the process of automatically generating a high-quality entity summary, to be used for a specific task or general purpose. Several systems and approaches have been proposed [1, 2, 4] in the recent past in tackling this problem but we believe it is still far from being solved. Therefore, this ENtity Summarization Evaluation Campaign (ENSEC) is initiated and organized for the first time to assess strengths

and weaknesses of entity summarization systems, compare performance of techniques, and enhance communication among researchers and developers. In the coming years, we intend to strengthen the gold standards and create benchmarks as new requirements appear.

ENSEC 2016 consists of two tracks: the DBpedia-50 track and the LinkedMDB-30 track, and we invited researchers to try out their new proposals. DBpedia-50 and LinkedMDB-30 tracks consist of 50 and 30 entities, respectively. A system had the opportunity to participate in both or either of the tracks. Each system is evaluated against newly created gold-standard summaries by human judges.

The remainder of this report is organized as follows. In Section 2, we describe the two tracks. In Section 3, we characterize gold-standard entity summaries given by human judges. In Section 4, we present evaluation metrics. In Section 5, we report evaluation results. Finally, in Section 6, we conclude this report.

2 Tracks

ENSEC 2016 consists of two tracks: the DBpedia-50 track and the LinkedMDB-30 track.

2.1 The DBpedia-50 Track

The DBpedia-50 Track consists of 50 entities in DBpedia (version 2015-04)[4]. Their descriptions are restricted to the following datasets:

- core data: DBpedia Ontology, Mapping-based Types, Mapping-based Properties, Titles,
- domain-specific data: Images, Geographic Coordinates, Homepages, Person-data,
- categories: Article Categories, Categories (Labels), Categories (Skos), and
- YAGO types: YAGO types, YAGO type hierarchy.

For diversity purposes, the 50 entities are composed of ten entities randomly selected from each of the following five classes in DBpedia: Agent, Place, Work, Species, and MeanOfTransportation.

The description of each entity consists of at least twenty RDF triples; an entity can be either the subject or the object of a triple. All the 50 entity descriptions can be found online[5].

Each participating system should select a subset of five triples from the description of each entity, as a summary for general purposes. Considering that some systems can be configured in different ways (e.g., under different parameter settings), each participating system is allowed to submit the results of two runs under different configurations.

[4] `http://wiki.dbpedia.org/Downloads2015-04`
[5] `http://km.aifb.kit.edu/ws/sumpre2016/dbpedia50.zip`

2.2 The LinkedMDB-30 Track

This track consists of 30 entities in LinkedMDB (version 2012-02-10)[6].

For diversity purposes, the 30 entities are composed of ten entities randomly selected from each of the following three classes in LinkedMDB: Film, Actor, and Director.

The description of each entity consists of at least twenty RDF triples; an entity can be either the subject or the object of a triple. All the 30 entity descriptions can be found online[7].

Each participating system should select a subset of five triples from the description of each entity, as a summary for general purposes. Considering that some systems can be configured in different ways (e.g., under different parameter settings), each participating system is allowed to submit the results of two runs under different configurations.

3 Gold-standard Entity Summaries

Entity summaries generated by participating systems are compared against the gold-standard entity summaries created by a group of students at Karlsruhe Institute of Technology, Nanjing University, and Wright State University, as human judges. In the rest of the report, we refer to these gold-standard summaries as *ideal summaries*. We asked 10 independent human evaluators to generate ideal summaries of length 5 for each entity. That is, each ideal summary consists of five triples selected from the description of an entity for general purpose (not task-specific summaries).

Participating systems in the ENSEC 2016 campaign showed less interest in the DBpedia-50 track and hence we focused on completing ideal summaries for the LinkedMDB-30 track[8]. For the LinkedMDB-30 track, each entity received at least 6 different ideal summaries from 6 different independent evaluators. These ideal summaries and the summaries generated by the participating systems are available online[9]. In the following evaluation and results section, we present results for only the LinkedMDB-30 track.

4 Evaluation Metrics

We use the evaluation metrics as presented in Equations 1 and 2. When there are n ideal summaries denoted by $Summ_i^I(e)$ for $i = 1, .., n$ and an automatically generated summary denoted by $Summ(e)$ for entity e, the agreement on ideal summaries is measured by Equation 1 and the quality of the automatically generated summary is measured by Equation 2. In other words, the quality of

[6] `http://www.cs.toronto.edu/~oktie/linkedmdb/linkedmdb-latest-dump.zip`

[7] `http://km.aifb.kit.edu/ws/sumpre2016/linkedmdb30.zip`

[8] We will complete and strengthen the DBpedia-50 track in future. Currently, it consists of at least two different ideal summaries per entity.

[9] `http://km.aifb.kit.edu/ws/sumpre2016/ENSEC2016_LinkedMDB.zip`

an entity summary is its average overlap with the ideal summaries for the entity in the gold standard.

$$Agreement(e) = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=i+1}^{n} |Summ_i^I(e) \cap Summ_j^I(e)| \quad (1)$$

$$Quality(Summ(e)) = \frac{1}{n} \sum_{i=1}^{n} |Summ(e) \ \cap \ Summ_i^I(e)| \quad (2)$$

If we are considering only outgoing triples of an entity as its description, we simply compare the property-value pairs for agreement and summary quality. Because we are considering both incoming and outgoing triples for each entity, we compare the whole triple in computing the values using Equations 1 and 2.

System	Average Summary Quality
System 1-A	1.9722
System 1-B	1.9388
System CD	0.6444
Average Agreement	1.8199

Table 1: Average summary quality and average agreement (of ideal summaries) for the LinkedMDB-30 track for summary length of five.

5 Results

Table 1 presents the evaluation results of the systems participating in this campaign for LinkedMDB-30 track. System 1 [3] has two variations that we named *System 1-A* and *System 1-B*, received from the East China University of Science and Technology. The submission received from Nanjing University is named as *System CD* [5]. These system descriptions are included in the SumPre 2016 workshop proceedings. Table 2 presents summary quality results running each system for each entity in the LinkedMD-30 track.

6 Conclusion and Remarks

ENSEC 2016 consists of two entity samples taken from DBpedia and LinkedMDB datasets. We created a gold standard for LinkedMDB entity sample and evaluated the participating systems. We hope to continue this campaign together with the SumPre workshop series to support the community in creating better entity summaries that can lead to improvements in real-world practical systems. Further, we will investigate on new evaluation metrics to measure the quality of the entity summaries in future, which can complement the existing measures proposed in the literature.

Entity Type	Entity ID	System 1-A	System 1-B	System CD
Actor	84	1.6666	1.6666	0.6666
	766	1.5	1.5	0.5
	30242	2.3333	2.3333	0.3333
	30510	2	2	0.6666
	33739	2.8333	2.8333	0.6666
	35586	2.1666	2.1666	1
	36732	2.3333	2.3333	0.5
	39131	1.6666	1.6666	0.6666
	41033	1.3333	1.3333	0.6666
	43934	2.6666	2.6666	0.8333
Director	261	1.6666	1.6666	0.6666
	8420	1.3333	1.3333	0.6666
	8532	1.3333	1.3333	0.6666
	8685	1.6666	1.6666	0.6666
	9424	1.5	1.5	0.6666
	9562	2.3333	2.3333	0.5
	10576	2.3333	2.3333	0.3333
	11372	1.3333	1.3333	0.3333
	11556	2.1666	2.1666	0.6666
	20668	2.1666	2.1666	0.8333
Film	1511	1	1	0.1666
	7751	1.8333	1.8333	0.8333
	38753	2.3333	1.3333	0.6666
	40691	2.3333	2.3333	0.1666
	45556	2.3333	2.3333	0.1666
	49486	2.3333	2.3333	0.6666
	55491	1.5	1.5	0.5
	66221	2.1666	2.1666	1.5
	68798	2.3333	2.3333	1.1666
	81553	2.6666	2.6666	1

Table 2: Summary quality results of each system for each entity in the LinkedMDB-30 track. Entity ID reflects the actual ID in the LinkedMDB database.

Acknowledgments. We thank all the participants in the creation of gold-standard entity summaries. Gong Cheng and Qingxia Liu were supported in part by the NSFC under Grant 61572247 and 61223003, and in part by the Fundamental Research Funds for the Central Universities. Andreas Thalhammer was supported by the German Federal Ministry of Education and Research (BMBF) within the Software Campus project "SumOn" (grant no. 01IS12051). Kalpa Gunaratna received partial support from the National Science Foundation (NSF) award: EAR 1520870: Hazards SEES: Social and Physical Sensing Enabled Decision Support for Disaster Management and Response.

References

1. Cheng, G., Tran, T., Qu, Y.: RELIN: Relatedness and Informativeness-based Centrality for Entity Summarization. In: Aroyo, L., Welty, C., Alani, H., Taylor, J., Bernstein, A., Kagal, L., Noy, N., Blomqvist, E. (eds.) ISWC 2011, Part I. LNCS, vol. 7031, pp. 114–129. Springer, Berlin Heidelberg (2011)
2. Gunaratna, K., Thirunarayan, K., Sheth, A.: Diversity-aware Entity Summarization Using Incremental Hierarchical Conceptual Clustering. In: 29th AAAI Conference on Artificial Intelligence, pp. 116–122 (2015)
3. Li, Y., Zhao, L.: A Common Property and Special Property Entity Summarization Approach Based on Statistical Distribution. In: 2nd International Workshop on Summarizing and Presenting Entities and Ontologies (2016)
4. Thalhammer, A., Toma, I., Roa-Valverde, A.J., Fensel, D.: Leveraging Usage Data for Linked Data Movie Entity Summarization. In: 2nd International Workshop on Usage Analysis and the Web of Data (2012)
5. Xu, D., Zheng, L., Qu, Y.: CD at ENSEC 2016: Generating Characteristic and Diverse Entity Summaries. In: 2nd International Workshop on Summarizing and Presenting Entities and Ontologies (2016)