ASM Sc. J., 16, 2021

https://doi.org/10.32802/asmscj.2021.758

Building the Knowledge Graph for Zakat (KGZ) in Indonesian Language

Husni Teja Sukmana[1*], JM Muslimin[2], Asep Fajar Firmansyah[3] and Lee Kyung Oh[4]

[1]*Dept. Informatics Engineering, Syarif Hidayatullah State Islamic University Jakarta, Indonesia*

[2]*Graduate School, Syarif Hidayatullah JM Muslimin State Islamic University Jakarta, Indonesia*

[3]*Dept. Information System, Syarif Hidayatullah State Islamic University Jakarta, Indonesia*

[4]*Dept. Computer Engineering, Sun Moon University, South Korea*

In Indonesia, philanthropy is identical to Zakat. Zakat belongs to a specific domain because it has its characteristics of knowledge. This research studied knowledge graph in the Zakat domain called KGZ which is conducted in Indonesia. This area is still rarely performed, thus it becomes the first knowledge graph for Zakat in Indonesia. It is designed to provide basic knowledge on Zakat and managing the Zakat in Indonesia. There are some issues with building KGZ, firstly, the existing Indonesian named entity recognition (NER) is non-restricted and general-purpose based which data is obtained from a general source like news. Second, there is no dataset for NER in the Zakat domain. We define four steps to build KGZ, involving data acquisition, extracting entities and their relationship, mapping to ontology, and deploying knowledge graphs and visualizations. This research contributed a knowledge graph for Zakat (KGZ) and a building NER model for Zakat, called KGZ-NER. We defined 17 new named entity classes related to Zakat with 272 entities, 169 relationships and provided labelled datasets for KGZ-NER that are publicly accessible. We applied the Indonesian-Open Domain Information Extractor framework to process identifying entities' relationships. Then designed modeling of information using resources description framework (RDF) to build the knowledge base for KGZ and store it to GraphDB, a product from Ontotext. This NER model has a precision 0.7641, recall 0.4544, and F1-score 0.5655. The increasing data size of KGZ is required to discover all of the knowledge of Zakat and managing Zakat in Indonesia. Moreover, sufficient resources are required in future works.

Keywords: Knowledge Graph; ontology; named entity recognition; relation extraction; zakat, managing zakat

I. INTRODUCTION

The utilising of KGs in the industry has been popular since 2012 when Google released Google Knowledge Graph (Ehrlinger & Wöß, 2016; Kärle *et. al.*, 2018; Fensel *et al.*, 2020). Furthermore, the development of Knowledge Graphs (KG) is followed by many industries, including Airbnb, Amazon, eBay, Facebook, etc. (Fensel *et al.*, 2020). Besides Industry areas, KGs have also been built for research purposes and are open and accessible by any researchers. As examples of these are YAGO, Freebase, Wikidata (Krötzsch, 2017; Fensel *et al.*, 2020), DBpedia (Fensel *et al.*, 2020), etc.

Those KGs are known as Open Knowledge Graph (OpenKG), in which the content can be accessed by the public (Fensel *et al.*, 2020). However, from extensive research related to KGs, there have several definitions of terms of the KG, and there is no formal definition of a KG. This research adopts the KG definition from (Fensel *et al.*, 2020). The term KG is a graph of data, which stored knowledge of the real-world represented by nodes as entities (e.g., people, places, objects) and edges as a relation between

*Corresponding author's email: husniteja@uinjkt.ac.id

entities (e.g., located in, capital of). KGs can be implemented on various applications, including integrating, managing data from multiple data sources (Fensel *et al.*, 2020). Compared with relational databases, KGs have many advantages, powerful tools to manage extensive scale knowledge (Lv Qingjie *et al.*, 2014). The KG leveraged graph theory to capture hidden value from knowledge stored to help users get insight from their knowledge and become the system more intelligent.

KGs also can be implemented in a specific domain. In this research, we have chosen the Zakat domain as a case study. Zakat is the 3rd pillar in Islam (Farah Aida, Rashidah and Normah, 2012) that an obligation for a Muslim who has assets that have reached the limit for Zakat. Basically, some components must be fulfilled in Zakat processes: the first is *mustahik*, which is a Muslim that receives Zakat based on certain criteria (*asnaf*). Second is *Muzakki*, a Muslim or a group that has the capability to pay Zakat. The third Amil Zakat is a Muslim or institution to satisfy five basic human requirements based on *MaqasidSyariah* or goals and objectives in Islamic laws: religion, physical self, knowledge, dignity, and wealth (Farah Aida, Rashidah and Normah, 2012). Amil Zakat is assigned to manage Zakat both collect and distribute Zakat. The next component is *Nisab* of Zakat. *Nisab* of Zakat is a threshold or minimum of assets owned by *muzakki*, which is eligible to pay the Zakat. Last but least, the Zakat tariff is calculated based on the *Nisab* of assets.

In Indonesia, the Government and society have been working together managing Zakat, which the Government represents the authority to BAZNAS. A community can create an organisation in the form of LAZ (Lembaga AmilZakat) to help the Government. In this way, Managing Zakat refers to the third model of Islamic managing Zakat, which is defined by (K, 2015). As a representation of the Government, BAZNAS has several roles in managing Zakat. The first is as a regulator, which is with the Government managing rules of Zakat. It also recommends to the Government for an organisation that needs a permit to operate managing zakat. In another role, the BAZNAS becomes a coordinator and operator in managing Zakat. As a coordinator, BAZNAS has duties to ask Zakat's report implementation from operators and report it to the Government.

Meanwhile, as an operator, BAZNAS and also with LAZ are working together to run managing Zakat including collecting, distributing, and utilising Zakat. There are many BAZNAS and LAZ that spread in 34 provinces and 514 cities in Indonesia. Furthermore, there is a significant gap between the potential Zakat around 217 trillion rupiahs and the realization of zakat collection, only about 3.67 trillion (Beik & Zaenal, 2016). The vast gap is caused by many transaction zakat payments direct to individual *amil* zakat who does not register as LPZ (Sukmana *et al.*, 2017, 2019). These facts encourage digitalization in managing Zakat in Indonesia to accelerate Zakat's dissemination of knowledge to society. Building a KG for managing Zakat which is called KGZ. It is a solution towards creating an ecosystem for zakat society to increase society's awareness to pay Zakat to the official institutions and involve the community to collaborate with LPZ in the collection, distribution, and utilisation of Zakat.

This research aims to study how to create the KG, which appropriates for Zakat in which it is a specific domain. The remainder section of this paper, section 2, explains the motivation and related works, section 3 talk about the research method, section 4 discusses results and section 5 presents the conclusion and future works.

II. MOTIVATION AND RELATED WORKS

Indonesian is a kind of low resource language (Ikhwantri, 2019), where is the research of application of this language on KGs still rarely performed. There are some issues about it including existing Indonesian NER research relatively few (Wibawa & Purwarianti, 2016). The Indonesia experiment of NER is not that good compared to English (Gunawan *et al.*, 2018). Recent research was done by (Sukardi *et al.*, 2020) using the neural network combination using BiLSTM and CNN. Indonesia NER researches dominate to recognize the named entity of PERSON, ORGANISATION, LOCATION (Syaifudin & Nurwidyantoro, 2016; Gunawan *et. al.*, 2018; Ikhwantri, 2019) including MISC (Gunawan *et al.*, 2018) and EVENT (Ikhwantri, 2019). (Wibawa & Purwarianti, 2016) introduces 11 new classes of a named entity that can be used in Indonesian NER, i.e., God, Facility, Product, etc.

However, some concepts of Zakat cannot be categorized into existing named entity, i.e., instances of things that can

be used to pay the Zakat like money, rice, wheat, and cows, etc. depend on the kind of Zakat. For those reasons, the implementation of Indonesian on a knowledge graph creation is more challenging than English. Regarding this research and based on the best of our knowledge, there is no research performing in building a knowledge graph for the Zakat domain yet. Those facts encourage us to start studying how to create the knowledge graph for Zakat. This section explains investigations on building knowledge graphs that are dominant under English.

Several kinds of research related to building KGs in a specific domain have been found in various fields, including media, tourism, Government, geography, life sciences, etc. (Fensel *et al.*, 2020). And other research, the knowledge graph is implemented for fighting human trafficking (Szekely *et al.*, 2015). For a detailed example, researchers (Kärle *et al.*, 2018) built the KG for tourism that is called Tirolean Tourism Knowledge Graph. They used schema.org and extend it to define a set of vocabulary for domain specification or called DS. They were mapping data source both static data (e.g., phone number, address) and dynamic data (e.g., accommodation offers) to schema.org. They used GraphDB to build the KG after they had validated the annotation results from mapping processes.

Another example is the KG implemented for fighting human trafficking (Szekely *et al.*, 2015). In this research, the authors define five steps to build a knowledge graph. The first task is a data acquisition and extraction processes. The second is mapping to the ontology, they used schema.org to map the data acquisition results to ontology and used *geonames.org* as an entity linking to resolve the city's problem mentioning appropriating entities in GeoNames in entity linking and similarity steps. Furthermore, to build the knowledge graph, the authors used GraphDB and elastic search in the knowledge graph deployment process. Finally, they build the interface to do queries and visualization of the information or knowledge.

The KGs creation is also applied for incidental events (Mirza, Darari and Mahendra, 2018), which is required three steps to build, namely, document preprocessing, including word sense disambiguation, named entity recognition, semantic role labelling, etc. The second step is incidental event extraction and coreference resolution through document clustering which is followed by the KG construction and population. In the KG construction step, the authors built OWL ontology to discover knowledge about incidental events and documents and then mapping the output from event extraction and coreference in step 2 to their ontology using Apache Jena.

From those related works, they have different cycles to build the knowledge graph. However, we identified the general process in their cycles involving the data acquisition extraction process, mapping to their ontology, and deploying the knowledge graph.

For Indonesian, a paper is related to the KG for traditional Indonesian herbal called Jamu (Susanto, Situmorang and Virginia, 2019). This research proposed a system to manage the knowledge base about Jamu comprehensively. They designed ontology for herb-based on a taxonomy that is defined based on their survey. Therefore, the authors used the Resource Description Framework (RDF) and RDF-Schema (RDFS) to map data sources to ontology schema.

To build a knowledge graph for Zakat, there is no reference regarding knowledge graph construction in Zakat domain. Based on this fact, after we identified the general process to build a knowledge graph, we adopt (Szekely *et al.*, 2015) method with some adjustment to fit with our case. Since there is no knowledge base regarding Zakat and managing Zakat in Indonesia, we skip the implementation of entity linking and jump to the deployment knowledge graph. The main contribution of this paper can be summarized as follows:

- Present the initial knowledge graph for Zakat (KGZ) that the first knowledge graph for this domain
- Provide a NER model to identify the named entity for the Zakat domain and define 17 new named entity classes for the Zakat domain that can be referenced by researchers in future works in the same domain.
- Created a labelled dataset for KGZ-NER that is publicly accessible.

III. MATERIALS AND METHOD

Section 3 will discuss how to build the KG, in which the methodology of the KG creation depends on what domain is what will be applied. It is also adapted to the actors involved, data sources provided, the envisaged applications,

etc. (Fensel *et al.*, 2020). This section describes our approach to build the KG for Zakat (KGZ) as our case. However, our method can be applied in the general domain, especially to create the Indonesian knowledge graph.

A. *Data Acquisition*

The first task in the data acquisition step determines the data seed that will be crawled. We divided Zakat's primary source into online and offline data sources. The online sources comprise the web (e.g., BAZNAS, Rumah Zakat, LAZ Al-Azhar). And the offline sources include a book about Zakat and Wakaf; and the regulation about managing Zakat. Data acquisition is then executed to find relevant data regarding Zakat. We used the data crawler to collect online data. Then, the results of the data crawler are stored on the text files todo the next processes.

Table 1. Data sources information

	Online Sources	Offline Sources
Number of Documents	20	5
Number of Sentences	615	2,823
Number of words	7,118	17,439
Number of unique words	3,143	5,194

On the other hand, we collect the books and the regulations related to managing Zakat and convert them to digital files (PDF) for offline data sources (see Table 1). And then, those data are converted to text files. All of the data sources are in Indonesian. The entire documents are 24 documents with 15,979 words comprising 20 documents from online sources and four offline data documents. Furthermore, the next task is applying the preprocessing task to the data.

A typical Natural Language Processing (NLP) process is executed on pre-processing involving extracting data from documents. There are some tasks to do the NLP process like tokenization that separates a piece of text into smaller units or tokens, for example from the text to sentences or sentences to words. Furthermore, text cleansing is a task to remove special characters from strings.

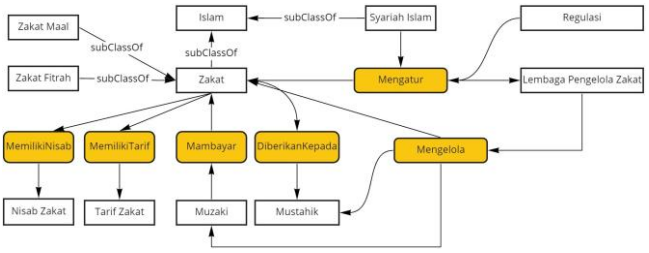

Figure 1. The taxonomy of managing Zakat in Indonesia

B. *Identification Taxonomy for Zakat and Building Named Entity Recognition (NER)*

In this step, we define the taxonomy of Zakat based on our investigation. Taxonomy aims to capture knowledge structured about Zakat. The managing Zakat taxonomy was created by study literature from several references about Zakat and managing Zakat in Indonesia. Developing of managing Zakat taxonomy is part of ontology engineering. It is provided to support defining named entity classes. Figure 1 depicts the taxonomy of managing Zakat in Indonesia. We used a judgment expert to evaluate the taxonomy through discussion with the expertise in the Zakat domain.

Zakat is divided into two types, namely, Zakat *Maal* and Zakat *Fitrah*. Zakat is a subclass of Islam and is regulated by *Syariah* Islam (Shariah of Islam) and *Regulasi* (regulations). Furthermore, Zakat is managed by Lembaga *Pengelola* Zakat (LPZ Zakat Institute). Besides Zakat, LPZ managed *Muzaki* and *Mustahik*. *Muzaki* pays Zakat based on *Nisab* and *Tarif* Zakat. LAZ distributes Zakat to *Mustahik*.

Table 2. The named entity class list

No.	Named Entity Class	Label
1	Azas_Pengelolaan_Zakat	APZ
2	Binatang_Ternak	ANI
3	Harta	HAR
4	Haul	HAU
5	Lembaga_Amil_Zakat	LAZ
6	Lembaga_Pengelola_Zakat	LPZ
7	Location	LOC
8	Logam_Mulia	MET
9	Makanan_Pokok	FOO
10	Mustahik	MUS
11	Muzakki	MUZ
12	Nisab	NIS
13	Organisation	ORG
14	Pekerjaan	PEK
15	Pengelolaan_Zakat	PZA
16	Person	PER
17	Regulasi	REG

18	Tarif_Zakat	TZA
19	Waktu	TIM
20	Zakat	ZAK
21	Zakat_Maal	MAL

Based on literature, Zakat is included in a specific domain. In this case, identifying entities on the documents in Indonesia is required to build the new NER model because the available NER (Azkiya *et. al.*, 2015; Syaifudin & Nurwidyantoro, 2016) is not sufficient to recognize the entity for Zakat. To make NER for Zakat requires a training dataset, which is not available at this moment. In this research, the dataset training is provided manually. In annotated documents training to identify the entity, we implement the BIO method to annotate the named entity. The IOB tagging is the abbreviation from Beginning (B), Inside(I), and Outside (O). This method is used to distinguish the entity either included in the type of word or phrase

For example, given a sentence:

"BAZNAS menerima pembayaran zakat fitrah dan akan disalurkan langsung dalam bentuk beras kepada mustahik"
(Literally means, "BAZNAS receives zakat *fitrah* payments and will be distributed directly in the form of rice to *mustahik*")

BAZNAS will be predicted as "B-LPZ", zakat as "B-ZAK", *fitrah* as "B-ZAK", *beras* as "B-FOO", and *mustahik* as "B-ZAK", and the remainder of the sentence will be predicted as outside that labelled by O.

To create dataset training, require a named entity class that is annotated on each word/phrase. The named entity class list is shown in Table 2. After that, we use Stanford-NER to build NER and train the model. Furthermore, we implement K-Fold cross-validation to validate KGZ-NER. This new NER model is called KGZ-NER.

C. Identification Relationship Between Entities

Based on our literature searching, there is only one research regarding relation extraction in Indonesian, done by (Gultom & Wibowo, 2017) In Indonesian Open Domain Information Extractor (ID OPENIE). We modified ID-OPENIE by replacing NER with KGZ-NER and then rebuild it using *Gradle* to get the new distribution of ID-OPENIE. The steps of how to do extraction relations can be seen in Figure 2.

ID-OPENIE using Stanford-CoreNLP as the core for NLP pipelines (e.g., Tokenizer, POS-tagger, Lemmatizer, NER, and Dependency parser). Furthermore, to extract relationships between entities, ID-OPENIE has three modules that execute consecutively. It is from a triple candidate generator with a rule-based, triple selector that applies (classifier) using machine learning with a random forest classifier. It then implements a token expander with a rule-based to produce triples as output. We analyzed the triples manually and divided them into correct, partially correct, and incorrect. The partially correct category comprises triples that have potentially correct after the refinement process, which is conducted manually based on sentences. In this task, we provide correct triples that will be used in the further task.

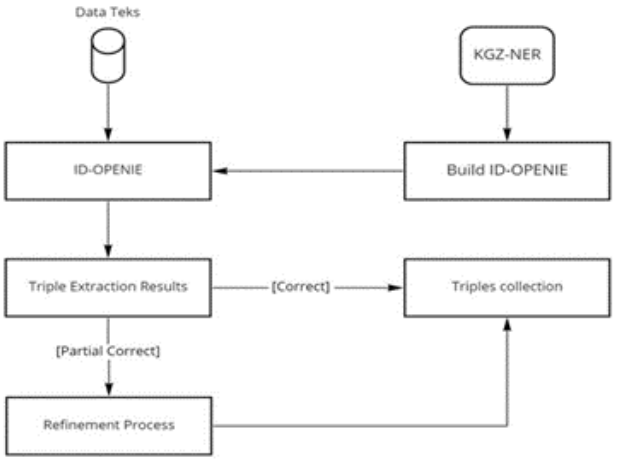

Figure 2. Identification of the relationship between entities procedure

D. Mapping to Ontology

We use the taxonomy for Zakat to create an ontology for KGZ. The purpose of ontology aims to represent nodes and edges in the knowledge graph. We adopt RDF to represent nodes and the relationship between nodes in the knowledge base model. For example, to represent ontology in RDF, we define class Zakat Maal and its properties:

```
<rdf:Class rdf:ID="Zakat_Maal">
  <rdfs:label xml:lang="id">Zakat Maal</rdfs:label>
  <rdfs:subClassOf rdf:resource="#Zakat"/>
</rdf:Class>
```

After KGZ-ontology has been created, the next step is mapping all triples to KGZ-ontology automatically by the module in python programming. And It is stored in the n-triple file formatted. We implement automatic mapping to mapping the instances from the correct triples collections that are produced by the relation extraction processes. The output of this task is generating a file of n-triples RDF that will be used as the input file to create KGZ

E. Deployment KG and Visualization

The last step in building KGZ is building the KG based knowledge base that has been created. We use GraphDB to deploy and visualize KGZ using import features. We can use query features on the SPARQL feature to evaluate KGZ with query knowledge regarding managing Zakat in Indonesia.

IV. RESULT AND DISCUSSION

The ontology for Zakat that called KGZ-Ontology fully applied to RDF and RDFS. We define class and subclasses based on the taxonomy that has been defined. All classes and their relationship are shown in Table 3 and Table 4.

Different from (Harun, Nordin and Hussain, 2008), which defines Zakat ontology focusing on applying of Zakat management system. KGZ- Ontology focuses on managing Zakat in Indonesia as interactions between stakeholders to create the Zakat domain's digital ecosystem. KGZ-Ontology consists of 8 classes, including *Zakat, Lembaga Pengelolaan Zakat, Regulasi, Pengelolaan Zaka*t, Thing, Person, Organisation, and Location. From those classes, Zakat has six subclasses that comprise *Haul, Mustahik, Muzakki, Nisab, Tarif Zakat*, and *Zakat_Maal.*

Zakat class represents all instances regarding the rule of thumbs Zakat. Furthermore, *Pengelolaan_Zakat* has two subclasses, which represent all instances of how Zakat managed in Indonesia that involve collecting, distributing, and utilising the Zakat. Because their concepts do not have derived classes, we do not make them subclasses. We define subclasses of this class as are principal (azaz) and goals (tujuan) of managing Zakat.

Lembaga_Pengelola_Zakat has one subclass, which represents all instances regarding institutions that are managing Zakat in Indonesia. The thing has four subclasses, which represent all instances as supporting other classes that have been defined. For example, Zakat can be paid by rice, wheat, or money. We define those items as instances. Furthermore, we put those instances in a class of *Makanan_Pokok* (Food) for rice and wheat; And money in the class of *Harta* (assets). All classes and subclasses are shown in Table 3.

Besides, KGZ-Ontology implements the classes, it has some fundamental properties that are implemented. Some properties have been defined as a relationship between entities. We use RDFS domain and RDFS range to define a relationship in which RDFS domain is an instance of properties to state a resource and RDFS range refers to an instance of properties to declare the value of a property. Those properties can be seen in Table 4.

Table 3. Class and Subclasses list

No	Class	Subclasses
1	Zakat	Haul, Mustahik, Muzakki, Nisab, Tarif Zakat, Zakat_Maal
2	Lembaga_Pengelola_Zakat	Lembaga_Amil_Zakat
3	Regulasi	
4	Pengelolaan_Zakat	Azas_Pengelolaan_Zakat, Tujuan_Pengelolaan_Zakat
5	Thing	Binatang_Ternak, Harta, Logam_Mulia, Makanan_Pokok,
6	Person	
7	Organisation	
8	Location	
9	Pekerjaan	
10	Waktu	

Table 4.Some properties list of KGZ ontology

No	Property	Domain	Range	Property
1	diatur_oleh	Lembaga_Pengelola_Zakat, Pengelolaan_Zakat	Regulasi	Object
2	memiliki_nisab	Zakat	Nisab	Object
3	memiliki_tarif	Zakat	Tarif_Zakat	Object
4	memiliki_jenis_pembayaran_zakat	Zakat	Thing	Object
5	mengelola	Lembaga_Pengelola_Zakat	Zakat	Object
5	mengumpulkan_zakat	Lembaga_Pengelola_Zakat	Zakat	Object
6	mendaftarkan	Lembaga_Pengelola_Zakat	Muzakki	Object
7	memberikan_bantuan_usaha	Lembaga_Pengelola_Zakat	Mustahik	Object
8	mendistribusikan_zakat	Lembaga_Pengelola_Zakat	Mustahik	Object
9	melakukan_pendayagunaan_zakat	Lembaga_Pengelola_Zakat	Mustahik, Organisation, Person	Object
10	membayar_zakat	Muzakki	Zakat	Object
11	menerima_zakat	Mustahik	Zakat	Object

We identified 272 entities after we extracted 24 documents using KGZ-NER. To build KGZ-NER model, we use stanford-NER library (Finkel, Grenage, and Manning, 2005). We designed the datasets based on a five-fold cross-validation scenario that consists of training and testing data. Afterward, we produced 5 KGZ-NER models and evaluated the model. We evaluate the model using a command provided by stanford-NER that produces outputs i.e. precision, recall, and F1-score. After a five iteration evaluation of the model, we averaged each of the evaluated results. Based on experiments on it, the KGZ-NER model has a precision, recall, and F1-score of about 0.7641, 0,4544, and 0.5655, respectively. Moreover, The distribution entities that can be identified can be seen in Figure 3.

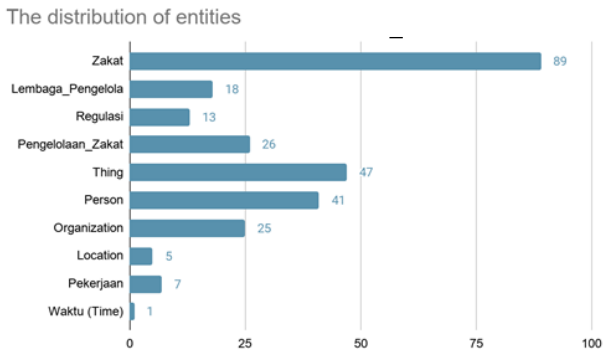

Figure 3. The identification of entities results of 24 documents using KGZ-NER

For relation extraction represented by the triple list, we identified 2,477 triples, which comprise the triples are 78 correct, 205 partially correct, and 2194 incorrect. After the refinement process, the total of correct triples is around 169 triples. These facts show that the effectiveness of using ID-OPENIE in KGZ-relation extraction, only seven percent.

At this point, All of the correct triples have been converted to RDF. Then, triples containing unique instances and relationships are merged. Figure 4 provides a snapshot of KGZ as a simple example. In Which, BAZNAS is an instance of LPZ class with a relationship with Pemerintah (Government) on a relation of *dibentuk oleh* (isFormedBy). Also, BAZNAS has a relation of *diangkat oleh* (isAppointedBy) with Presiden (President) and has a relation of *diatur dalam* (isRegulatedBy) with *Peraturan Pemerintah* (Government regulations).

The managing Zakat data stored in KGZ can be reached through the SPARQL feature of GraphDB. We can create queries depending on the information needed, such as in the list below that shows a query related to Zakat *Maal*. The result can be seen on Figure 5

```
PREFIX kgz: <http://kgz.uinjkt.ac.id/ontology#>
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
select * where {
        ?s rdf:type kgz:Zakat_Maal .
} limit 100
```

Leverage the application programming interface SPARQL query that GraphDB provides. It is allowed to integrate this knowledge graph with other applications to disseminate information about Zakat.

Our KGZs comprises 89 Zakat instances, 47 Thing instances, 18 *Lembaga Pengelola Zakat* (Zakat institution) instances, 26 *Pengelolaan Zakat* (Zakat management) instances, 41 Person instances, 25 Organisation instances, 13

Regulasi (Regulation) instances, 7 *Pekerjaan* (Profession) instances, 5 location instances, 1 Time Period instances. Those instances are depicted in Figure 6.

As an initial knowledge graph, KGZ focuses on providing linked data in the form of RDF. There is no reason for implementation yet to make the knowledge graph more intelligent, which can derive new knowledge from existing facts.

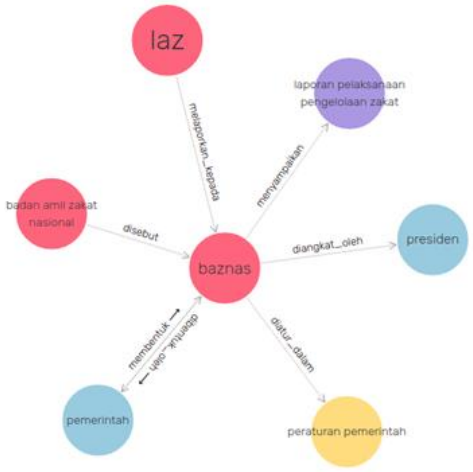

Figure 4. Snapshot of KGZ

Filter query results

1	http://kgz.uinjkt.ac.id/#zakat_barang_tambang
2	http://kgz.uinjkt.ac.id/#zakat_emas
3	http://kgz.uinjkt.ac.id/#zakat_emas_dan_perak
4	http://kgz.uinjkt.ac.id/#zakat_ghanam
5	http://kgz.uinjkt.ac.id/#zakat_harta
6	http://kgz.uinjkt.ac.id/#zakat_hewan_ternak
7	http://kgz.uinjkt.ac.id/#zakat_mata_uang
8	http://kgz.uinjkt.ac.id/#zakat_penghasilan
9	http://kgz.uinjkt.ac.id/#zakat_perdagangan
10	http://kgz.uinjkt.ac.id/#zakat_perindustrian
11	http://kgz.uinjkt.ac.id/#zakat_perniagaan
12	http://kgz.uinjkt.ac.id/#zakat_pertanian
13	http://kgz.uinjkt.ac.id/#zakat_pertanian_dan_perkebunan
14	http://kgz.uinjkt.ac.id/#zakat_perusahaan
15	http://kgz.uinjkt.ac.id/#zakat_rikaz
16	http://kgz.uinjkt.ac.id/#zakat_ternak
17	http://kgz.uinjkt.ac.id/#zakat_uang
18	http://kgz.uinjkt.ac.id/#zakat_uang_dan_surat_berharga

Figure 5. An example query using SPARQL to find the information related to *Zakat Maal*

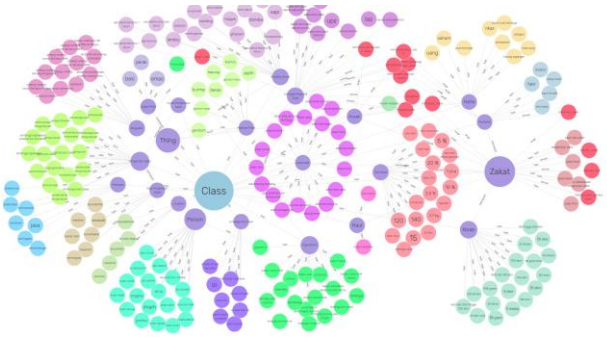

Figure 6. All instances and their relationships in KGZ

V. CONCLUSION AND FUTURE WORK

In this research, we present the first Indonesian knowledge graph in Zakat domain. As an initial work, we have defined 17 related classes and at the same time, used ENAMEX NE types, including Person, Location, TimePeriod, and Organisation to build Zakat ontology. Furthermore, KGZ-NER has been designed to detect named entity mentions and assign an appropriate class for each named entity. Due to limited resources, we were able to create a small size of labelled data for KGZ-NER. Conveniently, an existing Open source relation extraction tool for Indonesian has been utilised together with KGZ-NER to extract the relation between NEs. This, finally, contributed to generating 16 relations, and by involving 272 entities we have produced 169 triples in KGZ. Based on this research, we have many future challenges: first, we need to increase the Zakat domain's dataset benchmark including data training, data development, and data testing. It is required to improve KGZ-NER model accuracy.

Second, the existing Indonesian relation extraction is aimed at a general-purpose. This is the reason why the Indonesian relation extraction in our experiments had poor results. It appears that researching Indonesian relation extraction, especially in the Zakat domain, will be interested. Besides, we need to increase the size of KGZ to discover all of the knowledge regarding managing Zakat in Indonesia. Furthermore, sufficient resources are required in future works.

VI. ACKNOWLEDGEMENT

This research was supported by the Research and Development Center (Puslitpen) UIN Syarif Hidayatullah

Jakarta in a budget research grant of 2020, No: Un.01/KPA/ 1346/2019.

VII. REFERENCES

Azkiya, II et al. 2015, 'Pengenalan Entitas Bernama pada teks bahasa Indonesia menggunakan Hidden Markov Model', PhD Proposal, Elsevier Masson SAS, 31(0274), pp. 51–58. doi: 10.1017/CBO9781107415324.004.

Beik, IS & Zaenal, MH 2016, Outlook_Zakat_2017_Puskasbaznas. 1st edn. Jakarta: Puskas Baznas.

Ehrlinger, L & Wöß, W 2016, 'Towards a Definition of Knowledge Graphs', in Martin, M, Cuquet, M & Folmer, E (eds) Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems - SEMANTiCS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCCESS'16) co-located with the 12th International Confer. CEUR-WS.org ({CEUR} Workshop Proceedings), Available at: http://ceur-ws.org/Vol-1695/paper4.pdf.

Farah Aida, AN, Rashidah, AR & Normah, O 2012, 'Zakat and Poverty Alleviation: Roles of Zakat Institutions in Malaysia', International Journal of Arts and Commerce, vol. 1, no. 7, pp. 61–72.

Fensel, D, Şimşek, U, Angele, K, Huaman, E, Kärle, E, Panasiuk, O, Toma, I, Umbrich, J & Wahler, A 2020, Knowledge Graphs, Springer International Publishing. doi: 10.1007/978-3-030-37439-6.

Gultom, Y & Wibowo, WC 2017, 'Automatic open domain information extraction from Indonesian text', in 2017 International Workshop on Big Data and Information Security (IWBIS), IEEE, pp. 23–30. doi: 10.1109/IWBIS.2017.8275098.

Gunawan, W, Suhartono, D, Purnomo, F & Ongko, A 2018, 'Named-Entity Recognition for Indonesian Language using Bidirectional LSTM-CNNs', Procedia Computer Science, Elsevier BV, 135, pp. 425–432. doi: 10.1016/j.procs.2018.08.193.

Harun, H, Nordin, N & Hussain, A 2008, 'Ontology of Zakat Management System', in Conference: Knowledge Management International Conference 2008 (KMICE'08), Langkawi, pp. 353–358. Available at: https://www.researchgate.net/publication/271843308.

Ikhwantri, F 2019, Cross-Lingual Transfer for Distantly Supervised and Low-resources Indonesian NER. Available at: http://arxiv.org/abs/1907.11158.

K, Amiruddin 2015, 'Model-Model Pengelolaan Zakat Di Dunia Muslim', Ahkam: Jurnal Hukum Islam, vol. 3, no. 1. doi: 10.21274/ahkam.2015.3.1.137-164.

Kärle, E, Şimşek, U, Panasiuk, O & Fensel, D 2018, 'Building an Ecosystem for the Tyrolean Tourism Knowledge Graph', in Pautasso, C. et al. (eds) Current Trends in Web Engineering, Cham: Springer International Publishing, pp. 260–267.

Krötzsch, M 2017, 'Ontologies for knowledge graphs?', CEUR Workshop Proceedings, 1879(Dl).

Lv Qingjie, Lingyu, X, Jie, Y, Lei, W, Yunlan, X, Suixiang, S & Yang, L 2014, 'Research on domain knowledge graph based on the large scale online knowledge fragment', in 2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA), IEEE, pp. 312–315. doi: 10.1109/WARTIA.2014.6976259.

Mirza, P, Darari, F & Mahendra, R 2018, 'KOI at SemEval-2018 Task 5: Building Knowledge Graph of Incidents', in Proceedings of The 12th International Workshop on Semantic Evaluation. Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 81–87. doi: 10.18653/v1/S18-1010.

Sukardi, S, Susanty, M, Irawan, A & Putra, RF 2020, 'Low Complexity Named-Entity Recognition for Indonesian Language using BiLSTM-CNNs', in 2020 3rd International Conference on Information and Communications Technology (ICOIACT), IEEE, pp. 137–142. doi: 10.1109/ICOIACT50329.2020.9331989.

Sukmana, HT, Lestiani, D, Anggraeni, N & Soetarno, D 2017, 'The prototype of Zakat management system in Indonesia by using the social society approach: A case study', 2017 5th International Conference on Cyber and IT Service Management, CITSM 2017, pp. 1–4. doi: 10.1109/CITSM.2017.8089311.

Sukmana, HT, Suprihadi, R, Rosyadi, T & Subchi, I 2019, 'Applying of Recommendation and Ratting Criterion in Evaluation of Mustahik Using AHP Method', 2018 6th International Conference on Cyber and IT Service Management, CITSM 2018, IEEE, (Citsm), pp. 1–4. doi: 10.1109/CITSM.2018.8674351.

Susanto, B, Situmorang, HY & Virginia, G 2019, 'RDFS-based information representation of Indonesian traditional jamu', Lecture Notes in Engineering and Computer Science, 2239, pp. 352–357.

Syaifudin, Y & Nurwidyantoro, A 2016, 'Quotations identification from Indonesian online news using rule-based method', in 2016 International Seminar on Intelligent Technology and Its Applications (ISITIA), IEEE, pp. 187–194. doi: 10.1109/ISITIA.2016.7828656.

Szekely, P, Knoblock, CA, Slepicka, J, Philpot, A, Singh, A, Yin, C, Kapoor, D, Natarajan, P, Marcu, D, Knight, K & Stallard, D 2015, 'Building and Using a Knowledge Graph to Combat Human Trafficking', in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 205–221. doi: 10.1007/978-3-319-25010-6_12.

Wibawa, AS & Purwarianti, A 2016, 'Indonesian Named-entity Recognition for 15 Classes Using Ensemble Supervised Learning', Procedia Computer Science, Elsevier Masson SAS, 81(May), pp. 221–228. doi: 10.1016/j.procs.2016.04.053.