Journal of King Saud University – Computer and Information Sciences xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Linked open data for halal food products

Nur Aini Rakhmawati *, Jauhar Fatawi, Ahmad Choirun Najib, Azmi Adi Firmansyah

Jl. Raya ITS, Department of Information Systems, Institut Teknologi Sepuluh Nopember, Surabaya 60111, Indonesia

ARTICLE INFO

Article history:
Received 29 November 2018
Revised 23 March 2019
Accepted 5 April 2019
Available online xxxx

Keywords:
Linked data
Food product
Halal
Nutrition

ABSTRACT

The life of a Muslim cannot be separated from the concept of halal in his daily life, especially in food. Indonesia is a country that its major population is Muslim. In Indonesia, Institute For Foods, Drugs, And Cosmetics Indonesian Council Of Ulama as known as LPPOM MUI is an organisation that provides halal food product information. However, lack of information is presented on LPPOM MUI site. Therefore, LODHalal, a Linked Open Data system for halal products is proposed. We introduce a halal vocabulary which extends two food vocabularies. Then, collected data from LPPOM MUI, E-Number, Open Food Facts are transformed to RDF. Also, the RDF data are connected to DBpedia, PubChem and Mesh. We develop a web application and an Android application which provides recommendations on the halal and nutritional status of a food product.

1. Introduction

The number of Muslims worldwide in 2010 reached 23.9% of the world's population or approximately 1.6 billion (Kettani, 2010). Those number is predicted to increase over time because the annual growth rate for the World Muslim population is estimated at 1.705%, while the world population annual growth rate is only at 1.194%. The increasing number of the Muslim population of the world will also increase the need for halal products as well as the consumption and promotion of the halal food (Sham et al., 2017).

Halal refers to a code of conduct which is permitted by Shariah which it applies to every activity carried out by a Muslim (Rezai et al., 2012). It is widely known that consumption of alcohol, pork, blood, dead meat and meat that had not been slaughtered according to Islamic rules are forbidden for all Muslims (Salman and Siddiqui, 2014). Therefore, a Muslim must make sure their food consumption that is halal (Hashim and Othman, 2011).

Indonesia is one of the largest Muslim-majority countries in the world. Percentage of Indonesian Muslims reach up to 12.7 % of the world's population. 88.1 % of the 205 million Indonesians are a Muslim (Indrawan, 2015). Therefore, Indonesian government established the Institute For Foods, Drugs, And Cosmetics Indonesian Council Of Ulama (LPPOM MUI), an authorised institution to supervise hal al food products that distribute in Indonesia.

To date, LPPOM MUI provides a website http://www.halalmui.org/, where users can search a halal certificate of a product based on either product name or company name. The search results show the list of a product name, including its company name, certificate number and expiry date of the halal certificate. Also, the full list of the halal-certified product is only available in PDF format which is difficult to extract the data and to integrate to another dataset. For instance, product information should provide a list of ingredients of the product and any information related to the ingredients. It might be useful if all related product information can be displayed on a single website. Therefore, we exploit Linked Data technology for retrieving data from multiple sources.

Our contribution can be explained as follows:

1. proposing a halal food vocabulary that is enhanced from two food existing vocabularies.
2. transforming set of related halal food data into Linked Data
3. integrating data to others Linked Open data
4. building LODHalal, a web application and an Android application that are able to search a food product and predict a halal status of an uncertified-halal product.

This paper is structured as follows: Section 2 describes related work in halal and nutrition. Data sources used in this system are explained in Section 3. The design of vocabulary and interlinking between datasets can be found in Section 4 and Section 6 respectively. Our halal application is detailed in Section 7.

* Corresponding author.
E-mail address: nur.aini@is.its.ac.id (N.A. Rakhmawati).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2019.04.004
1319-1578/

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

Table 1
Comparison of Linked Data Systems for Halal and Nutrition.

Research	Ontology	App	Features
Kasim et al. (2017), Halal Checker (Kasim et al., 2017)	E-number ontology	Mobile	E-Code checker
Open Food Facts	LIRMM Food ontology	Mobile & web	search product,food information
Ibrahim et al. (2015)	Halal institutions ontology	-	query an ontology server
LODHalal	LIRMM Food ontology + food ontology + halal ontology	Mobile & web	predicting product, entity search

Table 2
Comparison of dataset used in Linked Data Systems for Halal and Nutrition.

Research	Pubchem	Chebi	E-Number	Mesh	Crowd-sourcing	Halal Institution
Kasim et al. (2017), Halal Checker (Kasim et al., 2017)	Yes	Yes	Yes	No	No	No
Open Food Facts	No	No	No	No	Yes	No
Ibrahim et al. (2015)	No	No	No	No	No	Yes
LODHalal	Yes	Yes	Yes	Yes	Yes	Yes

2. Related works

Several related studies have been conducted by other researchers related to Linked Data on the domain of food, nutrition, and halal.

Albukhitan and Helmy (2013) introduced an automatic methodology to annotate the Arabic textual content of Web resources related to domains of food, nutrition and health. The proposed methodology uses Arabic OWL ontology. It uses linguistic patterns to discover relevant relationships between the named entities in the Arabic Web resources. The RDF results does not contain any halal property of the halal products.

El-Dosuky et al. (2012) proposed food recommendation using an ontology, TF-IDF and cosine similarity where user profiles as a dataset to build the recommendation system and the recommendation data does not produce from ingredients. Ontology is also used in Snae and Bruckner (2008) for building an expert system for finding a suitable food based on an individual nutritional profile.

Predicting a halal status of food has been done in several works such as (Kasim et al., 2017; Kasim et al., 2017; Ibrahim et al., 2015). Kasim et al. (2017) examines E-number for assisting the Muslim community in verifying the halal status of substances used as the ingredients for the manufacture of a food and beverage product. They build an ontology and database based on PubChem and Chebi dataset. Besides, an Android app, Halal Checker (Kasim et al., 2017) is also provided where a user can check a halal of the status of E-Code. Hashim et al. (2017) designed a halal framework, for tracing a halal status of flavouring ingredient which aims to increase a halal certification process.

One of the notable efforts in publishing food data in RDF is made by Open Food Facts (https://world.openfoodfacts.org/) project which uses LIRMM food ontology. The Open Food Facts project aims to create a free, open and collaborative database of food products from the entire world. Open Food facts provide a halal label. However, halal integrity becomes the central issue since everyone can give a halal label to a product without verification.

Another research on the implementation of Linked data on Halal food case study was conducted by Ibrahim et al. (2015) by presented a framework for ensuring interlocking institutional worlds integrity by using the paradigm of ontology server. The framework provides interlocking halal institutions in the worlds with a consensus vocabulary to enable them to inter-operate semantically on the one hand. Moreover, it provides institutional worlds participants with an API to query the ontology server.

Our work, called LODHalal introduces halal ontology, consist of halal vocabulary which extend two food vocabulary and combine other ontology such as LIRMM Food and Food Ontology. The halal vocabulary contains a set featured of halal information such as certification code, expired date of the certification, and halal organisation. We combine several datasets to adjust our proposed halal ontology and generate RDF. Our dataset consists of LPPOM MUI, E-Number, PubChem, and Mesh.

We will be comparing related work using these criteria: ontology, application, features and dataset. We develop a web app and an Android application. Features criteria show the functionality of the system. Our main features are entity search and predicting product. Table 1 compares our system with others system. The detail of our system will be explained in the next sections. In addition, the datasets used in those systems are described in Table 2.

3. Data source

Our system collects data from multiple sources. Food halal status is mainly obtained from LPPOM MUI. The e-number dataset is from three Islamic sites. Some of the food information are imported from World Open Food Facts. Also, three datasets: DBpedia, MeSH, PubChem and are also accessed for getting more information about the products.

3.1. Institute For Foods, Drugs, And Cosmetics Indonesian Council Of Ulama (LPPOM MUI)

LPPOM MUI is an institution under the Ministry of Religious Affairs which issues halal certificates for products consumed by Indonesian including foods, drugs and cosmetics. Also, LPPOM MUI is also responsible for publishing halal product information both through printed and electronic media (Atmadi and Widati, 2015). Currently, LPPOM MUI has provided halal product information that can be accessed at http://halalmui.org/, an SMS gateway and an Android application. In general, the products inserted in our system are labelled as a halal product based on LPPOM MUI. However, a user can input a product in our halal system that without a halal certificate or a halal certificate which is not from LPPOM MUI since our system provides a component that detects a halal status of a product. The following is a sample of LPPOM MUI certificated halal product:

happy.tos Rasa Jagung Bakar 55 g
Nomor Sertifikat: 00100061230412.
Produsen: Sinar Kencana Agung, PT.
Berlaku hingga: 24 July 2020.

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

Table 3
E-Number Classification.

E-Number	Purpose
100–199	food colors
200–299	preservatives
300–399	antioxidants, phosphates, and complexing agents
400–499	thickeners, gelling agents, phosphates, humectants, emulsifiers
500–599	salts and related compounds
600–699	flavor enhancers
700–899	not used for food additives (used for feed additives)
900–999	surface coating agents, gases, sweeteners
1000–1399	miscellaneous additive
1400–1499	starch derivatives

3.2. E-Number

E-Number represents an additive of foods used by the food industry in the manufacture of various food products. E-Number is formulated by the European Economic Community (EEC) and universally adopted by the food industry around the world (Muslim Customer Group). The classification of E-Number is described in the Table 3.

In Islam, it is necessary to investigate additives especially the additives that come from animals. For example, the code E-441 for Gelatin is an additive derived from bone and cow skin and/or pig. Therefore, Muslims must be aware of E number before consuming a food product. The information of the halal status of E Number is obtained from the following three sites:

- http://www.alahazrat.net/islam/e-numbers-listing-halal-o-haram-ingredient.php
- http://special.worldofislam.info/Food/numbers.html
- http://www.muslimconsumergroup.com

The sample data of E-number from the Worldofislam:

Number: E102
Name: Tartrazine.
Description: Color.
Halal Status: Halal if used as 100% dry colour. Mushbooh if used as liquid colour, the solvent has to be Halal.

3.3. Open food facts

Open food facts (https://world.openfoodfacts.org/) is an initiative for establishing a collaborative, free and open database of food products from around the world. Open food facts provides information about food products including food, manufacture, ingredients and nutrition. It also presents a halal label for a food product, but the clarity of the halal status of a product cannot be accountable. Our system also uses food information in this dataset since the dataset is available on CSV and RDF format. The sample data of Open Food Facts RDF is presented in Listing 1.

Listing 1: Sample Open Food Facts Turtle.

```
@prefix foodlirmm: <http://data.lirmm.fr/ontologies/food#> .

<http://world-en.openfoodfacts.org/product/0000000000017/vitoria-
    crackers>
  a foodlirmm:FoodProduct ;
  foodlirmm:code "0000000000017" ;
  foodlirmm:name "Vit ria crackers" ;
  foodlirmm:IngredientListAsText "" ;
  foodlirmm:saturatedFatPer100g "3.08" ;
  foodlirmm:carbohydratesPer100g "70.1" ;
  foodlirmm:proteinsPer100g "7.8" ;
  foodlirmm:energyPer100g "1569" ;
  foodlirmm:fatPer100g "7" ;
  foodlirmm:saltPer100g "1.4" ;
  foodlirmm:sodiumPer100g "0.551181102362205" ;
  foodlirmm:sugarsPer100g "15" .
```

3.4. DBpedia

DBpedia is a community to get structured information from Wikipedia and make it available on the web. DBpedia allows others to search by using queries on Wikipedia and also connecting DBpedia dataset to other datasets. The full sample data can be accessed at http://dbpedia.org/data/Monosodium_glutamate.ttl. A piece of Monosodium Glutamate entity data presented in Listing 2

Listing 2: Sample DBpedia Turtle.

```
@prefix dbr: <http://dbpedia.org/resource/>.
dbr:Monosodium_glutamate rdfs:label "Monosodium glutamate"@en ;
rdfs:comment "Monosodium glutamate (MSG, also known as sodium
    glutamate) is the sodium salt of glutamic acid, one of the most
     abundant naturally occurring non-essential amino acids.
    Monosodium glutamate is found naturally in tomatoes, cheese and
     other foods."@en .
```

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

3.5. MeSH

MeSH is a medical vocabulary dictionary managed by the Library of Medicine (NLM). MeSH data structured in a hierarchy and allows searching at various specific levels of detail. MeSH already has RDF format and provides an API to connect with other datasets. A piece of Riboflavin entity data presented in Listing 3

Listing 3: Sample MeSH Turtle.

```
@prefix mesh: <http://id.nlm.nih.gov/mesh/vocab#> .
@prefix meshr: <http://id.nlm.nih.gov/mesh/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

meshr:D012256
  mesh:treeNumber meshr:D08.211.474.650, meshr:D03.633.300.507.650,
       meshr:D23.767.405.650, meshr:D03.633.100.733.315.650 ;
  mesh:allowableQualifier meshr:Q000506, meshr:Q000009, meshr:
      Q000592, meshr:Q000494, meshr:Q000633, meshr:Q000031, meshr:
      Q000276, meshr:Q000627, meshr:Q000378, meshr:Q000502, meshr:
      Q000302, meshr:Q000493, meshr:Q000191, meshr:Q000737, meshr:
      Q000134, meshr:Q000037, meshr:Q000138, meshr:Q000008, meshr:
      Q000235, meshr:Q000145, meshr:Q000528, meshr:Q000032, meshr:
      Q000819, meshr:Q000266, meshr:Q000652, meshr:Q000600, meshr:
      Q000096, meshr:Q000097 ;
  mesh:pharmacologicalAction meshr:D014803, meshr:D017319 ;
  mesh:identifier "D012256" ;
  mesh:dateCreated "1999-01-01"^^xsd:date ;
  mesh:broaderDescriptor meshr:D005415 ;
  mesh:dateEstablished "1966-01-01"^^xsd:date ;
  mesh:active true ;
  a mesh:TopicalDescriptor ;
  mesh:preferredConcept meshr:M0019064 ;
  mesh:dateRevised "2016-05-31"^^xsd:date ;
  rdfs:label "Riboflavin"@en ;
  mesh:nlmClassificationNumber "QU 191" ;
  mesh:preferredTerm meshr:T036408 .
```

3.6. PubChem

PubChem is a database provide information about chemical molecules. PubChem supports for RDF format and provides an API to connecting with other datasets. A piece of PubChem Riboflavin entity data presented in Listing 4

Listing 4: The results of PubChem SPARQL for Riboflavin.

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix biopax: <http://www.biopax.org/release/biopax-level3.owl
    #>.

<http://rdf.ncbi.nlm.nih.gov/pubchem/compound/CID493570> a biopax
    :SmallMolecule ;
rdfs:label "RIBOFLAVIN"@en .
```

4. Vocabulary

To design a vocabulary model, two related vocabularies are investigated: (1) http://purl.org/foodontology# and (2) http://data.lirmm.fr/ontologies/food#. The first vocabulary provides food class and nutritional content of food. The second vocabulary focuses on food nutrition. Based on two vocabularies, we combine several classes and add a halal certificate class of food products and halal sources of an additive.

Fig. 1 describes a relationship between classes in our vocabulary. A circle in the figure shows a class while the arrows show the relationships between classes. The vocabulary consists of seven classes, namely: *food*, *ingredient*, *food Additive*, *Source*, *certificate*, *Organization*, and *Status*.

Class FoodProduct and ingredient are extended from http://purl.org/foodontology# andhttp://data.lirmm.fr/ontologies/food#. Based on the LPPOM MUI data, each food has a certificate which issued by an authorised institution (*foaf:Organization*). A certificate has an expired date. In order to cover the food product that is not certified by any institution, each ingredient has a halal status based on its origin (*halalv:Source*). Regarding Food additives, a halal status of an additive is obtained from an organisation. There are three halalv:Status declared as *rdf:Alt* namely: Halal, Haram and Musbhooh. Foods are labelled Halal when it is allowed to eat, while they are labelled Haram

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

when it is not permitted to eat. Mushbooh is an unclear condition whether food is halal or not. For instance, given a product in Table 4, the turtle for the product is described in Listing 5. Our system generates 48176 entities which can be explained in Table 5.

Listing 5: Happy Toss Product in Turtle.

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>.
@prefix owl: <http://www.w3.org/2002/07/owl#>.
@prefix xsd: <http://www.w3.org/2001/XMLSchema#>.
@prefix foaf: <http://xmlns.com/foaf/0.1/>.
@prefix halalv: <http://halal.addi.is.its.ac.id/halalv.ttl#>.
@prefix halalf: <http://halal.addi.is.its.ac.id/foodproducts/>.
@prefix halali: <http://halal.addi.is.its.ac.id/ingredients/>.
@prefix halalc: <http://halal.addi.is.its.ac.id/certificates/>.
@prefix halals: <http://halal.addi.is.its.ac.id/sources/>.
@prefix halalm: <http://halal.addi.is.its.ac.id/manufactures/>.
@prefix food: <http://purl.org/foodontology#> .
@prefix foodlirmm: <http://data.lirmm.fr/ontologies/food#> .
@prefix gr: <http://purl.org/goodrelations/v1#>.

halalf:Happy_Tos_Rasa_Jagung_Bakar  a halalv:FoodProduct;
    foodlirmm:code "08993027163764";
    rdfs:label "Happy Tos Rasa Jagung Bakar";
    gr:hasManufacturer halalm:PT_Sinar_Kencana_Agung;
    food:containsIngredient halali:Whole_Corn, halali:Palm_Oil,
        halali:Monosodium_glutamate;
    food:ingredientsListAsText "Whole Corn, Palm Oil, Flavour
        Enhancer";
    foodlirmm:energyPer100g 280^xsd:integer;
    foodlirmm:fatPer100g 14^xsd:decimal;
    foodlirmm:saturatedFatPer100g 6^xsd:decimal
    foodlirmm:sodiumPer100g  0.12^xsd:decimal
    foodlirmm:carbohydratesPer100g 35^xsd:decimal;
    foodlirmm:fiberPer100g 4^xsd:decimal;
    foodlirmm:sugarsPer100g 1^xsd:decimal;
    foodlirmm:proteinsPer100g 4^xsd:decimal;
    halalv:certificate halalc:1.

halalm:PT_Sinar_Kencana_Agung a foaf:Organization;
    rdfs:label "PT. Sinar Kencana Agung".

halali:Whole_Corn a halalv:Ingredient;
    foodlirmm:rank 1^xsd:integer;
    rdfs:label "Whole Corn";
   halalv:halalSource halalv:Halal .

halali:Palm_Oil a halalv:Ingredient;
    foodlirmm:rank 2^xsd:integer;
    rdfs:label "Palm Oil";
 halalv:halalSource halalv:Halal .

halali:Monosodium_glutamate a halalv:FoodAdditive ;
    foodlirmm:rank 4^xsd:integer;
    rdfs:label "Monosodium glutamate";
    rdfs:comment "E621";
    halalv:halalSource
        halals:1.

halals:1 a halalv:Source;
    halalv:Estatus halalv:Mushbooh
    rdfs:comment "Miscellaneous - Flavour Enhancers. Suitable for
        the vegetarian label on the package indicates the source of
         Monosodium Glutamate is from vegetable protein or it has
        to be under Halal or kosher certification.";
   halalv:OrgSource halals:Muslim_Customer_Group.

halalc:1 a halalv:HalalCertificate;
halalv:halalvStatus halalv:Halal;
    halalv:halalCode "00100061230412";
    halalv:halalExp "24-05-2018"^xsd:date;
    halalv:OrgCert halals:Majelis_Ulama_Indonesia.
```

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

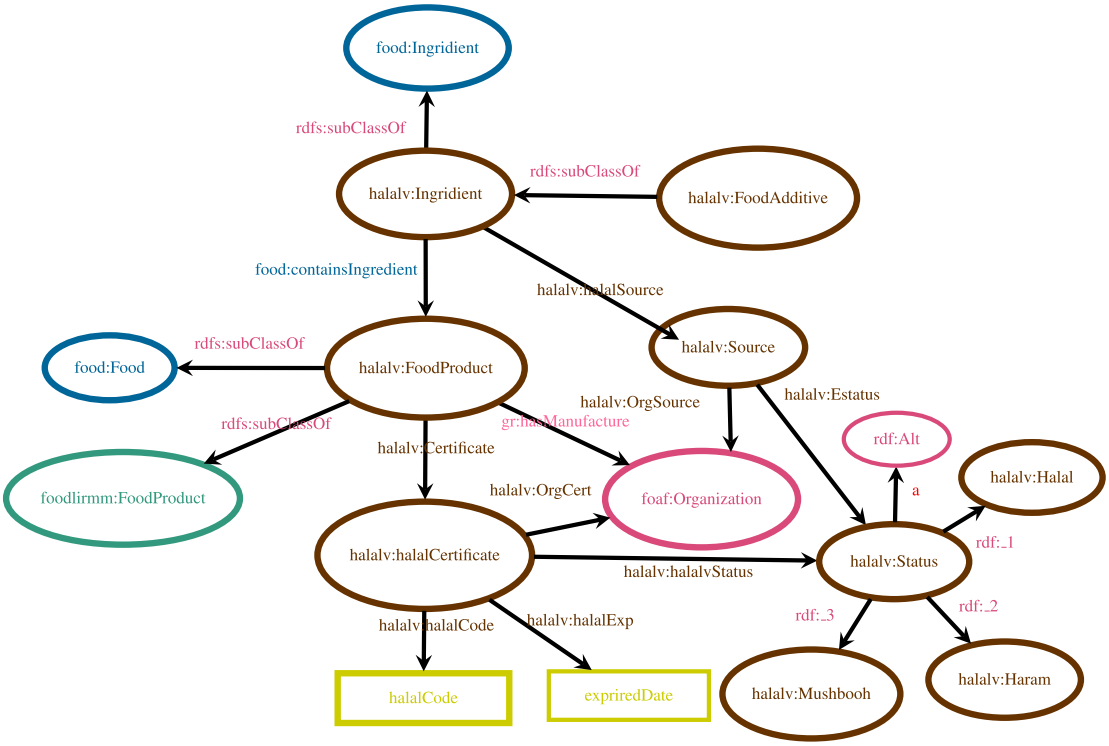

Fig. 1. Halal Vocabulary.

Table 4
Happy Toss Snack.

Property	Value
Food Code	8993027163754
Food name	Happy Tos Rasa Jagung Bakar
Manufacture	PT. Sinar Kencana Agung
Ingredients	Whole Corn, Palm Oil, MSG, Sunset Yellow FCF
Nutrition	energyPer100g:280kkal, fatPer100g:14 g, saturatedFatPer100g:6gr, sodium:120 mg, carbohydratesPer100g:35 g, fiberPer100g:4 g, sugarsPer100g:1 g, proteinsPer100g:4 g.
Certificate	00100061230412, Expired Data: 24 May 2018, LPPOM MUI

Table 5
Dataset Statistics.

Class	Number of Entities	Number of Triples
Certificates	2968	59360
Food products	42672	63262
Halal Sources	94	1786
manufactures	2442	41531
ingredient	598	2990

5. Linking

The concept of Linked data is to connect data with other data from different sources. At this stage, our dataset is connected to three datasets: DBpedia (http://dbpedia.org), PubChem, and MeSH. Mesh is a medical vocabulary dictionary administered by the Library of Medicine (NLM). PubChem provides information on the biological activity of small molecules. PubChem is managed by connecting three lined databases to the NCBI Entrez Information Retrieval System. PubChem consists of PubChem Substance, PubChem Compound, and PubChem BioAssay (PubChem).

The food additives are linked to those three datasets using *owl:sameAs*. The food additives descriptions are taken from DBpedia (Fig. 2 part number 1) and Mesh (Fig. 2 part number 2). The chemical structure of additives is retrieved from PubChem (Fig. 2 part number 3). Fig. 2 is a screenshot of Tartrazine information available at http://halal.addi.is.its.ac.id/additive/3.

6. Halal food application

Our architecture can be depicted in Fig. 3. At first steps, data from LPPOM MUI and E Number from three websites are converted to RDF. Although Open Food Fact already provided data in RDF format, the data are dirty since this data are collected by using a crowded sourcing method. For instance, a list of ingredients of a product is sometimes in an irregular pattern. As a result, it is difficult for the next computation to process the data. Thus, we clean Open Food Facts data before integrating the data. The following are the steps of data cleaning:

- change all the product and ingredients to lowercase.
- remove punctuation, numbers, and special characters.
- tokenize of ingredients part with white-space, comma and semicolon.
- remove the stop words such as *contain*, *warning*, *calorie*, etc
- compare the ingredients and product name with the existing ingredient and product with string matching methods. If the product name or ingredient exist in LODHalal, we do not insert it in LODHalal.

6.1. Ranking system

Our system provides entity searching by using the combination of Query Independent Ranking and Query-Dependent Ranking that are proposed by Delbru et al. (2010). The searching component consists of two subcomponents namely: indexing component and ranking components.

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

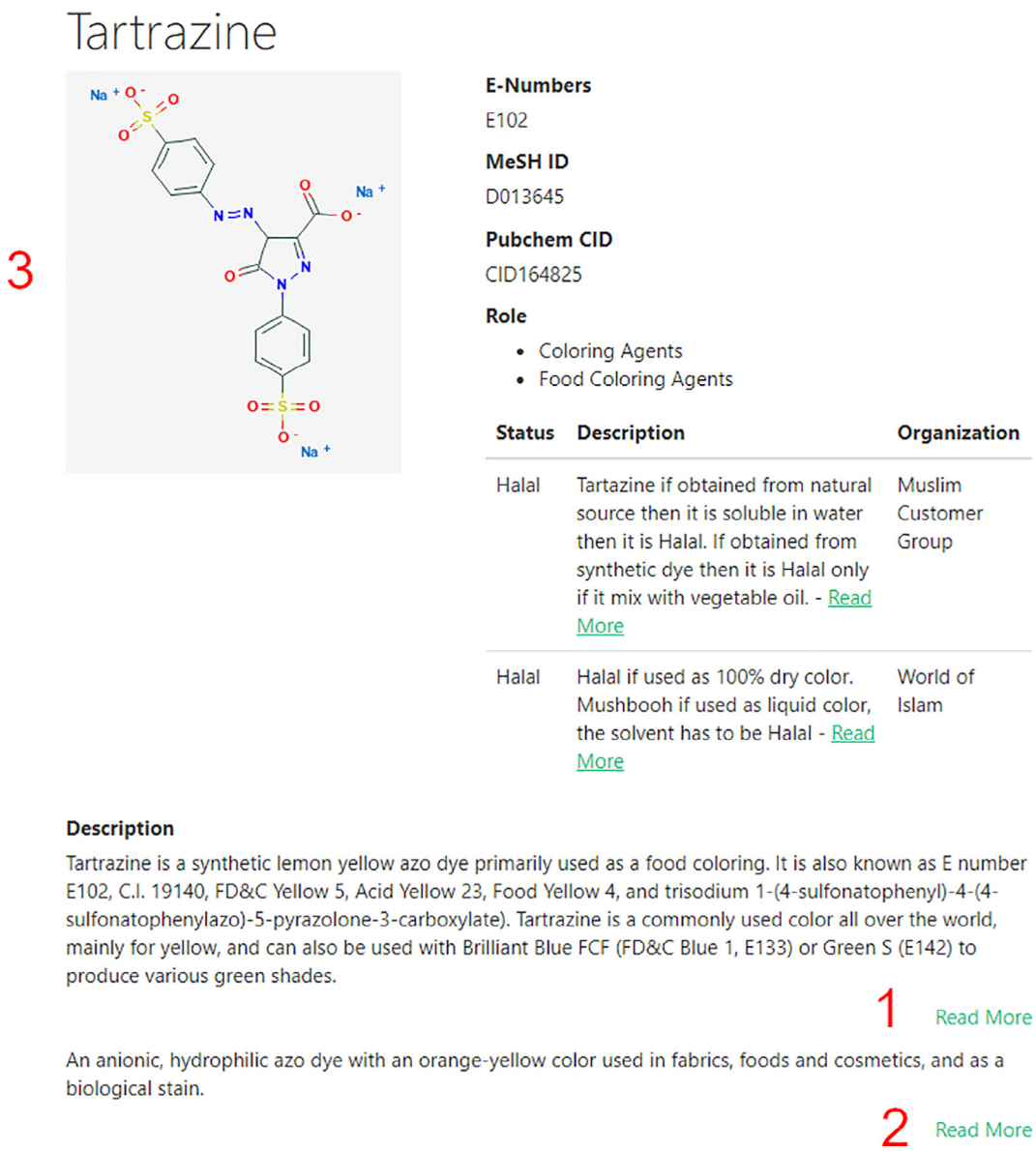

Fig. 2. Tartrazine Information displayed on LODHalal Web app.

Before ranking process, each entity is indexed in Apache Lucene (https://lucene.apache.org/). We only store the value of *rdfs:label* and *food:ingredientsListAsText* in Apache Lucene. Prefix is removed from these value. In addition, each value is tokenized and stemmed before being stored in Apache Lucene. Fig. 4 describes *Happy Toss* snack indexed in Apache Lucene, while Fig. 5 describes ingredients *Monosodium Glutamate* indexed in Apache Lucene.

In the dependent ranking, we calculate *Query Score* (Q_s) from Term Frequency-Inverse Entity Frequency (*TF-IEF*) then being normalised by the keywords entered by user (Delbru et al., 2010). Given t as term query delivered by a user, e as an entity and E is set of entities in our dataset, the *Query Score* (Q_s) can be defined in Eq. (1).

$$Q_s(t,e,E) = \frac{n(t,e)}{\sum_{k \in E} n(t,k,E)} \times log \frac{|E|}{1 + n(t,e)} \quad (1)$$

We calculate *Static Score* (S_s) by counting incoming links and outbound links each entity in the independent ranking figured by Eq. (2).

$$S_s(e,L) = \sum_{l \in e \wedge l \in L} w(l) \quad (2)$$

where L is a set of links in our dataset and $w(l)$ is weight of the link l.

The combination of static score and query score produces a final score (F_s) of a product. Final score normalize the query score by using logarithm $\log(Q_s)$, and the static score using sigmoid function (Craswell et al., 2005) with parameters $w = 1.8, k = 1, a = 0.6$x. The final score is presented in Eq. (3)

$$F_s = log(Q_s) + w * \frac{{S_s}^a}{k^a + {S_s}^a} \quad (3)$$

Those two rankings performed in our system can be found at our repository[1].

Halal Prediction component is intended for predicting a product that has not been certified any halal institution based on similarity with another certified product. In order to find similarity value between two products, we exploit two similarity measures: Cosine

[1] Github https://github.com/halaladdiits/halalnutritionfood-ranker

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

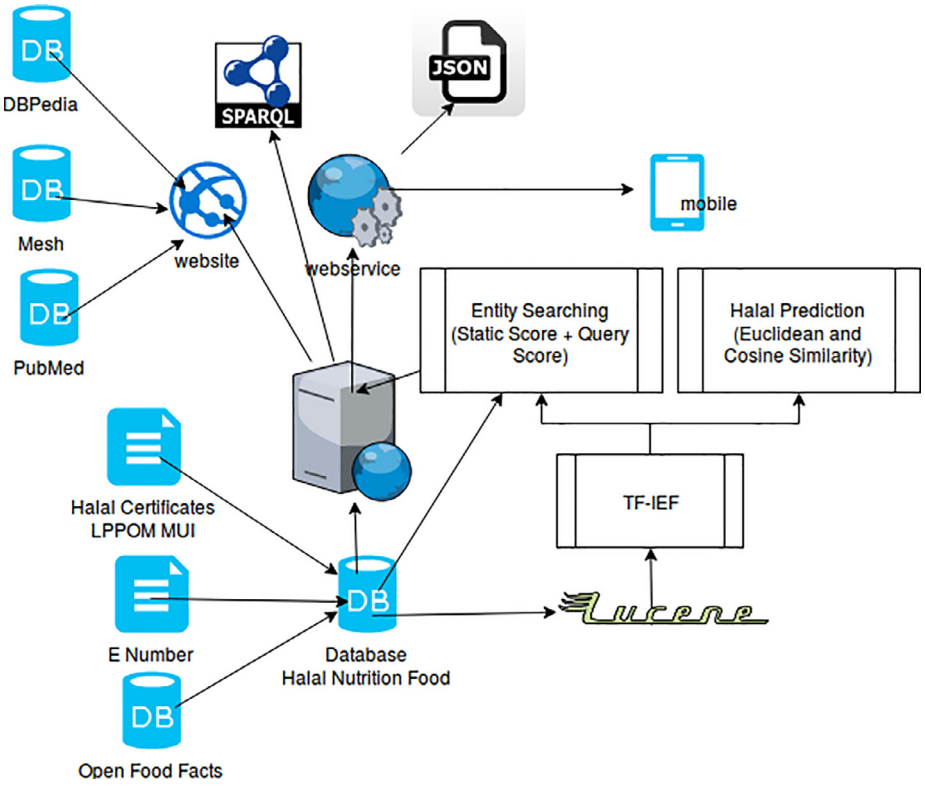

Fig. 3. Architecture System.

Doc : Happy Tos Rasa Jagung Bakar

label
bakar;happy;jagung;rasa;tos;

containsIngredient
corn;enhacer;fcf;flavour;glutamate;monosodium;oil;
palm;sunset;whole;yellow;

Fig. 4. Food Product Indexing.

Doc : Monosodium Glutamate

label
monosodium;glutamate

eNumber
E621

Fig. 5. Ingredient Indexing.

similarity and Euclidean similarity. Both of similarity measure based on the value of *TF-IEF* of each ingredient term in a product that is also used in our searching component. The halal detection of a product is explained in Rakhmawati et al. (2018).

As seen in Fig. 6, *Wafer Selamat Chocolate* has a high similarity value with *Selamat Sandwich Wafer Double Chocolate*. The notification will appear when our system finds a halal-certified product that has a similarity value higher than the pre-defined threshold value with the current browsed product. This value can be adjusted by the web administrator.

6.2. Data access point

As seen in Fig. 3, there are three points for accessing our dataset namely SPARQL endpoint, RDF dump and web API. The RDF converted data can be accessed via a SPARQL endpoint and RDF dump. Suppose that a user can look for three products that contains of the Monosodium Glutamate by using the query in Listing 6. The URL of SPARQL endpoint can be accessed in http://halal.addi.is.its.ac.id/api/sparql, while the URL of RDF dump is available at http://halal.addi.is.its.ac.id/rdf-dump.

Listing 6: SPARQL for querying three products that contain Monosodium Glutamate.

```
select * {
?s ?p <http://halal.addi.is.its.ac.id/resources/ingredients/
    Monosodium_Glutamate>
}
limit 3
```

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

Listing 7: The results of SPARQL query for Monosodium Glutamate.


```
{ "head": { "vars": [ "s" , "p" ] } , "results":
{ "bindings": [
{ "s": { "type": "uri" , "value": "http://halal.addi.is.its.ac.id/
    resources/foodproducts/boucheries_andre_cur_de_porc" } ,
"p": { "type": "uri" , "value": "http://purl.org/foodontology#
    containsIngredient" } } ,
{ "s": { "type": "uri" , "value": "http://halal.addi.is.its.ac.id/
    resources/foodproducts/
    elivia_bijou_madeleines_nature_burgers_surgeles_15_m_g_" } ,
"p": { "type": "uri" , "value": "http://purl.org/foodontology#
    containsIngredient" } } ,
{ "s": { "type": "uri" , "value": "http://halal.addi.is.its.ac.id/
    resources/foodproducts/thai_dancer_bambou_pimente_legumes_454g" } ,
"p": { "type": "uri" , "value": "http://purl.org/foodontology#
    containsIngredient" } } ] } }
```


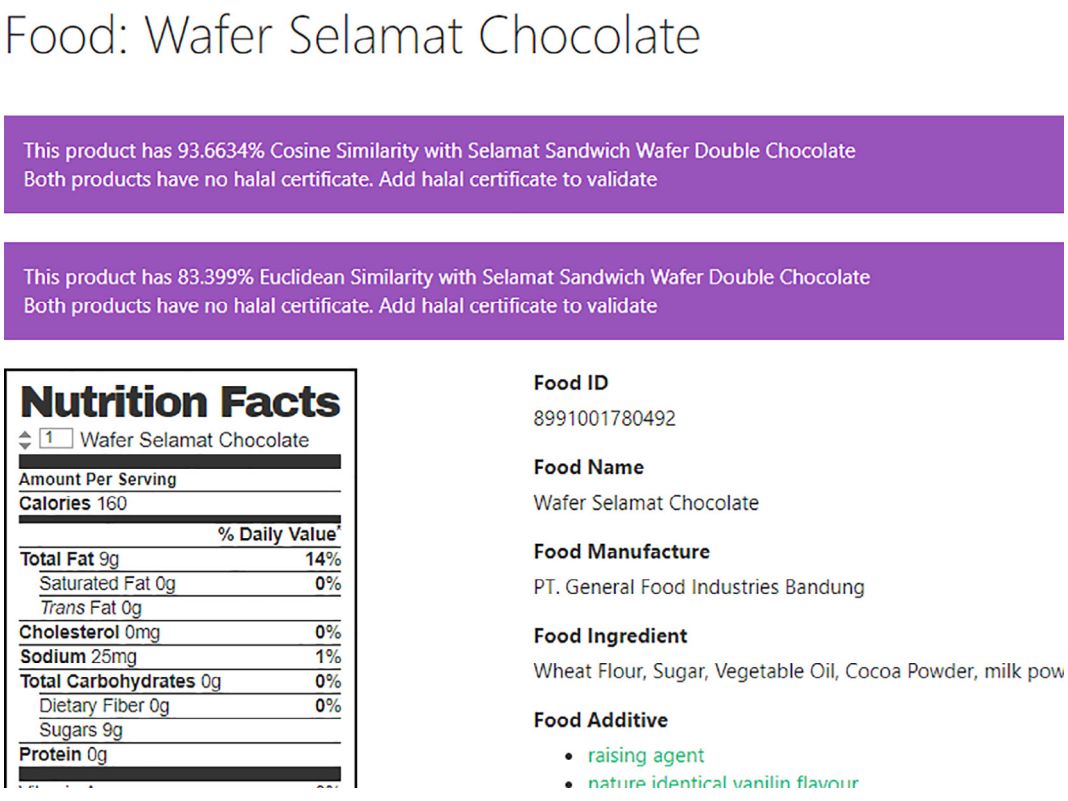

Fig. 6. Product comparison.

The aforementioned two access points only provide the RDF data without the ranking score, a user can retrieve our dataset with the rangking score using a web API (http://halal.addi.is.its.ac.id/apiv2) which uses REST technology[2]. The API server returns entities related to a keyword. For instance, http://halal.addi.is.its.ac.id/apiv2?q=monosodium&result=1 is for retrieving one of products that contains *Monosodium*.

[2] REST:https://www.w3.org/2001/sw/wiki/REST

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

Listing 8: The results of API for Monosodium Input.


```
{
    "message": "Only results 1 total matching documents collected
        . Time: 324ms",
    "entityData": [{
        "score": 2.1413969531465415,
        "label": "Monosodium Glutamate",
        "statsScore": "Ss : 81.0 - DocScore : 4.5981116 -
            NumbOfTerms : 2.0 - TFIEF : 2.2990558 - Spread : 1.25
            92592592592593 - Qs : 2.895107322269016 - Fs : 2.1413
            969531465415",
        "atribute": {
            "path": "/var/www/halal/public/resources/ingredients/
                Monosodium_Glutamate.ttl",
            "halal": "283",
            "rank": "283",
            "comment": "E621",
            "label": "Monosodium Glutamate",
            "type": "food#Ingredient",
            "sameAs": "Monosodium Glutamate"
        }
    }]
}
```


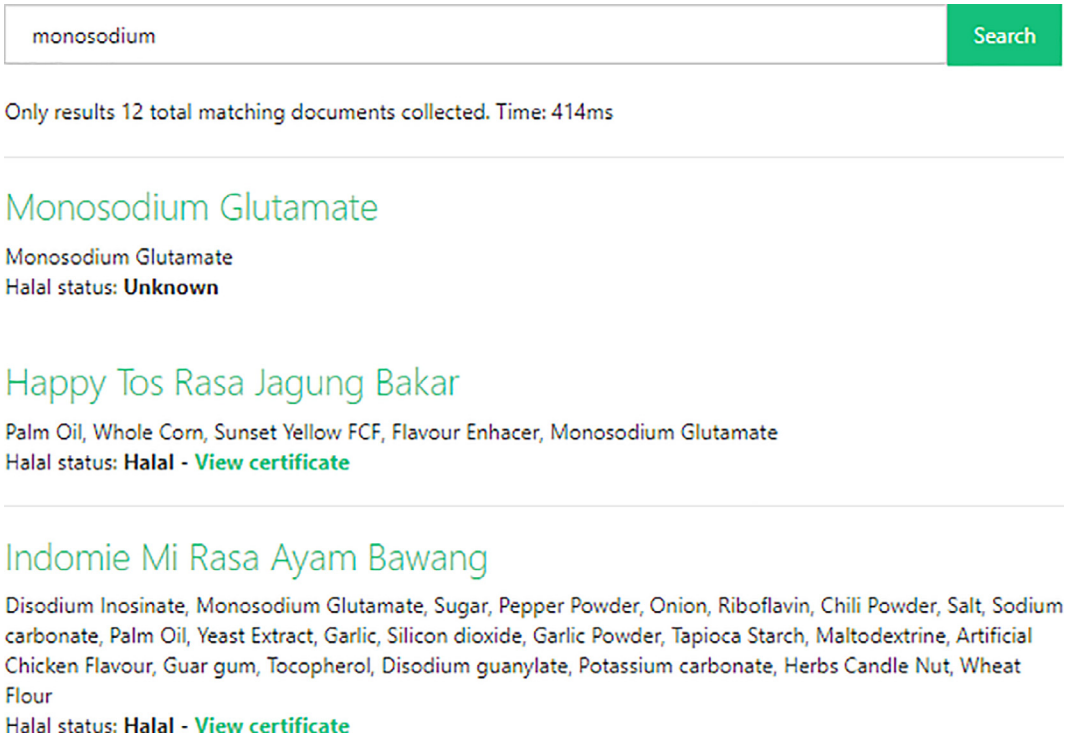

Fig. 7. The search results for "monosodium" query.

6.3. Web and android application

There are two user interfaces provided: web application and Android application. The main features of those two applications are entity search and predicting product. Fig. 7 depicts the results when user search entities that are containing "monosodium" keyword.

The Android app communicates with our system using a web service. Fig. 8 shows the android app for *monosodium* query.

The repository of our application can be seen in the Github where the web application repository is in https://github.com/halaladdiits/halalnutritionfood while the Android repository is in the https://github.com/halaladdiits/halalnutritionfood-android.

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004

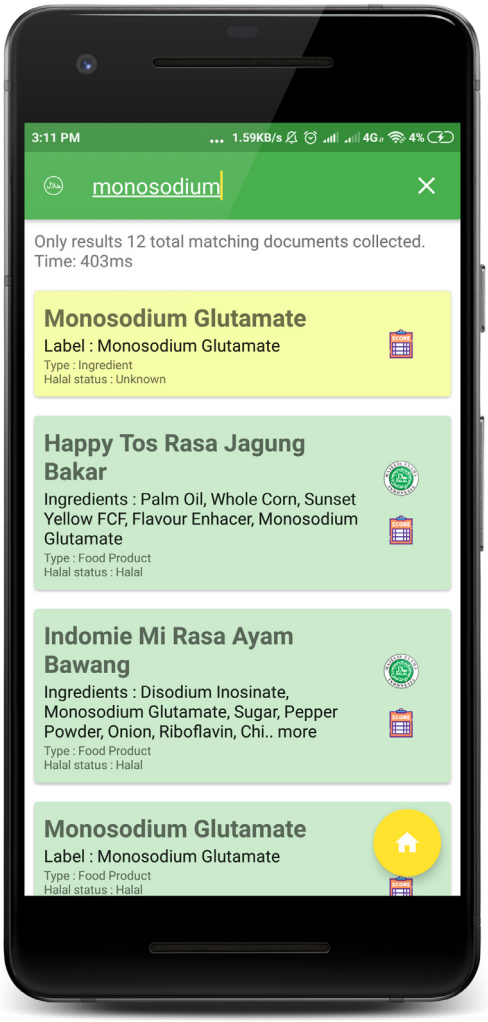

Fig. 8. The search results for "monosodium" query in the Android app.

The web application can be accessed at http://halal.addi.is.its.ac.id. The Android app that can be downloaded at Google Play store (https://play.google.com/store/apps/details?id=id.ac.its.is.addi.halal).

7. Conclusion

We have presented our system that exploits Linked Data technology in a halal food system. A halal food vocabulary is extended two existing food vocabularies. Our dataset is linked to three Linked Open Data datasets: DBpedia, PubChem and Mesh.

The searching process is used as an entity ranking method that is intended for RDF data. In order to predict a halal status of an halal-certified product, the system employs Euclidean and Cosine similarity for finding a related certified food product.

Our system provides various data access methods such as a SPARQL endpoint, RDF dump and web service. Moreover, a web app and an android app are two user interfaces provided for a user.

According to Batu and Regenstein (2014), there are 200 halal certification bodies in the world. However, there are some challenges for integrating the information of halal products from those certification bodies:

- Few the certification bodies provide a list of products along with the certification number. In the majority, they only present the procedure of halal certification. Moreover, it is difficult to find the ingredients of a product.
- A few of halal certification bodies present a list of products name in English in their website.
- the format of the data is various, further the format data is not machine-readable.
- As the best of our knowledge, there is no API provided by those halal certification bodies.

Therefore, we will advocate our vocabulary to be widely used in halal certification bodies. Our current system only obtains halal-certified data from LPPOM MUI. In order to increase the usage of our system, we would like to retrieve more halal-certified products from others halal institution as one of future works such as Halal Malaysia (http://www.halal.gov.my), Majlis Ugama Islam Singapura (https://www.muis.gov.sg/Halal), Islamic Food and Nutrition Council of America (https://www.ifanca.org/), Halal UAE (http://halal.ae/), Halal Accreditation Council of Sri Lanka (http://www.hac.lk) and Association A Votre Service (http://avs.fr). Those halal certification bodies are selected since they provide the data publicly in English language in their site. Moreover, the Open Food Facts data is dirty; therefore only a few data can be integrated with our system. As the next steps, the Open Food Facts data will be cleaned. Once we obtain more data, the prediction system can be enhanced. In the end, our system can assist a Muslim while choosing a food product wherever he goes around the world.

Acknowledgements

This research is being conducted and was supported by funding from Lembaga Penelitian dan Pengabdian kepada Masyarakat, Institut Teknologi Sepuluh Nopember (LPPM - ITS) and Kementrian Riset, Teknologi, dan Pendidikan Tinggi (or Ministry of Higher Education Indonesia) with the scheme of Department Research and the Grant No.: 1156/PKS/ITS/2017.

References

Albukhitan, S., Helmy, T., 2013. Automatic ontology-based annotation of food, nutrition and health arabic web content. Procedia Comput. Sci. 19, 461–469.

Atmadi, G., Widati, S.R.W., 2015. Strategi pemilihan media komunikasi lppom mui dalam sosialisasi & promosi produk halal di indonesia. SERI PRANATA SOSIAL 2 (2), 87–97.

Batu, A., Regenstein, J.M., 2014. Halal food certification challenges and their implications for muslim societies worldwide. Electron. Turkish Studies 9 (11).

Craswell, N., Robertson, S., Zaragoza, H., Taylor, M., 2005. Relevance weighting for query independent evidence. In: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 416–423.

Delbru, R., Toupikov, N., Catasta, M., Tummarello, G., Decker, S., 2010. Hierarchical link analysis for ranking web data. In: Proceedings of the Extended Semantic Web Conference (ESWC).

El-Dosuky, M.A., Rashad, M.Z., Hamza, T.T., 2012. A.H. EL-Bassiouny, Food recommendation using ontology and heuristics. In: Hassanien, A.E., Salem, A.-B.M., Ramadan, R., Kim, T.-H. (Eds.), Advanced Machine Learning Technologies and Applications. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 423–429.

Hashim, A.H., Othman, M.N., 2011. Halal food consumption: a comparative study between arab muslims and non arab muslims consumers in malaysia. In: Australian and New Zealand Marketing Academy (ANZMAC) Conference, Perth, Australia, November, available at: http://anzmac. org/conference/2011/Papers% 20by% 20Presenting% 20Author/Hashim,% 20Azura% 20Hanim% 20Paper, Vol. 2035.

Hashim, S., Salim, J., Noah, S., Mustapha, W., 2017. A framework for tracing the flavouring information to accelerate halal certification. J. Telecommun. Electron. Computer Eng. (JTEC) 9 (2–9), 147–153.

Ibrahim, A.A., Colomb, R.M., Ahmed, A.H., 2015. A framework of an ontology server for enhancing interlocking institutional worlds' integrity. Int. J. Metadata, Semantics Ontologies 10 (4), 219–228.

Indrawan, A., 2015. Inilah 10 negara dengan populasi muslim terbesar di dunia. http://khazanah.republika.co.id/berita/dunia-islam/islam-nusantara/15/05/27/noywh5-inilah-10-negara-dengan-populasi-muslim-terbesar-di-dunia.

Kasim, S., Azahar, U.A., Samsudin, N.A., Fudzee, M.F.M., Mahdin, H., Ramli, A.A., Suparjoh, S., 2017. E-code checker application. In: Herawan, T., Ghazali, R., Nawi, N.M., Deris, M.M. (Eds.), Recent Advances on Soft Computing and Data Mining. Springer International Publishing, Cham, pp. 570–578.

Kasim, S., Omar, N.A., Fudzee, M.F.M., Ramli, A.A., Salamat, M.A., Mahdin, H., 2017. The development of ontology from multiple databases. IOP Conference Series: Materials Science and Engineering, vol. 226. IOP Publishing, p. 012108.

Kettani, H., 2010. 2010 world muslim population. Proceedings of the 8th Hawaii Internafional Conference on Arts and Humanifies, pp. 12–16.

Muslim Customer Group, FOOD INGREDIENTS NUMBERS (e-numbers) – WorldOfIslam – halal & haram food. https://special.worldofislam.info/Food/numbers.html.

National Library of Medicine, Fact SheetMedical subject headings (MeSH). https://www.nlm.nih.gov/pubs/factsheets/mesh.html.

PubChem, About PubChem. https://pubchem.ncbi.nlm.nih.gov/about.html.

Rakhmawati, N.A., Firmansyah, A.A., Effendi, P.M., Abdillah, R., Cahyono, T.A., 2018. Auto halal detection products based on euclidian distance and cosine similarity. Int. J. Adv. Sci., Eng. Inform. Technol. 8 (4–2), 1706–1711.

Rezai, G., Mohamed, Z., Shamsudin, M.N., 2012. Assessment of consumers confidence on halal labelled manufactured food in Malaysia. Pertanika J. Social Sci. Humanity 20 (1), 33–42.

Salman, F., Siddiqui, K.A., 2014. An exploratory study for measuring consumers awareness and perceptions towards halal food in Pakistan. Interdisciplinary J. Contemporary Res. Business 3.

Sham, R., Rasi, R.Z., Abdamia, N., Mohamed, S., Bibi, T.T., 2017. Halal logistics implementation in malaysia: A practical view. IOP Conference Series: Materials Science and Engineering, vol. 226, p. 012040.

Snae, C., Bruckner, M., 2008. Foods: a food-oriented ontology-driven system. Digital Ecosystems and Technologies, 2008. DEST 2008. 2nd IEEE International Conference on. IEEE, pp. 168–176.

Please cite this article as: N. A. Rakhmawati, J. Fatawi, A. C. Najib et al., Linked open data for halal food products, Journal of King Saud University – Computer and Information Sciences, https://doi.org/10.1016/j.jksuci.2019.04.004