Multi-layered Open Data Processing Model for Hazelnut Farms

Şahin Aydın[1*], Mehmet Nafiz Aydın[2]

1 İstanbul Ayvansaray University, **2** Kadir Has University, * Corresponding author, aydinoglu.sahin@gmail.com

Abstract

In recent years several projects that are supported by information and communications technologies (ICT) have been developed in the agricultural domain to promote more precise agricultural activities. These projects account for different kinds of key ICT terms such as internet of things (IoT), wireless sensors networks (WSN), cloud computing (CC) etc. These projects are used for different agricultural products and it is a well-known fact that they can be essential to perform precise agricultural activities for the relevant agricultural products. It is important to emphasize that the success of implementing these projects depends on the extent to which support various stakeholders by leveraging relevant data, which are gathered from different kinds of data sources. Agriculture domain has a great deal of stakeholders. These stakeholders need sophisticated data and appropriate intelligence for getting benefits to performing precise agricultural activities. We agreed with scholars that "Open Data" idea, which means accessing data which is published on web and available in a machine readable format is an appropriate way to get benefit for precise agriculture by relevant stakeholders. In this paper, we shall investigate the open data term in an agriculture context and create an open data processing model. We also show viability of the proposed model by developing an ICT-based solution. Taking into account the socioeconomic importance of hazelnut for Turkey, the stakeholders of hazelnut domain still have problems such as availability, meaningful, accuracy of the hazelnut related data so we shall focus on hazelnut within the scope of this paper.

Keywords: Open Data, Multi-layered Open Data Processing Model, Open Data Processing Model for Hazelnut.

Başlık

Fındık Tarlaları İçin Çok Katmanlı Açık Veri İşleme Modeli

Özet

Son yıllarda tarım alanında hassas tarım faaliyetlerini geliştirmek için bilgi ve iletişim teknolojilerinin (BİT) kulanıldığı pek çok proje geliştirilmektedir. Bu projeler nesnelerin interneti, kablosus sensör ağları, bulut bilişim gibi BİT'in değişik teknolojilerini içermektedir. Bu projeler farklı tarımsal ürünler için kullanılmakta ve bu projeler ilgili ürünler için hassas tarım faaliyetlerinin gerçekleştirilmesini sağlamanın en bilinen yöntemidir. Değişik veri kaynaklarından elde edilen verinin işlenmesinin projelerin başarısını etkileyen faktör olduğunun altını çizmek gerekir. Tarım alanı pek çok paydaşa sahiptir. Bu paydaşlar hassas tarım faaliyetlerini gerçekleştirebilmek için uygun veriye ihtiyaç duymaktadır. Web üzerinde erişime açık ve makinaların anlayabileceği formatta olan "Açık Veri" kavramı hiç şüphesiz ki paydaşların hassas tarım aktivitelerine fayda sağlamalarının uygun bir yoludur. Bu çalışmada "Açık Veri" kavramını fındık tarımı kapasamında araştıracak ve bir açık veri işleme modeli oluşturacağız. Fındığın sosyoekonomik önemini göz önünde bulundurduğumuzda bu alanla ilgili çalışma yapan paydaşların fındık ile ilgili kaynaklara erişimi, kaynakların uygunluğu, doğruluğu ve anlamlılığı gibi konularda bir takım problemleri mevcuttur bu yüzden bu çalışma kapsamında daha çok fındık ürünü üzerinde duracağız.

Anahtar Kelimeler: Açık Veri, Çok katmanlı Açık Veri İşleme Modeli, Fındık için Açık Veri İşleme Modeli.

Citation: Şahin, A., Aydın, M. N. (2018, October) *Multi-layered Open Data Processing Model for Hazelnut Farms.* Paper presented at the Fifth International Management Information Systems Conference.

Editor: H. Kemal İlter, Ankara Yıldırım Beyazıt University, Turkey

Received: August 19, 2018, **Accepted:** October 18, 2018, **Published:** November 10, 2018

Introduction

Agriculture is a crucial sector given the fact that it contributes to employment, exportation and domestic income such that over one billion people are employed in agriculture sector all over the world. Turkey is the world's 7th largest agricultural producer overall and is the world leader in the production of dried figs, hazelnuts, sultanas/raisins, and dried apricots (URL 1). It is one of the few countries in the world with favourable weather conditions for hazelnut production, accounts for 75% of the global production and 70-75% of the exportation also 4.000.000 people are directly or indirectly related to hazelnut which has been produced in an area of 550-600 thousand hectares in Turkey, which is a fact that boosts the socio-economic importance of hazelnut (URL 2). Hazelnut data is usually published as HTML, Excel Files, and CSV formats by TUIK (Turkish Statistical Institute) in Turkey. When one looks at the context of published data, one can see that it includes only certain type of statistical information. As shall be elaborated later on, the stakeholders of hazelnut domain need to look into more sophisticated data regarding hazelnut such as hazelnut genetic resources data. Indeed, there are many problems to be solved by stakeholders for sustainable hazelnut production. Some of these problems are related to production stage, not to apply existing statutes and marketing. These problems may cause decrease in hazelnut production due to hazelnut gardens' age, structure of land and soil, treatment and fertilizing, less pollination, less pruning, blights and agricultural pests, not to preserve size of economic enterprise, increasing the hazelnut plantation area uncoordinatedly, and price fluctuation, not to apply persistent exportation policies (URL 3). Furthermore, weather conditions such as the sudden drop in temperature affect hazelnut production negatively. This type of data is generally inaccessible. It is produced by many scientific experiments. So that, scientists generally see published data as belonging to the scientific community, but many publishers claim copyright over data and will not allow its re-use without permission (Murray-Rust, 2013). In addition, it should be noted that both private and public organizations are responsible for publishing, accessing and re-using this data but this can be taken to mean that there is a permission barrier. Taking into account the needs of stakeholders, needed data related hazelnut should be accessible and reusable.

In this paper, we shall look in details of processing open data that gathered from different data sources and design a model for processing considering very exclusive layers in different perspectives.

This paper begins with an introduction section in which the research problem is explained. Second section is research background section that mentions some key terms regarding model's layers. Then, open data processing model is disclosed in model and findings section. Lastly, a case study which is developed as an exclusive system to demonstrate the model viability was explained in implementation section.

Research Background

Due to the fact that Black Sea region of Turkey has favourable weather conditions for hazelnut, hazelnut production activities are farmed in this region. This region consists of two sub-regions which are called First Standard Region (H-FSR) and Second Standard Region (H-SSR). H-FSR is the eastern part of Black Sea Region and H-SSR is middle and western part of Black Sea Region. Ordu, Giresun, Trabzon, Rize and Artvin provinces are in H-FSR. Samsun, Sinop, Kastamonu, Bolu, Düzce, Sakarya, Zonguldak and Kocaeli provinces are in H-SSR. Hazelnut is one of the most valuable agricultural products for Turkey. The statistics show that average hazelnut production is 557.270 ton according to last 14 years (from 2004 to 2017) in Turkey. Taking into account this statistical data, Turkey is the world leader in hazelnut production. Hazelnut is one of the most important agricultural exportation products in Turkey. It is one of the few countries in the world with 70-75% of the exportation. According to 2017 hazelnut exportation data, total income of hazelnut exportation is 1.866.877.685 dollars.

Multi-layered model provide to distinguish the data sources, data processing, semantic annotation, data storage, services, applications and users functions physically and logically. Taking into account the secure, easy to manage, scalable, flexible, more efficient development, easy to add new features and easy to use benefits of multi-layered model, it seems to be reasonable to creating it as multi-layered. That is why a multi-layered agricultural open data processing model is proposed in this paper. In literature, there does not exist a particular model with regard to hazelnut agricultural activities so a multi-layered open data processing model for such a valuable agricultural product is proposed the scope of this paper to manage hazelnut agricultural lifecycle properly.

The model proposed is essentially based on two pillars: the very idea of "Open Data", and "Semantic Web" approach to data modeling. We shall first provide basic understanding of these pillars so that conceptual foundation of the proposed model is established.

The term "Open Data", anyone can access, use and share. It must be accessible, which usually means published on the web, available in a machine-readable format and have a license that permits anyone to access, use and share it (Carolan et al., 2015).

Moreover, Information and Communication Technology (ICT)'s role is to facilitate distributing data to clients and stakeholders in the respective domain. All of this points to the fact that published data should be machine-readable format that can be processed by software without a labor force. The objective of the Semantic Web is to provide a new form of content that is meaningful and processable by both humans and computers (Szilaghyi and Wire, 2016).

On the other hand, Semantic web does not only refer to maintain data on the web, it is also establishing links to facilitate exploration of data by a person or machine. Under the favor of linked data, possessed information may lead to related ones. As far as Tim Berners-Lee concerned Linked Open Data has to be available on the web in different kinds of format but with an open license, to be Open Data. It should be noted that Linked Open Data must be as machine-readable structured data. For instance, excel instead of image scan of a table (Berners-Lee, 2006). This means that Linked Open Data should be as non-proprietary format. (e.g. CSV instead of Excel file.) In addition, it is important to emphasize Linked Open Data bases on open standards which are defined by World Wide Consortium (W3C). Two brief terms might clarify this definition. Resource Description Framework (RDF) is a directed, labeled graph data format for representing information in the Web and it is often used to represent, among other things, personal information, social networks, metadata about digital artifacts, as well as to provide a means of integration over disparate sources of information. This specification defines the syntax and semantics of the SPARQL query language for RDF (URL, 4). Last but not least providing context depends on linking your data to other people's data.

Understanding the challenge and difficulty of data availability can help reveal the significance of applying semantic web standards. RDF is a standard model for data interchange on the Web and it has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed (URL 5). Semantic web's RDF data support is the building blocks for developing an ontology of any domain of interest (Lata et al., 2013).

Model and Findings

As mentioned in previous sections we created a multi-layered open data processing model and focused on the hazelnut agricultural product in particular. We claim that even the contribution of proposed model within the scope of this paper rests on particularly hazelnut production lifecycle, it can be applied different kinds of agricultural products. This model is created by taking into account data sources, data processing, semantic annotations, data storage, services, applications, and users. The proposed multi-layered agricultural open data processing model contains following components; types of agricultural data sources (five types of data sources; farmers, sensors, statistical data from government, market data, and other data sources); data processing layer (raw data storage, data pre-processing and processing); semantic annotation layer (agricultural product ontology and data interchange); data storage layer (graph databases, relational databases and database services); services layer (REST web services, SOAP web services, mobile services, analysis services, reporting services and SPARQL query services); applications layer (open data web platform, open data mobile platform and open data desktop platform); and lastly end users layer (six different kinds of users; farmers, researchers, businesses, analysts, experts and other users). Figure 1 demonstrates these aforementioned components.

The agricultural data generally comes from heterogeneous sources such as farmers, sensors, government statistical data, market data and other data sources. Farmers have the main role in agricultural production cycle so their feedbacks are crucial for proposed model. Farmers have the most valid information with regard to hazelnut fields so that they are crucial data sources. A wide range of data can be obtained from farmers. For instance, age productivity and morbidity of trees, observation for soil fertility, the number of trees in orchards, location of orchards (southern slopes or northern slopes), pruning hazelnut trees, fertilization, agricultural spraying, cropping system, propagation method, irrigation method, overall vegetation surrounding the collecting site, stoniness, rockiness etc.

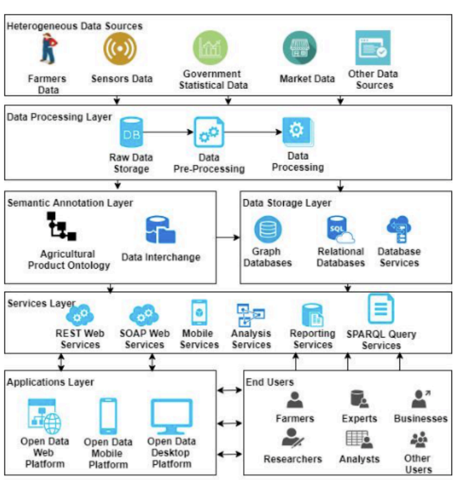

Figure 1. Proposed agricultural open data processing model.

Sensors can provide precise data about environmental effects on plant breeding. A great number of data can be obtained by sensors such as slope, weather temperature, velocity of wind, rainfall, soil moisture, soil Ph, frost (date of most recent frost, minimum temperature, duration of temperature below 0°C), relative humidity (diurnal and seasonal range), light intensity, leaf anatomy etc.

TUIK which compiles, evaluates, analyses and publishes statistics in the fields of economy, social issues, demography, culture, environment, science and technology, and in the other required areas in Turkey (URL 6). TUIK publishes data as Excel files, HTML output and CSV. The following information can be obtained from TUIK web application and services; summary of agricultural statistics, agricultural structure; production, price, value, census of agricultural holdings, general agricultural census village information, crop products balance sheets and agricultural statistics, series of official statistics with questions.

Agricultural marketing covers the services involved in moving an agricultural product from the farm to the consumer. Numerous interconnected activities are involved in doing this, such as planning production, growing and harvesting, grading, packing and packaging, transport, storage, agro- and food processing, distribution, advertising and sale (URL 7). The following information might be obtained from the market; supply, demand, crop monitor, policy developments, international prices, future market, and fertilizer outlook.

The experts that interest in relevant agricultural product might contribute to the system by loading some exclusive data. So these kinds of data sources are defined as other data sources within the scope of the proposed model.

According to NIST (National Institute of Standards and Technology), cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction (Mell and Grance, 2018). So considering the aforementioned definition of cloud computing it is convenient that data storages, services and applications of the proposed model should serve on the cloud.

Data Processing Layer (DPL) consists of three components; raw data storage, data pre-processing, and data processing. Raw data storage component is responsible for storing raw data that is gathered from all data sources. In general, storing raw data is essential due to the fact that stakeholders might need it. The data is stored on cloud in the format that gathered from data sources. This data can be structured, unstructured and semi-structured so that data storage types differ such as file system, RDBMS (Relational Database Management System), NoSQL database system, main-memory database system, graph database system etc.

The raw data is not usually significant or valid. There has to be a mechanism to identify and correct errors with respect to raw data which is collecting different data sources. Pre-processing component is responsible for decomposing data as significant or not. Data pre-processing is a means of transforming raw data, which can be incomplete or inconsistent and have many errors or unusable parts, into an understandable format that is ready for a more efficient analysis process (Siddiqa et al., 2016). The purpose of this component is implementing data pre-processing techniques. Within the scope of this component following pre-processing techniques are applied to improve data quality; data integration, data enhancements and enrichment, data transformation, data reduction, data discretization and last but not least data cleansing (Taleb, Dssouli and Serhani, 2015).

The purpose of implementing storing and pre-processing techniques is to make preparations for gaining insights that are significant and valuable that are parts of data processing. This point to the fact that data processing is essential within the scope of proposed model. The term data processing refers to producing significant information from raw data. Meanwhile, supplying meaningful information by analyzing data is essential for data users to make future predictions and support producing analytical solutions. According to Begoli and Horey's design principles (support a variety of analysis methods, one size does not fit all, and make data accessible) which are largely concerned with maximizing the controllable factors and thereby enabling researchers to explore, analyze, and interact with data in as easy manner as possible (Begoli and Horey, 2012). The processing layer provides users operational analytic solutions in respect to principles.

After aforementioned techniques are applied to data by components of DPL, data is converted RDF datasets as well as structured data by data interchange component which is a component of semantic annotation layer. Transformation of data is performed by considering agricultural product ontology. Our focus is on hazelnut agricultural product in this paper. Hazelnut ontology is intended to be created using "Descriptors for Hazelnut" which consists of five different categories; passport, management, environment and site, characterization, and evaluation (Koksal and Gunes, 2008). As mentioned before our focus is on creating an open data processing model in this paper so information about creating hazelnut ontology will not be mentioned. However, it is important to emphasize that data interchanging (from raw data to RDF triples) should be performed according to agricultural product ontology.

Data underlying RDF datasets that are gathered different data sources and processed by DPL has been storing in graph databases which are the components of data storage layer as part of proposed model. RDF datasets are crucial for proposed model due to the fact that the data heterogeneity. Data is obtained from different data sources so that data format is essentially different. That is why entire data is converted RDF triple which consists of subject, predicate, and object (or subject, property, value). It should be noted that RDF data management is crucial for proposed model and it can be examined in a variety of contexts. This variety is actually reflected in a richness of the perspectives and approaches to storage and indexing of RDF datasets, typically driven by particular classes of query patterns and inspired by techniques developed in various research communities (Luo et al., 2012). Relational databases component of data storage layer (DSL) is responsible to store relational data. Database services that is a component of DSL provide secure connection, session management, user mapping, user authentication and authorization between components of services layer (SL).

Other essential layer of proposed system is services layer which consists of six different types of services; REST web services, SOAP web services, mobile services, analysis services, reporting services, and SPARQL query services. According to W3Schools web services are defined as self-contained and self-describing application components that can be discovered using UDDI, be used by other applications and they communicate using open protocols (URL 8). Web Service Layer is a part of proposed model. It is a kind of API (Application Program Interface) that provides using and processing agricultural RDF and other datasets regarding hazelnut. Meanwhile, it supports REST (Representational State Transfer) operations thus it can be consumed by different software applications. When the power of mobile technologies is considered today, we have believed that open data processing model should include mobile services. So mobile services which provide to connect different kinds of mobile devices has to need to take part of proposed model.

Due to the fact that stakeholders of agriculture might have lack of data analysis methods, the model has analysis services. This services has used distinct algorithms to gain acquisition from agricultural data. The stakeholders particularly experts, researchers and analysts might get specific information using these services without a lot of labor. Reporting services provide to create analyzed data into the report format, manage and publish them. Also the stakeholders can deliver these reports to whoever they would like. SPARQL Query Service is a service of proposed model that enables querying large RDF datasets with regard to hazelnut. As it is mentioned in data management part, RDF datasets are stored in graph databases on the cloud. This provides a simple and fast querying service to users.

Applications Layer (AL) allows end users to access agricultural data which are gathered from heterogeneous data sources. AL consists of web, mobile and desktop applications that provide to import, export, report, analyze, and query data.

From the agricultural open data processing model figure, it is apparent that end users consist of farmers, researchers, businesses, analysts, experts, and other users. All of these end users are hazelnut agriculture stakeholders within the scope of proposed agricultural open data processing model. Users can obtain expressive information from services provided by services and applications layer. This data is used to gain benefit into sustainable hazelnut agricultural activities by researchers, analysts, and experts. It is worth bearing in mind that setting a certain estimate for hazelnut is a challenge due to its unpredictable harvest in every year. There are few institutions with regard to estimating the harvest of hazelnut in Turkey so that the estimates rapidly fluctuate. This affects hazelnut market prices negatively. Given this explanation for hazelnut market, the importance of users are crucial for proposed model. The farmers benefit from system output to prevent agricultural problems and plan their agricultural activities. The researchers, analysts, and experts make use of conferred data to investigate statistical analysis, scientific findings, and market analysis. Businesses which are related to hazelnut's production and market price have a need of making decision about their investment, employments, and market so that the information is inferred from presented data of proposed model.

On the other hand these stakeholders make contribution through services layer using SPARQL Query services. Farmers inform general situation with respect to their hazelnut fields such as fertile and infertile number of hazelnut trees, and total number of hazelnut trees in orchards and crop yield per orchard. Researchers, analysts, and experts give information about existing or recently discovered plant diseases, and fertility factors. Eventually businesses give feedback regarding production standards, market price, packaging, shipping and storage, industrial problems, foreign trade and product tagging.

Implementation

Up to this point an open data processing model is proposed. Viability of this model has been tested with a prototype of electronic system developed using wireless sensor network (WSN), different kinds of sensors, and other electronic devices. We focused on only the lifecycle of sensors data to test proposed system whether it is feasible or not in

this paper. There are two sensors nodes named router 1 and router 2 in the system respectively and these nodes transmit measured data to coordinator. This system structure is an implementation of ordinary and basic WSN. Weather temperature, weather humidity, weather pressure, carbon monoxide, nitrogen, oxygen, ultraviolet detection sensors have been plugged on router 1 and digital light intensity, soil moisture, and precipitation detection sensors have been plugged on router 2. While some of these sensors measure different types of environmental events or changes, others measure just one. Weather pressure, temperature, humidity changes have been detected by only one sensor named BME280. Likewise, rates of carbon monoxide, nitrogen and oxygen have been measured by only one sensor named MICS-4514. Arduino Uno micro-controller boards based on ATmega328P have been used to plug the sensors and measure the values of them in the system. XBee Pro S2C modules embedded solutions providing wireless end-point connectivity to devices have been used to transmit data from routers to coordinator. Besides, routers have solar panels to charge their lithium batteries. The coordinator consists of a Raspberry Pi 3 small computer, an Arduino Uno, and an XBee Pro S2C module. Raspberry Pi has been used to run applications and services. A universal Windows software application are running on it and this application demonstrates the sensor values. Arduino Uno and Raspberry Pi have been connected to each other using I2C communication protocol. XBee Pro S2C module has been plugged on Arduino and received data from all sensors of routers. The universal application was developed by using C# programming language in Visual Studio 2015. This application runs on Windows IoT Core operating system which was built for IoT. The routers transmit data in a particular format named frame. Frame data is in a ZigBee packet. ZigBee is an open global standard for low-power, low-cost, low-data-rate, wireless mesh networking based on the IEEE 802.15.4 standard and it represents a network layer above the 802.15.4 layers to support advanced mesh routing capabilities (DIGI, 2018). The frame data in a ZigBee packet means that data between the length LSB and the checksum. LSB represents the least significant byte and checksum refers to the last byte of the frame and helps test data integrity. Figure 2 representing the format of transmitted data from routers to coordinator.

Figure 2. Frame data format.

This frame data refers to "RID:1;SID:1;WT:32.25". RID, SID and WT have been meant router ID, sensor ID and weather temperature respectively. It has been processed to store into relational database management system by DPL and then stored on Microsoft Azure MS SQL Server which is a component of data storage layer. The data gathered from sensors and stored on cloud has to be published in a suitable format for applications running on different platforms. Therefore, REST web services were developed using Windows Communication Foundation (WCF), which is a framework for building service-oriented applications. Furthermore, web services for Windows platforms were developed in the same way. Figure 3 illustrates the formatted data in JSON (JavaScript Object Notation) format.

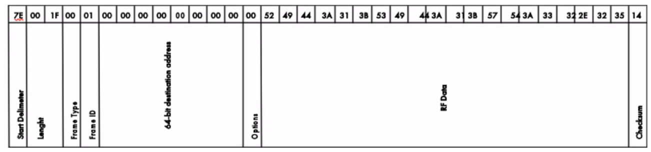

Figure 3. Formatted data storing relational database.

This data is accessible in an appropriate format for anyone who wants to use. Within the scope of this implementation an open data web or mobile platform have not been developed. We believe that understanding just one case which demonstrates the lifecycle of data can help reveal the validity of the proposed model in this paper.

Conclusions and Future Work

In this paper we proposed a multi-layered open data processing system that consists of different layers and components. All layers and its components were explained and an implementation regarding sensors was developed and summarized in implementation section. We examined only one aspect in this implementation. However, there are four other aspects in terms of heterogeneous data sources. These other aspects are worth discussing and examining in the future. Also, we tried to justify the suitability of the data life cycle in terms of our proposed multi-layered model by implementing aforementioned wireless sensor network system. We feel that our study enhances academic understanding of creating a multi-layered open data processing model. On the contrary the life cycle of other data from other data sources should be examined and justified the viability. In addition, we mentioned about creating hazelnut ontology but did not give detail information with respect to it. As a future work an ontology for hazelnut shall be created and adapted to the model. A number of restrictions of our study and areas for future research should be mentioned and the following ones might be given as examples; difficulty of gathering data from farmers (lack of knowledge of technology usage), defining the boundaries of gathering market data, determining the potential data sources which is defined as "Other Data Sources" in heterogenous data sources layer, and spreading the usage of applications mentioned in "Applications Layer".

References

URL1: http://www.invest.gov.tr/en-US/sectors/Pages/Agriculture.aspx [21.07.2018]

URL 2: http://www.ftg.org.tr/en/turkish-hazelnut-turkeys-hazelnut.html [21.07.2018]

URL 3: http://www.findikihracati.com/findik/findik-veriminde-ve-

uretiminde-temel-sorunlar.html: [21.07.2018]
Murray-Rust, P., 2013, "Open Data in Science", Serials Reviewi 34:1, 52-64
Carolan, L., Smithy, F., Protonotarios, V., Schaap, B., Broad, E., Hardinges, J., Gerry, W., 2015, How can we improve agriculture, food and nutrition with open data? , Open Data Institute
Szilaghyi, I., Wira, P., 2016, Ontologies and Semantic Web for the Internet of Things – A Survey, IEEE, 6949-6954
Berners-Lee, T., 2006, Linked Data, https://www.w3.org/DesignIssues/LinkedData.html [23.07.2018]
URL 4 : https://www.w3.org/TR/2013/REC-sparql11-query-20130321/ [23.07.2018]
URL 5: https://www.w3.org/RDF/ [23.07.2018]
Lata, S., Sinha, B., Kumar, E., Chandra, S., Arorar, R., 2013, "Semantic Web Query on e-Governance Data and Designing Ontology for Agriculture Domain", International Journal of Web & Semantic Technology (IJWesT) Vol.4
URL 6: http://www.turkstat.gov.tr/UstMenu.do?metod=gorevYetki [24.07.2018]
URL 7: https://en.wikipedia.org/wiki/Agricultural_marketing [24.07.2018]
Mell, P., G., Tim, 2018, National Institute of Standards and Technology, Information Technology Laboratory – a nonregulatory federal agency within the U.S. Department of Commerce – http://csrc.nist.gov/groups/SNS/cloud-computing
Siddiqa, A., Hashem, I.,A.,T., Yaqoob, I., Marjani, M., Shamshirband, S., Gani, A., Nasaruddin, F., "A survey of big data management: Taxonomy and state-of-the-art", 2016, Journal of Network and Computer Applications 71(2016) 151–166
Taleb, I., Dssouli, R., Serhani, M.,A., 2015, "Big Data Pre-Processing: A Quality Framework", 2015 IEEE International Congress on Big Data,
Begoli, E.,Horey,J.,2012.Design principles for effective knowledge discovery from big data.In:Proceedings of the IEEE 2012 Joint Working IEEE/IFIP Conference on Software Architecture(WICSA) and European Conference on Software Architecture(ECSA).
Koksal, A.,I., Gunes, N., Descriptors for Hazelnut, 2008, Bioversity International and FAO,
Luo, Y., Picalausa, F., Fletcher, G.H.L., Hidders, J., Vansummeren, S., Storing and Indexing Massive RDF Data Sets,
URL 8: https://www.w3schools.com/xml/xml_services.asp [25.07.2018]
DIGI International Inc., XBee/XBee-PRO S2C Zigbee RF Module User Guide, 2018.