Web of Needs Using Semantic Technology: Empowering Hawkers as Use Case

Swati Sen and Prithwiraj Sen

December 1, 2023

Web of Needs using Semantic Technology: Empowering Hawkers as Use Case

Swati Sen
Department of Management Studies
Dayananda Sagar College of Engineering, Bangalore
India
Email: swatisen@live.in

Prithwiraj Sen
Independent Researcher
Bangalore,
India
Email: prithwiraj@live.com

***Abstract*—Marginal business holders or hawkers have meager scope to grow their business without sufficient investments. This research focuses on the use case of Hawkers by prioritizing the need to build a platform that can connect Hawkers with Investors. Hawker data is not stored in a structured format in any repository. Hence, it is impossible to make the data accessible to any authorities. Overall, this results in the marginal businesses to be unreachable to the willing investors.**

In this paper, we aim at explaining a methodology to make abstract data available over a web of needs and run matching service on the web. By creating a Web of Needs with Hawker data we can make their requirements available for any investor to find a prospective business. We create an automated matchmaking service that can meet the expectation of both hawker and Investor. This can help Hawkers expand their business. Initial approach is to make the Hawker data structured in a unified format in the form of RDF triples. We will be using RDF to create the web of needs and then run a matching algorithm on it. This will enable different communities to interact with each other and exchange data between different systems.This will also enable accessibility by Investors from their domain. Using Web of Needs we would summarize the data in structured format and represent it in the form of URI (so that it is accessible irrespective of application). Data is represented in the form of RDF triple so that it can be linked with other systems to make it interoperable.

1. Introduction

Web services have progressed in such a way that every business function has its corresponding online application. Any operation that can be performed on a physical platform can be devised using a virtual medium. Businesses and customers are connected via the World Wide Web (WWW) and various Web services available over the internet. However, all these services are hosted by disjoint systems. Each service has its secluded system which provides a spectrum of resources. With evolving data requirements, the need for a connected system is realized which will provide different services under a single umbrella. This is where the technologies such as the Web of Needs and Semantic Web come into play.

The Internet has evolved from the World Wide Web to the Semantic Web where context and pragmatics become more predominant for a sensible search. Semantic Web, as coined by Sir Tim Berners Lee, has made it possible for machines to understand the underlying context of a search and led its way to provide with more accurate information to the user. Eventually, the web has evolved from linked documents to both linked documents and data. Semantic Web is built on the foundation of linked data. Connecting and publishing structured data over the Internet is called Linked Data [1]. The technology to develop semantic data on the web is RDF [2]. The underlying meaning of anything can be expressed in the form of RDF triples, that is a particular subject has properties with certain values [3]. To deploy user information in the form of RDF triples we create a Web of Needs node that can accumulate and process this information for efficient matchmaking facility.

E-commerce industry experts believe that the Semantic Web can be used in matchmaking for e-business [4]. Matchmaking is a process in which businesses are put in contact with potential business partners or customers. Traditionally, this process is handled by hired brokers, and many have suggested creating a matchmaking service that handles advertising services and querying for advertised services. Experts argue that only Semantic Web technologies can sufficiently meet these requirements, and they believe that the Semantic Web can automate matchmaking and negotiation [5]. The opportunities for maximizing business opportunities with Semantic Web technologies are limitless. The aim of the Semantic Web is to convert the web from a defined set of document links to a pool of information. This transformation of web of documents to web of data is supported by RDF (Resource Description Framework). Semantic Web supports information sharing in a form where the resource does not need to be available within the same environment. This opens the gateway for interoperable systems which can work together to present data from a wider scope. RDF structures the data contained in web pages, in a format that can be processed by machines without human interactions. RDF is a model for encoding semantic relationships between items of data so that these relationships can be interpreted computationally.

More than 60% of the world's working population is employed in the informal economy [6]. Everyday a massive amount of work is being handled by these informal employees and quantum amount of data is being generated by them to run their businesses smoothly. The entire data in this unorganized sector is in an unstructured format. There is no standardized platform for these informal employees to establish a wider scope for their business.

In the era of the semantic web, where all data are interconnected, knowledge sharing and communication are paramount in any organization but, as most organizations grow and constantly gather latest information, this becomes a major struggle.

Many businesses are operated by the unorganized or marginal sector. To date, no platform has been developed that can connect marginal business owners or hawkers and unorganized business executors with investors or an idea-generator. Anyone wanting to reach out to the small businessmen of the unorganized sector with the motive to help them either with funds or ideas needs to be able to connect with them directly. But there is no organized platform to establish this direct connection. There is a nominal scope of customer repeatability for these marginal traders. There is also very less market opportunity at the bottom of the pyramid. But that can be created if these vendors of the informal sector can be enabled with an orchestrator for their business to grow.

The purpose is to connect the idea givers and investors with the people running a marginal business. Another requirement is to increase the reachability of these hawkers, costermongers and pavement vendors to the right customers and vice versa.

2. Related Work

This section highlights some of the key contributions made by the semantic web community to understand Web of Needs technology for matchmaking service along with issues and methods to support the work presented here. The Web of Needs (WoN) is an approach for communicating and publishing human needs on the online-marketplace as linked data to allow automatic matching of the expressed needs by building connection between users who expressed the needs. The authors, in this paper proposes that finding matches for a certain need is a critical feature of the Web of Needs, and the system's utility is heavily reliant on the quality of the matches [7].

The Web of Needs vary from traditional e-marketplaces, in that it does not build numerous vertical marketplaces for different specialties; instead, it develops a one unified Web marketplace. They have been focused on motivating the Web of needs as a generic cooperation framework that can serve as a basis of a worldwide on-line marketplace [8].

In another paper [9], the authors implemented Web of Needs node that would store needs and matching service would crawl the Web of needs, which is the part of the linked open data graph located on WoN nodes. Whenever they find needs that satisfy each other's matching criteria, the respective need owners are given the appropriate information. We are trying to implement a similar type of concept on matching service in our use case. We were inspired by this paper [10], where a chat-bot connects users of the RDF-based WoN network with an existing web service. In this paper, the contributors designed an agreement protocol on top of the federated and completely RDF-based chat protocol used by WoN users to communicate in order to allow users from various domains or platforms to come to a mutually accepted agreement about how to interact with each other. This work [11] substantiates that when needs are generated, they are anonymously made available in public to be matched with other needs at the nodes on the web. Agents with similar requirements are connected through need matching.

Another work [12] was done on ontology matching between job recruiter and job searcher by creating human resource ontology from integrating existing widespread used standards and classifications. With the help of semantic similarity, job description and applicant profiles are compared, instead of mere keyword matching.

[13] proposed matching algorithm that is used to categorize the matches found between resource and job and create ranking for the best possible match that will connect a resource available with the job. We propose to implement a similar approach using weighted score method.

A Web services framework called Autonomous Matchmaking Web Services (AMWS) was introduced in another paper [14] that uses RDF messages as a communication protocol among services, requester, and the registry; to achieve dynamic service discovery, assembling and invocation in a large-scale, distributed environment. A more recent paper [15] shows an AMWS framework conforming to semantic message-based Web service is wrapping around existing web services that includes redesigning Web service to enable exchange of RDF messages. In our work, we propose to implement storage of data as RDF input and output.

3. Scope of Work

With the inception of the semantic web, different domains see prospective growth in terms of user accessibility and platform independence. Data is not siloed in one system under one business head, instead, it is made interoperable. Such technology is the need for elevating the marginal business runners. Existing research shows that matchmaking services can be implemented using the Web of Need technology, where the data is represented in the form of URIs.

Most e-Commerce businesses do not allow transactions across their boundaries [8]. Our attempt is to make a decentralized marketplace by connecting Investors to marginal business runners through a matchmaking service that would suffice the need of both the participants.

3.1. Contributions

- Create a platform to consolidate the Hawker related data.

- Convert the Hawker data into an RDF Triple format to support Web of Needs accessibility.
- Create a Score based Matching service to enable the Investors for searching investment opportunities without having to go through all the available requirements.
- Organize the data in a way that interoperability is possible in the future.

4. Problem Statement

This research attempts to focus on the marginal business sectors that generate huge revenue in a country's economy. From a social perspective, a large portion of the marginal business sector includes the hawkers who do not have the proper capital to build a stable business. The social status of hawkers is very low as they do not have a steady income due to their unsteady business. If business by the marginal sector can be monetized, then that could become a significant revenue stream for a country's economy. With the aid of the Web of Needs (WoN), we are publishing the Hawker needs-related data which is accessible to the Investors.

There is a lack of hawker data available on the web which means investors are not knowing the scope of investment in the business and its growth. From an Investor's perspective, with such a WoN based platform available to them, they can easily make investment decisions by analyzing the data. This WoN would also open gateways to bi-directional growth opportunities for both the Hawkers and Investors.

5. Architecture

The proposed model would basically accept inputs from both Hawkers and Investors in their respective input pages. The Hawker Input page would accept the hawker details and store them as an RDF Graph element in the Neo4j graph database. Whereas the Investor inputs would be accepted by the system as search criteria. Once there is a match is achieved, the Investor data would then be stored as a node element in accordance with the Investor investing in a Hawker. Both systems would create and access the Web of Need (WoN) to find a suitable match. The matching service algorithm mentioned in section 3.2 would explain the logic behind the score-based matching utilized in this scenario.

Figure 1 explains how the Hawker input data is converted into an RDF graph and stored into Neo4J. The system is designed to convert the user input of data from Hawkers into RDF data. This RDF data would feed into the Web of Needs that would connect to the matching service to find a suitable match as per the requirement criteria of both parties - Hawkers and Investors. The hawker endpoint is used to enter the data through the input portal or excel spreadsheets. The data entered would be converted into URI bindings in the format of RDF triples, subject, predicate, and object. The triple store entered is stored in Pandas data-frame of Python framework which is compiled to create the RDF Graph in Neo4j Aura database.

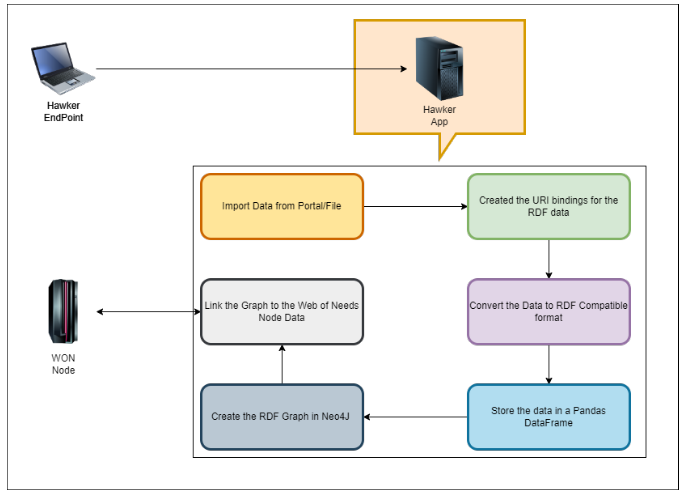

Figure 1. Application data flow for the conversion of Hawker information to RDF data.

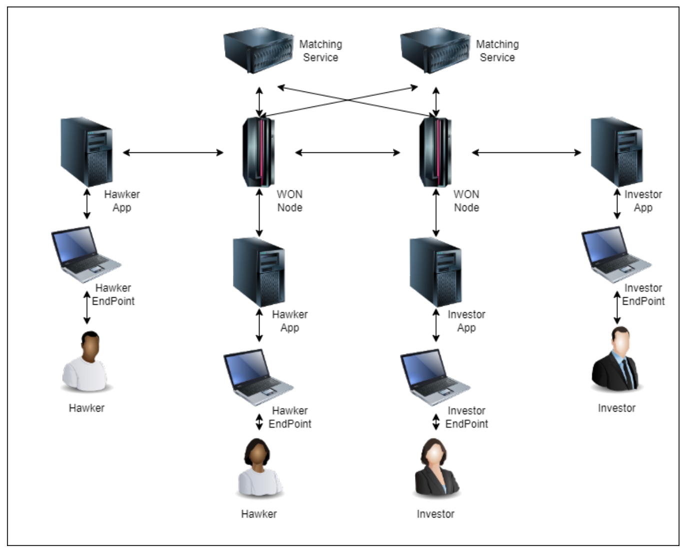

Figure 2. Flow of information between multiple systems through a Web of Need node. Each node is connected to either an Investor app or our Hawker app.

Figure 2 explains how the WoN created by the Hawker application interacts with the WoN created by the Investor apps to achieve a matching. The matching service is running in this case between the different WoNs to perform a score-based matching.

Figure 3 explains the schema of the Hawker data that is utilized to prepare the RFD Graph and stored into Neo4J.

5.1. RDF Schema

```
@prefix rdf:<http://www.w3org/1999/02/22
    -rdf-syntax-ns#>
@prefix rdfs:<http://www.w3org/2000/01/
```

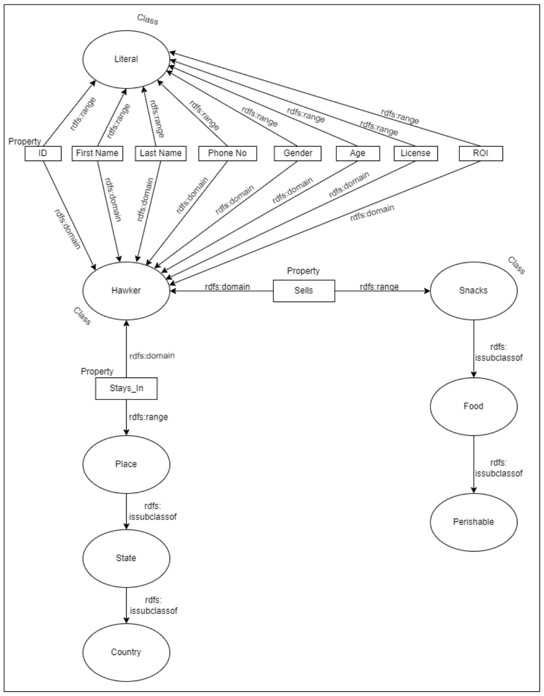

Figure 3. RDF Schema of Hawker Data.

```
   rdf-schema#>
: Place a rdfs:Class;
   rdfs:subClassOf:State.
: State a rdfs:Class;
   rdfs:subClassOf:Country.
: Country a rdfs:Class.
: Literal a rdfs:Class.
: Hawker a rdfs:Class.
: Snacks a rdfs:Class;
   rdfs:subClassOf:Food.
: Food a rdfs:Class;
   rdfs:subClassOf:Perishable.
: Perishable a rdfs : Class;

: Stays_In a rdf : Property;
   rdfs : domain : Place;
   rdfs : range : Hawker.
: Sells a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Snacks.
: Id a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal
: FirstName a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: MiddleName a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: LastName a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: Phn_No a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: Gender a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: Age a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: LicenseStatus a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: LicenseNo a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.
: ROI a rdf : Property;
   rdfs : domain : Hawker;
   rdfs : range : Literal.

: HAW00001 a : Hawker.
: Snacks a : Food.
: HAW00001 : sells : Snacks.

: HAW00001  a : Hawker.
: Kolkata  a : Place.
: HAW00001 : Stays_In : Kolkata.
```

5.2. Matching Algorithm

This algorithm uses a weighted average scoring mechanism to identify a suitable match between Investors and Hawkers. Each of the matching parameters has its individual scores which are then multiplied by their respective weights. Each Hawker is matched to the Investor and a score is assigned to the match. In our implementation we have used these scoring parameter, however a different set of scoring parameter can also be used. Following table displays the Weights of each scoring parameters:

Parameters	Weights	Weighing Factor
Investment Amount Score Factor	20%	0.2
Product Interest Score Factor	10%	0.1
Location Score Factor	10%	0.1
Education Score Factor	10%	0.1
License Score Factor	15%	0.15
Union Score Factor	10%	0.1
Work Location Score Factor	10%	0.1
Contact Info Score Factor	15%	0.15

TABLE 1. WEIGHTS AGAINST SCORING PARAMETERS

Each matching parameter assigns a score of the range of 0 to 1, 1 being the best match and 0 being no match. Each of the matching score is then multiplied with their respective weights. e.g. i_1 is the matching score for parameter 1 and w_1 is the weight of the parameter 1. Hence,

$$Parameter\_Score = i_1 * w_1$$

Once all the parameters are assigned with their respective Parameter_Score (p_n) then we will use the following formula to get the Matching Score:

$$Matching\_Score = \sum_{1}^{n}(p_1, p_2, .., p_n)/n$$

Any match that attains a 75% or higher percentile is considered a positive match. Following algorithm uses some of the parameters such as: Hawker_Requirement (HR) is matched with Investor_Investing (II), Hawker_Requirement_Amount (HRA) should be less than Investment_Amount (IA), hence assigning a score of 1 if they match or else 0 if they do not match. R and AMT are the two scoring parameters being used in this algorithm. 'n' number of parameters can be matched using this similar algorithm.

Algorithm 1: Outer Structure of the Algorithm

Input: HR, II, HRA, IA
Output: R,AMT

---

if $HR = II$ **then**
  R=1
  if $HR = Money$ **then**
    if $HRA < IA$ **then**
      $AMT = 1$
    else
      if $HRA = IA$ **then**
        $AMT = 0.5$
      else
        $AMT = 0$
      end
    end
  else
    $AMT = 0$
  end
else
  $R = 0$
end

6. Use Case

This paper focuses on a use case of upliftment of the hawker system around the world. There is a huge untapped market of hawkers and their needs to flourish. In this paper, we are proposing to create a Web of Needs specifically for upliftment of hawkers and their aspirations for a better livelihood. Investors can easily access this data from their respective systems as this Web of Needs is designed to be Semantically connected to the World Wide Web (WWW). We are using technologies such as RDF and Graph Database to enrich and store the data for creating a Semantic Web of Needs.

7. Implementation

Hawker Growth Initiative (HGI) has been implemented in Python using different existing frameworks like Pandas, Numpy, Py2neo, RDFPandas, and RDFlib. For the visualization of graphs, Neo4j AuraDB [16] which is a fully managed cloud service graph database for cloud developers is used. To read semantic service descriptions, we employ RDF and convert Hawker data into URIs. Data is first captured in a graph database and then converted into an RDF model, which is subsequently processed using the Python framework.

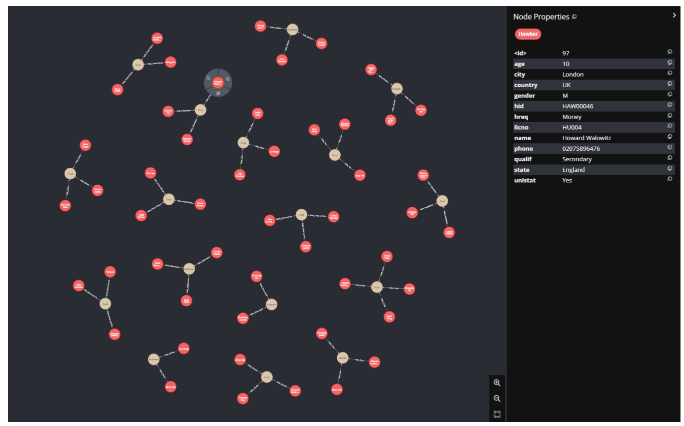

Figure 4. Graph created from RDBMS data of Hawkers.

Initially, we concentrated on capturing the data from an RDBMS [17] data model and importing them to a graph database (Neo4j). Then this data is converted to an RDF data model using Python RDFLib and RDFPandas modules. In Figure 4, we show the graph representation of Hawker data, where the nodes contain all attributes of the Hawkers. We also created some sample disjoint nodes for Investors. Once the data was ready for processing, we implemented our Matching Algorithm on the data to find out suitable matches between the Hawker Nodes and Investor Nodes based on the Matching Score. Once a match is made, we created a relationship between the respective Hawker and Investor as Investor—> invests_in —> Hawker as shown in Figure 5.

In our case study, the RDF Hawker data becomes a part of the Web of Needs [18] for Hawkers. This enables the Investors to search for investment opportunities based on the Matching Score provided from the Multidisciplinary Expertise Matching [19] to them in accordance with the respective Hawkers.

8. Conclusion and Future Scope

In this paper, we propose to create a Web of Needs node and a way to store Hawker data in RDF triples format in

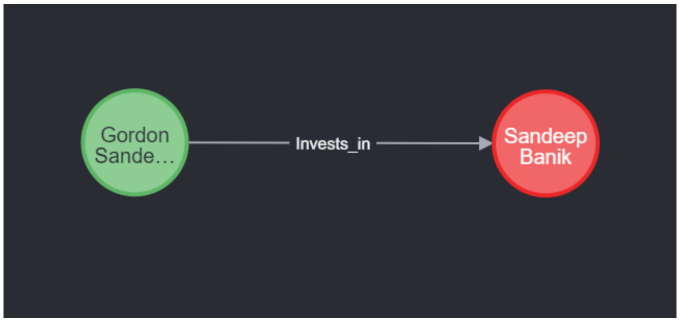

Figure 5. Relationship created between the Investor and the Hawker.

that Web of Needs. In the current data landscape, there is no one place where all the Hawker data can be found in a consolidated way. In our research, we make an attempt to consolidate data from the entire Hawker landscape, into a Web of Needs and make them available as interoperable data. We are also proposing a matching service that will utilize a weighted scoring mechanism to provide Investors with the most suitable choices for an investment opportunity. This system will also keep track of the Investors as they select the Hawkers to make an investment.

The platform can be further developed to incorporate the following functionalities:

- The application can be further developed to be interoperable with other systems.
- The RDF triple data can be linked to the UDDI (Universal Description, Discovery, and Integration) framework [20] for higher visibility.

References

[1] C. Bizer, T. Heath, and T. Berners-Lee, "Linked data: The story so far," in *Semantic services, interoperability and web applications: emerging concepts*. IGI global, 2011, pp. 205–227.

[2] W. Ali, M. Saleem, B. Yao, A. Hogan, and A.-C. N. Ngomo, "A survey of rdf stores & sparql engines for querying knowledge graphs," *The VLDB Journal*, pp. 1–26, 2021.

[3] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic web," *Scientific american*, vol. 284, no. 5, pp. 34–43, 2001.

[4] M. C. Daconta, L. J. Obrst, and K. T. Smith, *The Semantic Web: a guide to the future of XML, Web services, and knowledge management*. John Wiley & Sons, 2003.

[5] D. Trastour, C. Bartolini, and J. Gonzalez-Castillo, *A semantic web approach to service description for matchmaking of services*. Hewlett-Packard Laboratories, 2001.

[6] "2022 theme: Achieving social justice through formal employment," https://www.un.org/en/observances/social-justice-day, accessed: 2022-05-31.

[7] F. Kleedorfer, S. Human, H. Friedrich, and C. Huemer, "Web of needs: A process overview." in *SEMANTiCS (Posters, Demos, SuCCESS)*, 2016.

[8] F. Kleedorfer, C. M. Busch, C. Pichler, and C. Huemer, "The case for the web of needs," in *2014 IEEE 16th Conference on Business Informatics*, vol. 1. IEEE, 2014, pp. 94–101.

[9] F. Kleedorfer and C. M. Busch, "Beyond data: Building a web of needs," in *LDOW*, 2013.

[10] F. Kleedorfer, F. Suda, M. Stolze, and C. Huemer, "Hailing a taxi on the web of needs." in *SEMANTICS Posters&Demos*, 2018.

[11] H. Friedrich, F. Kleedorfer, S. Human, and C. Huemer, "Integrating matching services into the web of needs." in *SEMANTiCS (Posters, Demos, SuCCESS)*, 2016.

[12] C. Bizer, R. Heese, M. Mochol, R. Oldakowski, R. Tolksdorf, and R. Eckstein, "The impact of semantic web technologies on job recruitment processes," in *Wirtschaftsinformatik 2005*. Springer, 2005, pp. 1367–1381.

[13] V. N. Narayanan and S. Kailash, "Resource matchmaking in grid-semantically," in *The 9th International Conference on Advanced Communication Technology*, vol. 3. IEEE, 2007, pp. 2051–2055.

[14] H. Q. Yu, S. Dietze, and N. Benn, "Autonomous matchmaking web services," in *2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM)*. IEEE, 2010, pp. 420–425.

[15] H. Q. Yu, D. Liu, S. Dietze, and J. Domingue, "Developing rdf-based web services for supporting runtime matchmaking and invocation," in *2011 7th International Conference on Next Generation Web Services Practices*. IEEE, 2011, pp. 392–397.

[16] J. Barrasa, A. E. Hodler, and J. Webber, "Knowledge graphs: Data in context for responsive businesses," 2021.

[17] S. Albahli and A. Melton, "Rdf data management: A survey of rdbms-based approaches," in *Proceedings of the 6th International Conference on Web Intelligence, Mining and Semantics*, 2016, pp. 1–4.

[18] F. Kleedorfer and C. Huemer, "Towards a web based transportation infrastructure," in *Digitalization in Supply Chain Management and Logistics: Smart and Digital Solutions for an Industry 4.0 Environment. Proceedings of the Hamburg International Conference of Logistics (HICL), Vol. 23*. Berlin: epubli GmbH, 2017, pp. 55–73.

[19] P. Liu, J. Curson, and P. Dew, "Use of rdf for expertise matching within academia," *Knowledge and Information Systems*, vol. 8, no. 1, pp. 103–130, 2005.

[20] L. Li and I. Horrocks, "A software framework for matchmaking based on semantic web technology," *International Journal of Electronic Commerce*, vol. 8, no. 4, pp. 39–60, 2004.