ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR316

A Knowledg Graph Model for e-Government

Friday Orji[1]; Nuka Nwiabu[2]; Okoni Bennett[3]; Onate Taylor[4]
[1,2,3,4]Department of Computer Science
Rivers State University
Port Harcourt, Nigeria

Abstract:- Many governments around the world have invested huge amount of resource to build their e-Government capabilities, to meet government objectives of effective public service delivery and citizens engagement. The increase in size of an e-Government landscape has led to the increase in complexity of the infrastructure. This increasing complex infrastructure presents a challenge for governments to continue to meet its objectives. Knowledge Graph (KG), a constituent AI technology, has shown a lot of promise in helping governments meet its objectives in the midst of the complexity. A major aspect of this complexity is the need to maintain a single view of the world, in the form of a unified meaning of data, within a given e-Government instance, given the heterogeneity in data models used in the different departments within an e-Government instance. In this paper, we present a unique perspective in addressing the problem of deriving semantic meaning from disparate data in an e-Government context, using KG. Our aim is to advance the objectives of effective service delivery and citizens engagement in a complex e-Government instance. We focus on creating a data-centric architectural model that is single, simple and extensible, based on KG. We create a functional model based on architectural view and viewpoints from standards such as The Open Group Architectural Framework (TOGAF). The functional model highlights the various components that underpin the functions. We have developed our model within the context of a Design Science Research (DSR) approach, and we provide evaluation of same model within that context. An e-Government KG model guides the development of KG solutions in e-Government, in order to achieve the e-Government enterprise goals of effective service delivery and citizens engagement.

Keywords:- *Knowledge Graph, E-Government, Ontology, RDF, AI, OBDA, Architecture, Model, TOGAF, Data.*

I. INTRODUCTION

In recent years, the field of e-Government has seen significant growth, with governments around the world investing heavily in the development of digital platforms and services to provide efficient and effective public services. As e-Government services continues to grow in complexity, there is need for more advanced technologies to support their development [1]. Artificial Intelligence (AI) generally and Knowledge Graph (KG) in particular has shown a lot of promise in addressing the complexities problems and the computational requirement of e-Government [2], [3], [4], [5]. e-Government is the use of Information and Communication Technology (ICT) in the context of public policy, operations in public organization, citizens engagement, and government services [6]. The e-Government domain is large, heterogenous, dynamic, and shared, with different semantic world view [2]. This characterization of the domain is unique and provides an ideal context to explore the synergies between different AI technologies. Two main factors are driving the adoption and growth of e-Government around the world. One is the need by government to improve services and revenue, while reducing cost; and the second is advances made in the areas of AI in particular, and ICT in general. A Knowledge Graph (KG) is a graph representation of real-world entities – i.e., objects, events, situation, or concepts, and the relationships between these entities [7].

A major objective of e-Government initiatives is the delivery of services, and citizens engagement, through a One-Stop-Shop approach [8]. A One-Stop-Shop is a public service delivery model which entails having a single access point to information and services, for citizens and businesses[8], with the One-Stop-Shop being used as a government's front office for frontline services. The adoption of this public service delivery model around the world, is driven by its huge success in the private sector. Organizing government services and citizens engagement as a One-Stop-Shop, requires integration and interoperability across existing government departments, and devising a framework for seamlessly integrating future departments. One dimension of the complexity problem in e-Government is the heterogeneity of the landscape in terms of the data models used in the various departments of government, resulting in each department having a view of the world that is different from the others [9]. This is key because, in an environment where several services are delivered using different technologies, the lack of a common view of the world makes it challenging to integrate them seamlessly and ensure their interoperability. This problem arises from the mismatch between the data generated from a government's operations and engagement with its citizens, and any knowledge implicit in the data. In others words, so much data in an e-Government environment, but little or no knowledge from that data.

II. RELATED WORK

Data integration problem has been examined from a more theoretical perspective, where two major approaches have been identified [10] – global-as-view and local-as-view. The global-as-view is an approach that requires the global

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR316

schema be expressed in terms of the data sources. On the other hand, the local-as-view approach is where the global schema is specified independently from the sources and the relationship between the global schema and the sources are established by defining every source as a view over the global schema. This work identified three main components of a data integration system – global schema, sources and mappings. While [10] took a more theoretical approach to the data integration problem, [11] examined it from a generic approach in identifying and including components for semantic processing of the data integration tasks, in their Ontology Based Data Access (OBDA) system approach. In this approach, data access is mediated by domain-oriented conceptual models in the form of ontologies. This creates a semantic layer on top of the data access layer for the different data sources, and provides a unified global view of the data sources. The components in [11] model include an ontology, a dataset, a semantic services component and a mapping. The approach here is generic, in the sense that it is not a specific solution for data integration, but rather a generic description of semantic based data integration system. An overview of OBDA paradigm, as a means of semantic data integration, was provided by [12], highlighting the techniques that form the basis of the paradigm and the challenges that remain to be addressed. Some of these challenges are query rewriting optimization, metamodeling and meta-querying, non-relational data sources, OBDA methodology and tools and OBDA evolution. Optique [13] describes empirical, case-study, and real-world development and use of an OBDA system in an enterprise environment. In this work, the authors developed techniques to address some identified limitations of existing OBDA systems, namely how to create ontologies and mappings for a deployment of an OBDA system, how to ensure that OBDA query processing is efficient in practice, how to ensure that the target users are actually able to efficiently express their information needs against an OBDA system. MASTRO [14] is a Java-based OBDA tool that allows reasoning based on DL-lite and can be connected to a federated database containing multiple, disparate databases. Morph-RDB [15] is an extension of the query translation mechanism used for RDB2RDF mapping language, based on the Morph engine.

None of the approaches, to the best of our knowledge, specifically addressed the architectural modeling of an OBDA system in e-Government context, with a view to integrating it other sub-system of the e-Government landscape.

III. DOMAIN MODEL ANALYSIS

We examine the problems of existing solutions and the requirements of a solution that addresses these problems. We do this in the following sections.

A. Overview

The main problem with the vast majority of existing approach in government data management, as in many enterprise domains, is that there is a gap between the data and the meaning of the data. The problem arises and is more acute, mainly because of the inherent nature of governments, which is large, heterogenous, dynamic, shared, with different semantic interpretation of data. The core idea of our proposed approach is to create a data-centric architectural model that is single, simple, and extensible, based on Knowledge Graphs (KG), which addresses the data-semantic mismatch inherent in e-Government systems.

Creating a KG model is essentially an attempt at solving the problem of integration data from multiple sources with different schemas and datasets.

B. A Government Case Study

Rivers State Government is a subnational government in Nigeria, which is currently making huge investment in its e-Government infrastructure, including investment in advanced AI technologies to solve real-world problems arising from the increased complexities of its e-Government landscape. One major problem faced by this government is encapsulated by the problem of having a single, unified view of a citizen's activities – Furo Konyaa, who buys and registers a land with the land registry at the Ministry of Lands & Survey (Government Department A), and makes payment to the Ministry of Finance (Government Department B). Furo goes ahead to obtain a planning permission from the Ministry of Physical Planning & Urban Development (Government Department C) to develop a property on the land. Even though each of these activities -land registry, payment, and planning permit, exist in the KG of each of the individual ministries, only the combined KG has the knowledge of the whole scenario, as shown in Figure 1

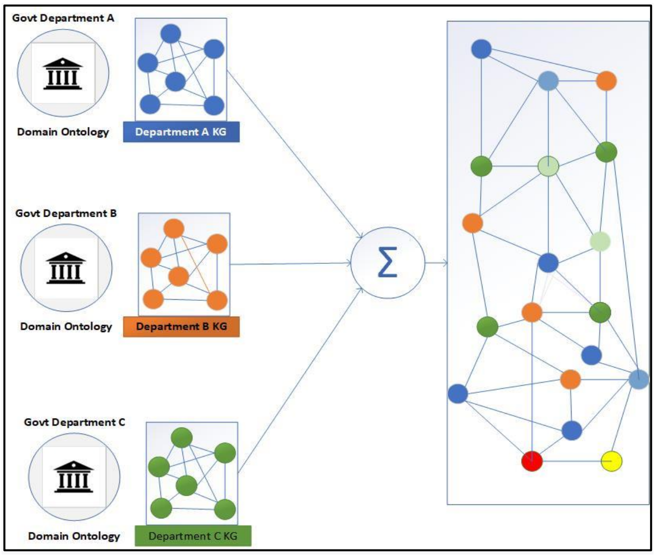


Fig 1: KG Combination Model in Rivers State Government Domain

C. The Proposed Model

High-Level architectural reference model of a new approach for government data management based on KG is shown in Figure 2. The three main components of the reference model are Knowledge Acquisition and Integration, Knowledge Storage, and Knowledge Consumption components. Each of these main components consists of many tasks, which can be performed using various approaches. The Knowledge Acquisition and Integration

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR316

component contains ontology development, data transformation, data annotation and quality assurance tasks. Tasks such as entity summary, graph summary, semantic search, query generation and question answering make up the Knowledge Consumption component. The Knowledge Storage component holds the materialized and persistent KG. The other components in the reference model are the data source component that provide the data for the transformation, and the users and application that access and consume the KG.

Each of these components is referenced in the specific instantiation in our proposed model - GOvernment Knowledge grAph DAta Model (GOKADAM) that follows from the high-level reference architectural model. GOKADAM is a functional view (Open Group Architecture Framework (TOGAF) and ISO's Reference Model of Open Distributed Processing (RM-ODP)) [16], [17] based on the viewpoint of the model designer, and its development is based on the analysis of the viewpoint of the existing solution artifacts as follows:

- Analysis of viewpoints of existing solution models
- Analysis of viewpoint of EA frameworks
- Synthesis of both viewpoints in (1) and (2) above, into a single viewpoint

The development of the reference model is based on:

- Analysis of the commonalities of existing models
- Analysis of reference models entities in EA frameworks
- Synthesis of both results from (1) and (2) above.

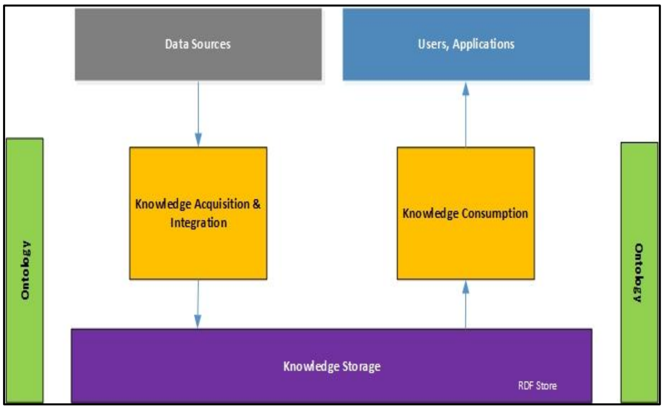


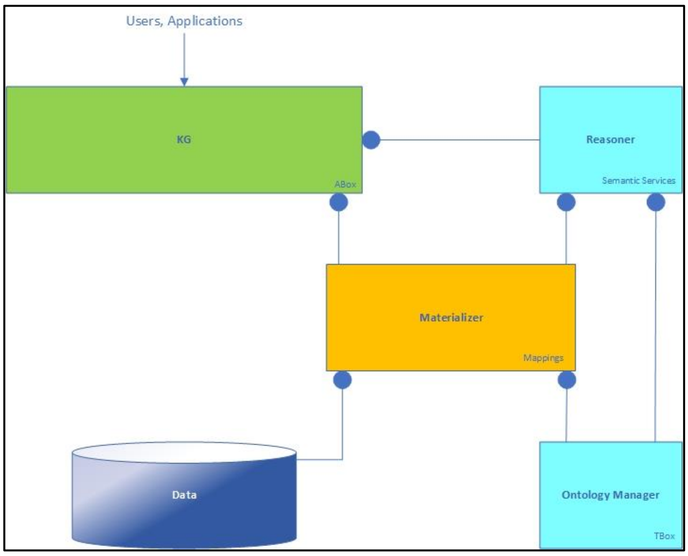
Fig 2: High-Level Architectural Reference Model of the Proposed Approach for Government Data Management Based on KGs

IV. DOMAIN ARCHITECTURE MODEL

This section is based on the synoptic review and analysis of specific existing models (sections 2 and 3) with a view to arriving at a new model that addresses the problem identified of the existing models. As part of this assessment, the requirements and components of the new model are identified. Existing models are examined with a view to determining if there are any attributes or components of existing models that could aid in the development of a new model, and that help to address the identified problem.

Based on the foregoing analysis, the following components have been identified as the necessary components for our proposed model – Government Object Based Data Access Model (GOKADAM) for the case study, as shown in Figure 3.

- e-Government global ontology manager
- A reasoner
- A materializer
- e-Government Data Sources
- e-Government KG



A. *e-Government Global Ontology Manager*

An ontology manager in this e-Government model is the TBox [18] and the component that stores the conceptual models for the top-level view, and the department and application views. It keeps the models up to date, makes use of an approximator to accommodate input ontologies with unsupported axioms, by reducing its expressivity.

B. *Reasoner*

A reasoner is a semantic service which implements system logic including basic reasoning tasks such as subsumption computation built on top of the TBox. The reasoner component also includes a query reformulator subcomponent that reformulate conjunctive queries and forwards them to the ABox [18] for execution.

C. *Materializer*

A materializer component makes use of the data transform mappings to extract and periodically update the ontology instances. A materialized ABox approach differs from a Virtual ABox approach, in the sense that the ontology instances are kept with the system in a triple store. On the contrary, a virtual ABox retrieves ontology instances directly from the data sources upon each request. We have opted for the materialized ABox because it provides more performance efficiency in the automatic reasoning that is normally done in a KG, compared to the virtual ABox.

D. e-Government Data Source

The e-Government data source component keeps data in several forms – relational, document repositories, NOSQL stores, and other formats. The focus of our model, and the mappings provided, is on relational data.

E. e-Government KG

The KG component is the materialized ABox in the model, and ABox is realized by an RDF based triple store. Like other DBMS, a triple store allows storing and querying of triples. In addition to other data management facility provided by a DBMS, a triple store supports inference computations, which are implemented based on formal semantics.

V. EVALUATION AND DISCUSSION

The GOKADAM model is an architectural model showing the various components indicating the computational viewpoint, which visually describes the distribution of the system through functional decomposition into computational objects, and also showing the interaction of these components at the interface level. This view describes the functionalities provided by the system using an architectural description framework. We have opted to use the TOGAF and the architectural method and language specified by ISO's RM-ODP. Using and architectural modeling approach that emphasizes a functional and computational viewpoint helps to achieve the goal of characterizing the difference in approach between GOKADAM and existing approaches, as opposed to a system development centric approach such as the Unified Modeling Language (UML) [19], which is better suited for software-intensive system development purposes, where emphasis is placed on the Software Development Life Cycle (SDLC) process.

GOKADAM is a KG-based data integration model in e-Government. It was developed following a Design Science Research (DSR) approach, EA architecture principles, ontology development methodology, and data transformation component development. The GOKADAM model was also developed within the context of a reference model for KG construction, storage, and consumption in e-Government. GOKADAM, and its associated ontology, are a solution that addresses the problem of KG-based data access in our e-Government context, and it is the outcome of model development process that attempts to address, incrementally, the issues raised about the short-comings of similar existing models, following a DSR approach. Evaluation of an artifact is an essential activity within DSR, to demonstrate its "utility, quality and efficacy" [20]. Various evaluation methods are available within the DSR research approach [20], [21], and the specific evaluation methods used depends on the design artifact and the selected evaluation metrics. At this stage of our work, we have opted for descriptive evaluation method, as more suitable approach validation process, and our task is to use informed argument and information from the knowledge base to build a convincing argument for the artifact's utility. The goal of GOKADAM is to provide a model to guide the development of KG solutions in e-Government. An evaluation of GOKADAM is done against this goal in various respect. EA principles suggest evaluating artifacts against criteria such as improving enterprise goals, improve communication between different stakeholder groups, help toward integration and interoperability, and reduce complexity [16], [22], [23], [24]. GOKADAM is simple in its design to aid communication between the modeler and stakeholder groups up and down the development cycle. This in turn helps to develop the KG system which satisfies the integration and interoperability objective. This aligns with the enterprise goals in our context.

VI. CONCLUSION

In this paper, we examined a proposal for Knowledge Graph based data model for e-Government. This proposal has an actual architectural model and a reference model on which the architectural model is based. We used Enterprise Architecture principles encapsulated in industry standard architecture standards, such as TOGAF. The DSR approach used in developing the model, provides a framework used to evaluate our work. The model development approach and the models itself helps to achieve the e-Government enterprise goals of effective service delivery and citizens engagement.

As future work, we plan to implement the functional components in our architecture, using dataset from our specific e-Government context, as a way of empirical validation of our model.

REFERENCES

[1]. A. Al-Besher and K. Kumar, "Use of artificial intelligence to enhance e-government services," *Meas. Sens.*, vol. 24, p. 100484, 2022.
[2]. A. Adadi, M. Berrada, D. Chenouni, and B. Bounabat, "Ontology based composition of e-Government services using AI Planning," in *2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA)*, Rabat: IEEE, Oct. 2015, pp. 1–8.
[3]. A. Adadi, M. Berrada, and N. El Akkad, "Artificial Intelligence based Composition for E-Government Services," in *Proceedings of the Third International Conference on Computing and Wireless Communication Systems, ICCWCS 2019, April 24-25, 2019, Faculty of Sciences, Ibn Tofaïl University -Kénitra- Morocco*, Kenitra, Morocco: EAI, 2019.
[4]. A. Adadi, M. Berrada, D. Chenouni, and B. Bounabat, "A SEMANTIC WEB SERVICE COMPOSITION FOR E- GOVERNMENT SERVICES," . *Vol.*, p. 8, 2015.
[5]. O. S. Al-Mushayt, "Automating E-Government Services With Artificial Intelligence," *IEEE Access*, vol. 7, pp. 146821–146829, 2019
[6]. IEEE Computer Society e-Gov STC, "About - IEEE Computer Society e-Government STC," About - IEEE Computer Society e-Government STC. Accessed: Nov. 25, 2022. [Online]. Available: https://sites.google.com/a/ieee.net/stc-egov/about
[7]. D. Fensel *et al.*, *Knowledge graphs*. Springer, 2020.

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/IJISRT24APR316

[8]. OECD, *One-Stop Shops for Citizens and Business*. in OECD Best Practice Principles for Regulatory Policy. OECD, 2020.

[9]. E. Hovy, "Data and knowledge integration for e-government," in *Digital government: E-Government research, case studies, and implementation*, Springer, 2008, pp. 219–231.

[10]. M. Lenzerini, "Data integration: A theoretical perspective," in *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, 2002, pp. 233–246.

[11]. J. Z. Pan, G. Vetere, J. M. Gomez-Perez, and H. Wu, Eds., *Exploiting Linked Data and Knowledge Graphs in Large Organisations*. Cham: Springer International Publishing, 2017.

[12]. G. De Giacomo, D. Lembo, M. Lenzerini, A. Poggi, and R. Rosati, "Using Ontologies for Semantic Data Integration," in *A Comprehensive Guide Through the Italian Database Research Over the Last 25 Years*, vol. 31, S. Flesca, S. Greco, E. Masciari, and D. Saccà, Eds., in Studies in Big Data, vol. 31. , Cham: Springer International Publishing, 2018, pp. 187–202.

[13]. E. Kharlamov *et al.*, "Ontology based data access in Statoil," *J. Web Semant.*, vol. 44, pp. 3–36, 2017.

[14]. D. Calvanese *et al.*, "The MASTRO system for ontology-based data access," *Semantic Web*, vol. 2, no. 1, pp. 43–53, 2011.

[15]. F. Priyatna, O. Corcho, and J. Sequeda, "Formalisation and experiences of R2RML-based SPARQL to SQL query translation using morph," in *Proceedings of the 23rd international conference on World wide web*, 2014, pp. 479–490.

[16]. V. Haren, *TOGAF Version 9.1*. Van Haren Publishing, 2011.

[17]. K. Raymond, "Reference model of open distributed processing (RM-ODP): Introduction," in *Open Distributed Processing: Experiences with distributed environments. Proceedings of the third IFIP TC 6/WG 6.1 international conference on open distributed processing, 1994*, Springer, 1995, pp. 3–14.

[18]. G. De Giacomo and M. Lenzerini, "TBox and ABox reasoning in expressive description logics," *KR*, vol. 96, pp. 316–327, 1996.

[19]. M. Fowler, *UML distilled: a brief guide to the standard object modeling language*. Addison-Wesley Professional, 2018.

[20]. Hevner, March, Park, and Ram, "Design Science in Information Systems Research," *MIS Q.*, vol. 28, no. 1, p. 75, 2004.

[21]. K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007.

[22]. F. S. De Boer, M. M. Bonsangue, L. P. J. Groenewegen, A. W. Stam, S. Stevens, and L. Van Der Torre, "Change impact analysis of enterprise architectures," in *IRI - 2005 IEEE International Conference on Information Reuse and Integration, Conf, 2005.*, Las Vegas, NV, USA: IEEE, 2005, pp. 177–181.

[23]. J. M. Morganwalp and A. P. Sage, "Enterprise architecture measures of effectiveness," *Int. J. Technol. Policy Manag.*, vol. 4, no. 1, p. 81, 2004.

[24]. J. Schekkerman, *The economic benefits of enterprise architecture: how to quantify and manage the economic value of enterprise architecture*. Trafford, 2005.