HOSTED BY unisel UNIVERSITI SELANGOR

Selangor Science &Technology Review
Special Issue: Science and Technology for Society
Vol 5 No 5 (2021)

Proposed Search of Intelligent Information System for Disease Diagnosis Using Semantic Web

S.R.Candra Nursari[1*], Setyawan Widyarto[2], Haslinda Sutan Ahmad Nawi[3]

[1]Departement of Informatics, Faculty of Engineering, Pancasila University, Jakarta Indonesia
sri.rezeki.candra.n@univpancasila.ac.id
[2]Faculty of Communication, Visual Art and Computing, Universiti Selangor (UNISEL), Selangor, Malaysia
swidyarto@unisel.edu.my
[3]Faculty of Communication, Visual Art and Computing, Universiti Selangor (UNISEL), Selangor, Malaysia
haslindasan@unisel.edu.my

Abstract: The development of Information Technology is not only on hardware and software but also in the development of technology in the field of medicine, based on AI. In the field of medicine, AI is developing rapidly. Information and Communication Technology in the health field is widely used, one of which is disease information data. Disease information data availability has increased electronically but the data is not categorized and stored semantically. No categorization and semantic store make the data difficult to find. Semantic Web is an intelligent service as an information intermediary, search agent, and information filter, which offers more functionality and interoperability than a standalone service. The Semantic Web context is meta-data that allows the machine to interpret it where the query execution time depends on this order. A good algorithm for determining query paths can thus contribute to making queries fast and efficient. Methods and analysis will use Ontology Web Language for disease domain modelling and meta-search systems for ontology mapping and Web services. The Mapping component includes domain ontologies, taxonomic information and collection databases and changes them in the Resource Description Framework. The study proposed a research direction to create semantic web ontology models and optimize disease information search using genetic algorithms that allow automatic meta-data matching. The ultimate purpose is a repository of knowledge related to such a mapping that disease information search can be found. A new algorithm or model will be proposed and it can optimize disease finding with the information needed and semantic Web ontology.

Keywords: Ontology Web Language, Genetic Algorithms, Disease Information

1. Introduction

The development of Information Technology is not only on hardware and software but also in the development of technology-based on Artificial Intelligent (AI) (Albabtain et al. 2014). In the field of medicine, AI is developing rapidly. Information and Communication Technology (ICT) in the health field is widely used, one of which is disease information data (He et al. 2019). Health data is an important part of health services. Data must be collected, stored, and must be easily accessed by all stakeholders in the form of Health Warehouses as a means of scientific communication (Braa, Monteiro, and Sahay 2004).

Disease information data availability has increased electronically but the data is not categorized and stored semantically (Paper et al. 2015) (Braa, Monteiro, and Sahay 2004). No categorization and semantic store make the data difficult to find. Semantic Web is an intelligent service as an information intermediary, search agent, and information filter, which offers more functionality and interoperability than a standalone service. The Semantic Web context is meta-data that allows the machine to interpret it where the query execution time depends on this order. A good algorithm for determining query paths can thus contribute to making queries fast and efficient.

In a disease information retrieval system, all disease information that is on the website can be read by the user via the internet anywhere and anytime. By using this system, users

can search for disease information anywhere and anytime. On average, a web that informs about a disease its content is not geographically relevant to local users or the web page has little or no authority if the user wants to get more detailed information.

This article is organized as follows. Section 2 discusses the intelligent analysis of information and web semantic and also presents previous studies/research on intelligent analysis of information. Finally, section 3 concludes the proposed research direction.

2. Intelligent Search Tool for Disease Information

2.1 Definition

The internet is a huge source of information, users can find about anything through a search engine. Most of the information on the internet is in the form of text. Text mining is a process of getting quality information from text or web documents so that extracting information can be focused and useful. (Xie et al. 2020). Health Information Systems still have weaknesses in the presence of data that is not connected/decentralized to each other. This is caused by the difficulty of exchanging data. There are new ways to exchange data. This can be done by allowing data to be accessed directly by other people. The new technology proposed is the Semantic Web which aims to share health data. The official definition of the Semantic Web World Wide Web Consortium (W3C) is a representation of data on the World Wide Web (www). Semantic Web does not replace current Web technology but is complementary. Semantic Web is the latest Web technology today. Where Semantic Web technology produces information that can be read and processed by machines (Decker et al. 2000) (W3C 2014) it's smarter than the Web today. So, using the Semantic Web, users can collect data provided by others and use it as desired. The Semantic Web can be used to share data, including health data. Each side only needs to create a Semantic website and connect each other. The Semantic Web consists of three layers: XML, RDF, RDF Schema, and Ontology

Ontology is a theory of the meaning of an object, the properties of an object and the relation of objects that might occur in a domain of knowledge (Ehrig 2007). Ontology is a model that represents declarative knowledge and provides a knowledge base that can be understood by machines, making it possible as machine intelligence. To automate its workflow, an ontology framework is needed that can later explain health/disease information and function as a knowledge base (Dang et al. 2008). The ontology-based approach can be a means to extract textual data semantics and enrich the data semantics with certain semantic conceptual representations of entities to obtain domain knowledge in the ontology. (Pornpit Wongthontham 2018).

2.2 Propose Methodology

As visualized in Figure 1, the analysis will use Ontology Web Language (OWL) for disease domain modelling and meta-search systems for ontology mapping and Web services. The Mapping component includes domain ontologies, taxonomic information and collection databases and changes them in the Resource Description Framework (RDF). Next, semantic web ontology models and optimize disease information search using genetic algorithms that allow automatic meta-data matching will be created. Using genetic algorithms for

implementing a good data scheme and metasearch will get optimal results in the search for disease information

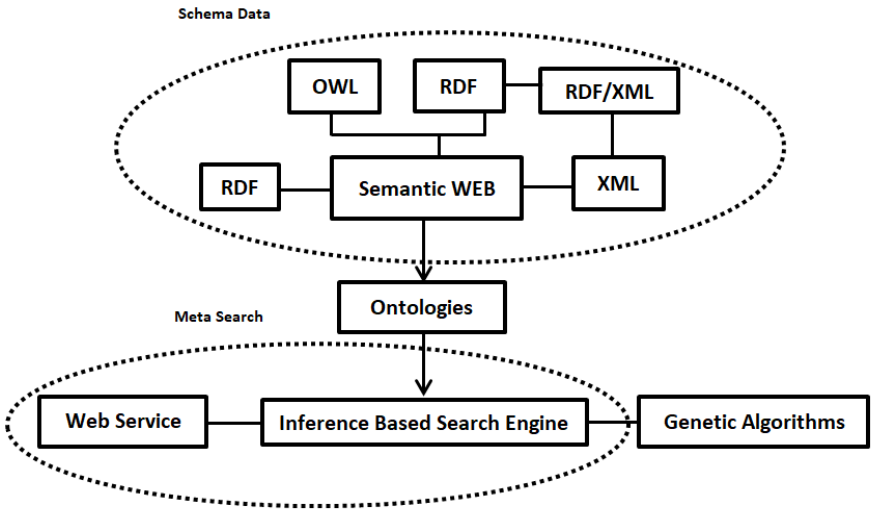

Fig.1 Method Meta Data Matching

The results of the research is a repository of knowledge related to such a mapping that disease information search can be found.

2.3 Previous Research

Table 1a. Summary the Previous Researches Showing Search Analytic Use of Genetic Algorithms

Approach/ Model/ Author	Brief Description	Search Analysis		Use of Ontology	Applied in Genetic Algorithms
		Entity Level	Domain-Level		
MLS (Aasiya Parveen 2019)	Improve prediction by using a genetic algorithm with MLS1 and MLS2 cases	No	Yes	No	Yes
MLSGA (Jennings et al. 2019)	Improve predictions for pattern classification and financial modelling using the MLSGA methodology to find optimal solutions	NO	Yes	No	Yes

Table 1a shows two papers summary discussed Multi-Level Selection (MLS) and Multi-Level Selection Genetic Algorithm (MLSGA). Both papers applied GA but it did not use ontology and non entity level.

Table 1b. Summary the Previous Researches With Semantic Topic

TOPIC	A : A_Title	B : B_Author	C : C_Year	D : D_Issue_Problem	E : E_Methods	F : F_Analysis&Results
Semantic	THE SEMANTIC WEB: The Roles of XML and RDF	STEFAN DECKER AND SERGEY MELNIK Stanford University FRANK	2018	we explain the role of ontologies in the architecture of	XML and RDF are the current standards for establishing semantic	The Web community currently regards XML as the most important
	Hybrid approach for big data localization and semantic	Waheed Yousuf Ramay1	2018	this issue in the web semantic web advances is assuming a remarkable	Relational Database (RDB) and RDF data models are becoming a	A system like a web needs updates to become capable of using any
	The Need for Semantic Web Service in the eHealth	Emanuele Della Valle	2018	The current trend in dealing with maintenance of an application	COCOON is a 6th Framework EU integrated project aimed at setting	The main lesson-learned in applying WSMO in the healthcare
	SemAxis: A Lightweight Framework to Characterize Domain-Specific	Jisun An† Haewoon Kwak† Yong-Yeol Ahn§	2018	the idea of defining a semantic axis and assessing the meaning of a	we propose SEMAXIS, a simple yet powerful framework to characterize	We have proposed SEMAXIS to examine a nuanced representation
	Direct Network Transfer: Transfer Learning of Sentence Embeddings	Li Zhang Steven R. Wilson University of Michigan	2018	we propose a transfer learning setting specialized for semantic	we compare several approaches to transfer sentence encoders to	We choose 3 sentence encoders, 8 semantic similarity tasks and 29
	SemAxis: A Lightweight Framework to Characterize Domain-Specific	Jisun An† Haewoon Kwak† Yong-Yeol Ahn§	2018	we propose SEMAXIS, a lightweight framework to characterize	Our framework, SEMAXIS, involves three steps: constructing a word	We have proposed SEMAXIS to examine a nuanced representation
	A Typedriven Vector Semantics for Ellipsis with Anaphora using Lambek	Gijs Wijnholds1 and Mehrnoosh Sadrzadeh1	2019	An explanation of the derivational processes resulting in these	Distributional semantics has a lot to say about the statistical colocation-	In this paper we incorporated a proper notion of copying into a
	Semantic Annotation of Arabic Web Documents using Deep Learning	Saeed Albukhitan, Ahmed Alnazer, Tarek Helmy	2018	investigates the feasibility of using deep learning, an emerging area of	related Web documents. For a given set of Arabic documents and	This paper demonstrates the benefits of using terms from the
	Semantic Web in data mining and knowledge discovery: A	Petar Ristoski *, Heiko Paulheim	2016	this survey is to give a survey on the field as broad	many approaches have been proposed in this area that combine	In this paper, we have given a survey on the usage of Semantic
	Semantic Interference and Facilitation: Understanding the	Ernesto Guerra1 * and Pia Knoeferle2,3,4	2018	We embrace the idea that the level of activation of perceptually and	inspired by conceptual metaphors such as 'similarity is closeness'	In two eye-tracking experiments we intended to further understand
	Estimates of incidence and mortality of cervical cancer in 2018: a	Marc Arbyn, Elisabete Weiderpass, Laia Bruni, Silvia de Sanjosé, Mona	2020	The knowledge that persistent human papillomavirus (HPV)	For this worldwide analysis, we used data of cancer estimates from 185	Approximately 570 000 cases of cervical cancer and 311 000 deaths
	Health, Food and User's Profile Ontologies for Personalized	Tarek Helmy*, Ahmed Al-Nazer, Saeed Al-Bukhitan, Ali Iqbal	2015	to make effective use for the annotation, ontologies concept	to achieve high relevancy and coverage we need to use ontologies	This paper presents food, health, nutrition and the user's profile

Table 1b shows the previous researches from 2015 to 2020 with semantic topic. We analysed their results and have addressed their issue problems and we will use suitable methods to search of intelligent information system for disease diagnosis using semantic web.

Table 1c. Summary the Previous Researches With Search Optimization Topic

TOPIC	A : A_Title	B : B_Author	C : C_Year	D : D_Issue_Problem	E : E_Methods	F : F_Analysis&Results
Search Optimization	Grid Search, Random Search, Genetic Algorithm: A Big	Petro Liashchynskyi	2019	we compare the three most popular algorithms for hyperparameter	In this experiment, we set population size to 2, the number of generations	The evolutionary algorithm took about ≈4.13 hours to run. The best
	An anatomy for neural search engines	Thiago Akio Nakamura a , b , *, Pedro H. Calais b , Davi de Castro	2018	we are interested on how deep learning can improve text based	We show that a full neural architecture, which employs neural	In this work, we have demonstrated a novel way of developing a search
	Implementation of hybrid pattern search-genetic algorithm into	C. Long. Lok1 · B. Vengadaesvaran1 · S. Ramesh1	2016	non-pollution energy alternative,	pattern search (PS) algorithm and genetic algorithm (GA)	get the best optimized model
	An efficient top-k ranking method for service selection based on ε-	Wei Yu1,3 · Shijun Li1 · Xiaoyue Tang1,2	2018	the actual success of a metasearch engine directly depends on the	Multi-objective evolutionary algorithms (MOEAs) have been	provide concrete evidence that the performance and efficiency of ε-
	Modeling dynamic urban land-use change with geographical cellular	Yongjiu Feng	2017	the dynamic spatiotemporal evolution of urban land-use change and urban	GPS optimization	the GPS method produced smaller residuals than the LR method in
	Pattern Search Multidimensional Scaling	Georgios Paraskevopoulos*†1, Efthymios Tzinis*1, Emmanuel-	2019	matrix of (similarities or) dissimilarities1 of the dataset objects	the classical multi-dimensional scaling (MDS) method	noisy manifold geometry shapes indicates that pattern search MDS
	An anatomy for neural search engines	Thiago Akio Nakamura a , b , *, Pedro H. Calais b , Davi de Castro	2018	build a neural model centric search engine	neural IR models	the neural IR show more important progress on recall-oriented queries
	MetaXplorer: an intelligent and adaptable metasearch engine using	Neha Dimri1,2 • Himanshu Kaul1 • Daya Gupta1	2018	metasearch engines, which combine the result	model: MetaXplorer	MetaXplorer has the highest average precision of 0.6641 when compared
	Retinal artery/vein classification using genetic-search feature	Fan Huang a , Behdad Dashtbozorg a , Tao Tan a , c , Bart M. ter Haar	2018	exploits the connectivity of vessels to improve the classification	a genetic-search based feature selection technique	an accuracy of 90.2%, the sensitivity of 89.6%, the specificity of 91.3% on
	LionRank: lion algorithm-based metasearch engines for re-ranking of	P. VIJAYA* & Satish CHANDER	2018,	very complicated for the users to know the most relevant information	lion algorithm-based metasearch engines	The maximum F-score of 80% was obtained for the proposed LionRank
	OmniSearch: a semantic search system based on the Ontology for	Jingshan Huang1*, Fernando Gutierrez2, Harrison J. Strachan1,	2016	we investigated the con-struction of an application ontology	We previously developed a miRNA domain-specific application	summarized as: (1) following a modularized ontology design (with
	Symbiotic Organisms Search Algorithm: theory, recent advances	Absalom E. Ezugwu1, Doddy Prayogo2	2018	The symbiotic organisms search algorithm is a very promising recent	despite all the remarkable achievements and rapidly expanding	this paper provides an overview of the research conducted on symbiotic
	AWeb service search engine for large-scale Web service discovery	Afnan Bukhari1 · Xumin Liu2	2018	the semantic Web services search engine	The proposed Web service search engine (WSSE) is based on the	to improve its efficiency and accuracy. The effectiveness of the
	BioTCM-SE: A Semantic Search Engine for the Information Retrieval	Xi Chen, Huajun Chen, Xuan Bi, Peiqin Gu, Jiaoyan Chen, and	2013	we present a novel semantic search engine	with the explosive growth ofbiomedical data on theWeb,	The BioTCM-SE project is still under development, but many
	PIBAS FedSPARQL: a web-based platform for integration and	Marija Djokic-Petrovic1		There are a huge variety of data sources relevant to chemical,	Running of predefined queries Information about initiatives and	PIBAS FedSPARQL is a web-based query builder and result set
	Performance-Similarity Reasoning as a Source for Mechanism Schema	Raoul Gervais1	2017	I explicate and discuss performance-similarity reasoning as a strategy for	Two types of such inferences are distinguished: one in which the	As I said in the introduction, the aim of this paper is to enlarge the stock of
	Introducing Complexity to Formal Testing	Ismael Rodríguez, Fernando Rosa-Velardo, and Fernando Rubio	2019	we will consider systems modelled accord-	We use this theory to analyze the √ testing complexity of some general	Inthispaperwehaveintroduced thenotionofcomplex-
	Annotation based automatic action processing	Elias K¨arle and Dieter Fensel	2018	we present the idea of using schema.org	The general idea is to collect the structured data which is available on	This work-in-progress paper describes a way to collect
	Why do we remember? The	Johannes B. Mahr	2018	In the present article, we argue that	We provide an account of the	A domain where such claims are

Table 1c shows the previous researches from 2013 to 2019 with search optimization topic. We analysed their results and have addressed their issue problems and we found GA could be used to search information intelligently for disease diagnosis using semantic web.

3. Conclusion

The creation of semantic web ontology models that optimize disease information search using genetic algorithms that allow automatic meta-data matching will ease the search

of disease information in knowledge repository. Apart from that a new algorithm will also be proposed to optimize the disease finding with the information needed and semantic Web ontology. This will support trusted interactions over the network that enable the health domain stakeholders to create data stores on the Web and organize this enrich data with additional meaning.

4. Acknowledgement

-

5. References

A.J. Sobeya, P.A. Grudniewski. 2018. “Re-inspiring the genetic algorithm with multi-level selection theory: Multi-Level Selection Genetic Algorithm (MLSGA)” *Bioinspir. Biomim. in press https://doi.org/10.1088/1748-3190/aad2e8*.

Albabtain, Arwa Fahad, Dina Ahmad Almulhim, Faisel Yunus, and Mowafa Said Househ. 2014. “The Role of Mobile Health in the Developing World : A Review of Current Knowledge and Future Trends.” *Journal of Selected Areas in Health Informatics* 4(2): 10–15.

Braa, Monteiro, and Sahay. 2004. “Networks of Action: Sustainable Health Information Systems across Developing Countries.” *MIS Quarterly* 28(3): 337. http://www.jstor.org/stable/10.2307/25148643.

Dang, Jiangbo, Amir Hedayati, Ken Hampel, and Candemir Toklu. 2008. “An Ontological Knowledge Framework for Adaptive Medical Workflow.” *Journal of Biomedical Informatics* 41(5): 829–36.

Decker, Stefan et al. 2000. “The Semantic Web: The Roles of XML and RDF.” *IEEE Internet Computing* 4(October): 63–74.

Ehrig, Marc. 2007. 4 *Ontology Alignment - Bridging the Semantic Gap*. http://link.springer.com/10.1007/978-0-387-36501-5.

He, Jianxing et al. 2019. “The Practical Implementation of Artificial Intelligence Technologies in Medicine.” *Nature Medicine* 25(1): 30–36. http://dx.doi.org/10.1038/s41591-018-0307-0.

Jennings, Paul C. et al. 2019. “Genetic Algorithms for Computational Materials Discovery Accelerated by Machine Learning.” *NJ Computational Materials* 5(1). http://dx.doi.org/10.1038/s41524-019-0181-4.

Paper, Discussion et al. 2015. *Information And Communication Technologies For Health Systems*. http://akb.africa-union.org/auc/handle/AKB/12191.

Pornpit Wongthontham, and Bilal Abu-Salih. 2018. “Ontology-Based Approach for Identifying the Credibility Domain in Social Big Data.” : 1–43.

W3C. 2014. “RDF Vocabulary Description Language 1.1:RDF Schema.” https://www.w3.org/TR/2014/REC-rdf-schema-20140225/.

Xie, Xia et al. 2020. “A Novel Text Mining Approach for Scholar Information Extraction from Web Content in Chinese.” *Future Generation Computer Systems* 111: 859–72.