

A Prototype of a Semantic Platform with a Speech Recognition System for Visual Impaired People

Jimmy Rosales-Huamaní¹, José Castillo-Sequera², Fabricio Puente-Mansilla¹,
Gustavo Boza-Quispe¹

¹National University of Engineering, Lima, Peru

²Alcala University, Madrid, Spain

Email: jrosales@uni.edu.pe, jluiscastillo@uah.es, fpuentem@uni.pe, gbozaq@uni.pe

Received 29 June 2015; accepted 26 September 2015; published 29 September 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In the world, 10% of the world population suffer with some type of disability, however the fast technological development can originate some barriers that these people have to face if they want to access to technology. This is particularly true in the case of visually impaired users, as they require special assistance when they use any computer system and also depend on the audio for navigation tasks. Therefore, this paper is focused on making a prototype of a semantic platform with web accessibility for blind people. We propose a method to interaction with user through voice commands, allowing the direct communication with the platform. The proposed platform will be implemented using Semantic Web tools, because we intend to facilitate the search and retrieval of information in a more efficient way and offer a personalized learning. Also, Google APIs (STT (Speech to Text) and TTS (Text to Speech)) and Raspberry Pi board will be integrated in a speech recognition module.

Keywords

Semantic Web, Ontology, ASR, Raspberry Pi

1. Introduction

In recent years the size of the World Wide Web (WWW) has grown dramatically; this has led to a considerable increase in the difficulty to find data about a particular issue, due to the ambiguity of terms used to make queries on the web.

How to cite this paper: Rosales-Huamaní, J., Castillo-Sequera, J., Puente-Mansilla, F. and Boza-Quispe, G. (2015) A Prototype of a Semantic Platform with a Speech Recognition System for Visual Impaired People. *Journal of Intelligent Learning Systems and Applications*, 7, 87-92. <http://dx.doi.org/10.4236/jilsa.2015.74008>

The Semantic Web, also known as the Data Web and Web 3.0 [1] pretend to solve this problem creating a mechanism for the exchange information with certain meaning. To provide a website with a comprehensible meaning by computers, it is necessary to have a knowledge representation. The Semantic Web proposes the use of collections of information called ontologies in order to have a structured knowledge.

By other hand, people with disabilities are nearly 10% in the world; these people have to deal with many physical barriers and some new barriers have been added: Technology. These are causing the digital gap, also known as the digital divide or digital stratification [2].

To be precise, in case of blind people, they require special help to work with any kind of computer system. Furthermore, these people needs appropriate tools to get relevant data from Internet.

From our point of view, in any case, the implementation of a web platform has to address certain problems like meaning understandable by computers, efficient retrieval of information and accessibility for everyone.

The World Wide Web Consortium (W3C) offers standards which are internationally accepted. They offer quantifiable rules, however, web developers often fail to implement them effectively. One of the reasons is that most of the available accessibility guidelines appear too costly [3].

According to [4], with the Semantic Web, there are now new opportunities to build flexible systems that will meet the needs of disabled students. Disability aware systems could be designed using Semantic Web technologies, leading to personalized environments that will enable disabled students to have relevant learning resources and to work independently, with little assistance from a tutor.

For these reasons, in this paper we make a prototype of a platform based on semantic web with speech recognition using natural language allowing blind students access to knowledge resources and learn independently of their tutor. The present work is divided as follow: in Section 2 we review related papers. The next section discusses the current problems. Afterwards, in Section 4 we present our proposal, in Section 5 we present the conceptual scheme of architecture and finally in Section 6 expected contribution is presented.

2. Review of Literature

The state-of-art was based in papers related to ontology, Semantic Web and accessibility for disabled people. In this section, we review papers from 2010 to date.

In [3] was developed an ontology called CO to represent user interaction in its context and improve web accessibility for all people. In this work, four important concepts are identified: User context, physical context, environmental context and computational context.

In [5] was developed a model based on ontologies for integrating several web services and their delivery to users with reduced mobility. Under this proposal, people with disabilities are able to search on web for services.

In [6] was developed a prototype system based on an ontology for Internet information retrieval for autistic people through learning styles. In this paper, the autistic user expect to retrieve information with different characteristics, but this is difficult for autistic people, because of their lack of ability to process and retain information. In this work the aim is to find the desired result, based on a suitable set of keywords that are based on their memories of people with autism.

In [7] was proposed the use of the URC framework (The Universal Remote Console) in form of UCH-oriented Gateway (The Universal Control Hub) and some services of Ontologies, such as ODP (Platform for dialogue based on Ontologies) to provide interactive services accessible to any TV architecture. This platform is called VITAL (Vital Assistance for the Elderly) and helps to elderly people to participate in dialogue.

In [4] was developed an ontology called ADOOLES (Abilities and Disabilities Ontology for Online Learning and Services) based on personalized online instruction for disabled students in higher education. This ontology was built on ADOLENA (Abilities and Disabilities Ontology for Enhancing Accessibility) that was developed by [8].

In [9], a middleware based on ontologies for personalize tourism to people with special needs was presented. Its function was to retrieve and classify information in places adapted for people with special needs. This was supported by PATRAC (Accesible Heritage) philosophy. Furthermore, they proposed a content manager based on ontologies divided in three designed modules. The content manager uses SOAP (Simple Object Access Protocol) web services.

Independently, in [10] was developed HIV (Heavyweight ontology Based Information Extraction for Visually impaired User) that provide a mechanism for highly precise information extraction using heavyweight ontology and built-in vocal command system for visually impaired internet users.

Finally, [11] presented the research program of the London Metropolitan University, which aims to use Semantic Web and mobile Internet for care of the disabled and elderly people using intelligent agents. The authors use as a conceptual background the International Classification of functionality, Disability and Health (ICF), it has been established as a standard for the classification of the various states of health. Where the core of the framework is a pilot whose ontological domain is disability.

3. Current Problems of Web Platforms for Accessibility

A review of the literature found the following problems:

- A big part of e-learning environments are not yet accessible for all users, because there are many electronic barriers that prevent access to online resources, and it does not have access technical aids (such as the use of screen reader) [3].
- The number of students with disabilities in UK higher education institutions increases every year. Delivering education online is becoming increasingly challenging as institutions encounter some disabilities requiring adjustments of learning environments. The law requires that people with disabilities be given equivalent learning experiences to their non-disabled peers through reasonable adjustments. Educational institutions have thus utilised assistive technologies to assist disabled students in their learning, but some of these technologies are incompatible with some learning environments, hence excluding some disabled students and resulting in a disability divide [4].
- Disabled people represent an important part of our society, they have different needs based on the type of disability that they presented, and being the assistance provided to them important and the use of new technologies should accommodate their needs [11].

4. Prototype of Semantic Platform with Speech Recognition System

In order to achieve the prototype we consider the following steps:

1) Building of Ontology. For an easy and rapid extraction of information by the user, the knowledge about a particular subject or domain will be stored in an ontology. Such ontology contain terms and the relationships among these terms. Terms are often called classes, or concepts; these words are interchangeable. The relationships between these classes can be expressed by using a hierarchical structure: superclasses represent higher-level concepts and subclasses represent finer concepts, and the finer concepts have all the attributes and features that the higher concepts have. The ontology is designed in the language OWL (Ontology Web Language) that is most popular language for creating ontologies today [12].

In addition for modelling the ontology we should follow the recommendations of Methontology that is a standard created by the Ontology Engineering Group of the Polytechnic University of Madrid (UPM), which comprises the following steps: Specification, conceptualization, acquiring knowledge, integration, implementation, maintenance, evaluation, and documentation [13].

2) Semantic Platform Implementation

The implementation of the Semantic Platform will be using Jena libraries API for Java that allow the management of ontologies in OWL code and support a reasoner engine.

According to [14] Jena is a library to develop applications based on RDF (Resource Definition Framework) and OWL documents. It is only used in application code as a collection of APIs, and there is no GUI of any kind. Also provides a framework to development Semantic Web applications using Java language.

Furthermore, the platform will be based on the Model-View-Controller scheme, which implement servlets to manage user queries. Such queries can be verified in an ontological editor using SPARQL (SPARQL Protocol and RDF Query Language) [15], so as W3C recommend, SPARQL will be used to consult RDF or OWL documents.

3) Adaptation of Speech Recognition System to Semantic Platform.

Automatic speech recognition systems (ASR) compared to other human-machine interaction systems like keyboard, mouse, etc. provide better naturalness. Speech recognition seems so natural and simple for people but for machines is quite complicated. For this reason, a recognition of patterns is used, these patterns are a set of linguistic units as (words, syllables, sounds, shapes).

There are studies that use queries patterns using SPARQL [16]. We propose a mechanism to facilitate the

interaction between the user and the platform through an interface that use natural language. We pretend to create a module that translates user queries in natural language to SPARQL queries. To solve this problem we will build a series of templates in SPARQL.

In our platform, a template consists of a SPARQL representation, which reflects the internal structure of the question from natural language. This module would be integrated into the server and also linked to the ASR system. A user with visual disabilities requires an accessible medium to interact with the platform. To achieve this, the following processes are required.

- a) Voice-to-text. The process used for adapt the voice recognition system is the ASR (Automatic Speech Recognition).

In [17], we can see a considerable number ASR systems, some open source, other privative and even based in cloud services. In the last case, we focus specifically in API Google STT Service, because it is a cloud computing system and does not compromise the performance of the local computer.

According to [18], thanks to the processing in the cloud, the ASR can be used in devices that have no high-performance processor and avoid the use of complex algorithms. In [19] was presented a system using API Google STT that operate together with reduced board computer Raspberry Pi, with excellent results (85% - 90% accuracy). The advantage of using this Raspberry Pi is its small size that makes ideal to be embedded anywhere, without saturating the available space. By other hand, the low cost of equipment enables this massive distribution.

- b) Text-to-Speech.

To convert Text to Speech (TTS) we propose to use the cloud services of Google, using Google Translator, although is a less complex process compared with ASR. Other possibility could be run locally on your Raspberry Pi, obviously with a reduced quality of voice synthesis.

5. Conceptual Scheme of Architecture

The conceptual scheme is roughed out in Figure 1 and explained as follow.

- 1) Information Request Module (IRM).

This module allows the user to makes consult through the keyboard and computer screen, this information is sent to a KRM. Once information is processed, the requested are shown to the user in the Prototype Portal.

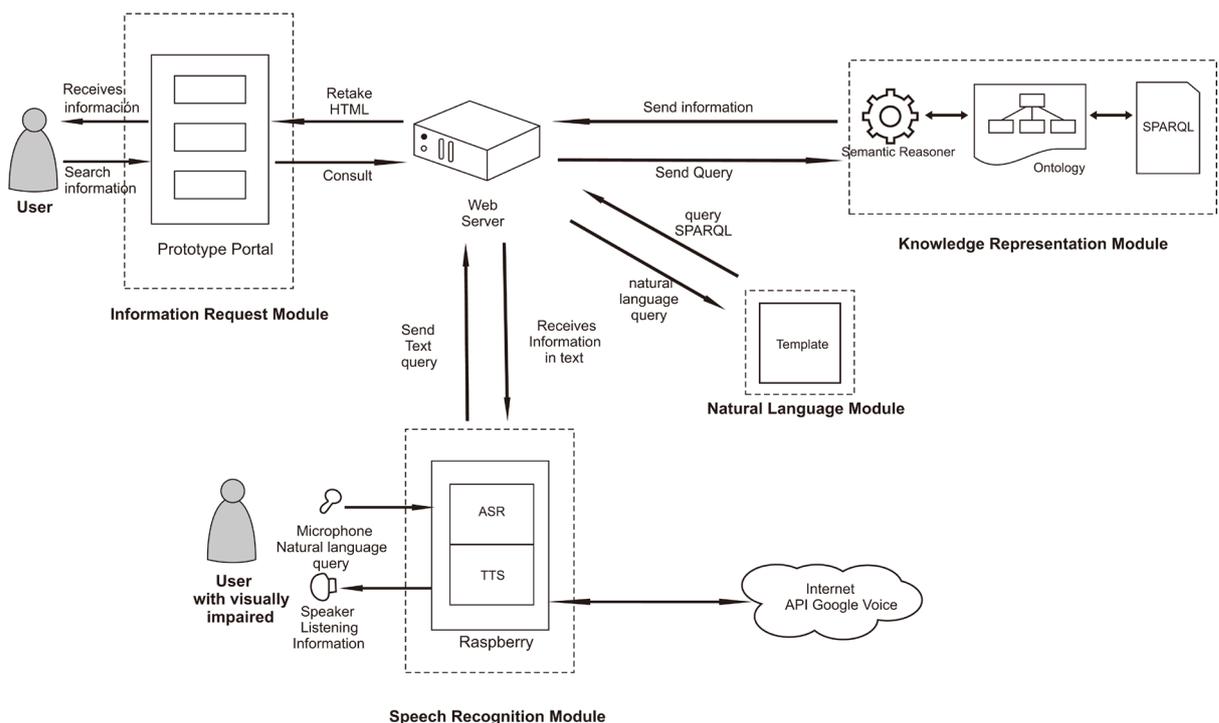


Figure 1. Conceptual scheme of architecture.

2) Knowledge Representation Module (KRM). The module consists of an ontology that contains all the concept models of study case, and is used to describe and represent a specific area of knowledge such as medicine, tourism, film, etc.

The ontology provides a way to encode knowledge and semantics such that machines can understand. Also, the ontological model contain logical rules. These rules depend on the selected domain, and it will need to be implemented as part of ontological model properties.

Depending on the source of the query, this module can interact with the other modules including IRM and NLM. In our prototype the SPARQL language will be used to extract different information from the ontology and answer the queries made by different modules.

3) Speech Recognition Module (SRM). This module allows to visually impaired people make queries in our Semantic Platform. For instance, the user ask for information in the microphone and such consult is converted to text in the ASR system. Then, this text is sent to the NLM to identify the appropriate template. After identifying the appropriate template, in KRM we make the SPARQL query to get the required information. This information is returned to the SRM in order to be converted to speech using TTS (Text-to-Speech). Finally, information is read for the user.

4) Natural Language Module (NLM). This module analyses and compare text queries come from SRM with certain patterns to create queries in SPARQL language and then it will be sent to KRM.

Although the process of formulating a natural language query and transform to SPARQL language is quite complex, it is possible to perform the query using query patterns [16].

In [20], they present a new approach based on a syntactic analysis of the questions to produce a SPARQL template that reflects the internal structure of the question.

With this in mind, we propose to use a series of templates associated with natural language. To give an illustration, in the query “Tell me the name of all transport to the city of Lima”, template natural language query would be: “Tell me the name of all transport to the city of %PLACE%” where %PLACE% is a variable that the system recognize and is associated in the following SPARQL template:

```

1 SELECT?name WHERE
2 {
3 ?transport table:name?name.
4 ?transport table:transport_offered_destination?transport_offered_destination.
5 Filter (?transport_offered_destination = %PLACE%)
6 }

```

The system will replace the variable %PLACE% with *Lima* and send SPARQL query resultant to the KRM, where the information initially is requested, will be obtained, finally this information will be read by the SRM.

6. Expected Contributions and Future Work

Actually, there are studies about the accessibility to internet for disabled people using tools of the Semantic Web. The reason of our proposal is the lack of works that use natural language to achieve the integration of semantics and ASR systems as a tool for people with visual disabilities.

The implementation must choose a particular domain and such ontology must contain all knowledge of this specific domain.

Once implemented the prototype of Semantic Platform with ASR system, the following benefits are expected:

- The integration of semantics and ASR system using natural language for assisting visually impaired people through the web.
- The developed platform architecture based on the prototype.
- The high accuracy with the use of the Raspberry Pi in ASR implementation.
- Generation of different templates SPARQL depending on the domain ontology for natural language queries.
- The expected SPARQL query templates will be quite similar to queries in the day-to-day language.
- As a complement and from the implemented prototype, a navigation support systems using voice commands, can be created to help people with visual disabilities, providing autonomy in its movement in an unfamiliar environment.
- The construction of a more flexible and personalized platform using tools of the Semantic Web.

As a future work, we expect to develop a prototype in the domain of tourism. The prototype will help people with visual disabilities to obtain information about several tourist attractions in Peru.

References

- [1] Berners-Lee, T., Hendler, J., Lassila, O., *et al.* (2001) The Semantic Web. *Scientific American*, **284**, 28-37.
- [2] Garcia-Crespo, A., Ruiz-Mezcua, B., Gonzalez-Carrasco, I., Lopez-Cuadrado, J., Hernandez, Z., Barahona, R. and Holnes de Toppin, L. (2014) Accessibility Services and Interactive Digital Television: An Opportunity to Reduce the Digital Gap. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, **9**, 8-16.
- [3] Zakraoui, J. and Zagler, W. (2010) An Ontology for Representing Context in User Interaction for Enhancing Web Accessibility for All. <http://pf-mh.uvt.rnu.tn/56/>
- [4] Nganji, J.T., Brayshaw, M. and Tompsett, B. (2011) Ontology-Based E-Learning Personalization for Disabled Students in Higher Education. *Innovation in Teaching and Learning in Information and Computer Sciences*, **10**, 1-11.
- [5] Kehagias, D. and Tzovaras, D. (2010) An Ontology-Based Framework for Web Service Integration and Delivery to Mobility Impaired Users. In: Lytras, M., Ordonez De Pablos, P., Ziderman, A., Roulstone, A., Maurer, H. and Imber, J., Eds., *Knowledge Management, Information Systems, E-Learning, and Sustainability Research*, Springer Berlin Heidelberg, Berlin, 555-563.
- [6] Gupta, S. and Garg, D. (2011) Ontology Based Information Retrieval for Learning Styles of Autistic People. In: Mantri, A., Nandi, S., Kumar, G., and Kumar, S., Eds., *High Performance Architecture and Grid Computing*, Volume 169 of Communications in Computer and Information Science, Springer Berlin Heidelberg, Berlin, 293-298. http://dx.doi.org/10.1007/978-3-642-22577-2_40
- [7] Epelde, G., Valencia, X., Abascal, J., Diaz, U., Zinnikus, I. and Husodo-Schulz, C. (2011) Tv as a Human Interface for Ambient Intelligence Environments. 2011 *IEEE International Conference on Multimedia and Expo (ICME)*, Barcelona, 11-15 July 2011, 1-6. <http://dx.doi.org/10.1109/ICME.2011.6012186>
- [8] Keet, C.M., Alberts, R., Gerber, A. and Chimamiwa, G. (2008) Enhancing Web Portals with Ontology-Based Data Access: The Case Study of South Africa's Accessibility Portal for People with Disabilities. *OWLED*, **432**. <http://ceur-ws.org/Vol-432/>
- [9] Alonso, K., Aginako, N., Lozano, J. and Olaizola, I. (2012) Ontology Based Middleware for Ranking and Retrieving Information on Locations Adapted for People with Special Needs. In: Miesenberger, K., Karshmer, A., Penaz, P. and Zagler, W., Eds., *Computers Helping People with Special Needs*, Volume 7382 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, Berlin, 351-354. http://dx.doi.org/10.1007/978-3-642-31522-0_53
- [10] Bukhari, A. and Kim, Y.-G. (2012) Ontology-Assisted Automatic Precise Information Extractor for Visually Impaired Inhabitants. *Artificial Intelligence Review*, **38**, 9-24.
- [11] Vassilev, V., Ulman, M., Ouazzane, K., Kazemian, H., Aigbodi, M. and Boyd, R. (2013) Ontocarer: An Ontological Framework for Assistive Agents for the Disabled. *The 3rd International Conference on Digital Information Processing and Communications (ICDIPC2013)*, Dubai, 30 January-1 February 2013, 404-416.
- [12] Yu, L. (2007) Introduction to the Semantic Web and Semantic Web Services. CRC Press, Boca Raton. <http://dx.doi.org/10.1201/9781584889342>
- [13] Fernández-López, M., Gómez-Pérez, A. and Juristo, N. (1997) Methontology: From Ontological Art towards Ontological Engineering. <http://oa.upm.es/id/file/88930>
- [14] Yu, L. (2011) A Developers Guide to the Semantic Web. Springer Science & Business Media, Berlin. <http://dx.doi.org/10.1007/978-3-642-15970-1>
- [15] Antoniou, G. and Van Harmelen, F. (2004) A Semantic Web Primer. MIT Press, Cambridge, USA.
- [16] Pradel, C., Haemmerl, O. and Hernandez, N. (2012) A Semantic Web Interface Using Patterns: The Swip System. In: Croitoru, M., Rudolph, S., Wilson, N., Howse, J. and Corby, O., Eds., *Graph Structures for Knowledge Representation and Reasoning*, Springer Berlin Heidelberg, Barcelona, 172-187. http://dx.doi.org/10.1007/978-3-642-29449-5_7
- [17] Duarte, T., Prikładnicki, R., Calefato, F. and Lanubile, F. (2014) Speech Recognition for Voice-Based Machine Translation. *IEEE Software*, **31**, 26-31.
- [18] Stefanovic, M., Cetic, N., Kovacevic, M., Kovacevic, J. and Jankovic, M. (2012) Voice Control System with Advanced Recognition. *20th Telecommunications Forum (TELFOR)*, Belgrade, 20-22 November 2012, 1601-1604. <http://dx.doi.org/10.1109/TELFOR.2012.6419529>
- [19] Naeem, A., Qadir, A. and Safdar, W. (2014) Voice Controlled Intelligent Wheelchair Using Raspberry Pi. *International Journal of Technology and Research*, **2**, 65.
- [20] Unger, C., Böhman, L., Lehmann, J., Ngonga Ngomo, A.-C., Gerber, D. and Cimian, P. (2012) Template-Based Question Answering over RDF Data. *Proceedings of the 21st International Conference on World Wide Web*, New York, 639-648.