Development of a Knowledge Base for smart screening of language disorders in primary care

Martín-Ruiz, M.L., Valero Duboy, M.A., Pau de la Cruz. I.

Ferrando Lucas, M.T., Peñafiel Puerto, M.

Abstract— **Monitoring of neuro-evolutive development from birth until the age of six is a decisive factor in a child's quality of life. Early detection of development disorders in early childhood can facilitate necessary diagnosis and/or treatment. Primary-care pediatricians play a key role in early detection of development alterations as they can undertake the preventive and therapeutic actions necessary in the interest of a child's optimal development. The focus of this research paper is the construction of a Knowledge Base for smart screening aimed to assist pediatricians in processes of early referral in language disorders. The proposed model provides health professionals with a decision-making tool that supports referral processes. In this way, essential diagnostic and/or therapeutic actions are triggered for a comprehensive individual development. The resulting system was developed on the basis of an analysis and verification of 21 cases of children with language disorders.**

I. INTRODUCTION

Pediatricians working in the Spanish healthcare system are primary care physicians who provide care in health centers to children between the ages of 0 and 14. As stated in the white paper on early attention [1], pediatricians must screen and correctly refer children with potential disorders towards diagnostic and monitoring centers. Present rates of detection of development disorders are lower than their real incidence [2], which means that early identification of children with such disorders remains a pending task. In certain cases, language disorders are the first symptoms a child will manifest of a possible development disorder [3]. In clinical practice, pediatricians do not have enough time to perform proper screening of the children's neuro-evolutive development, and their background in disabilities is not always as complete as it would be desirable [4]. For these reasons, it is considered to be particularly valuable for a pediatrician to have access to a Knowledge Based System (KBS) in primary care to enable efficient screening of such disorders. A smart screening system in such a healthcare scenario does require the construction of a substantial Knowledge Base (KB) to test for effective approaches to this problem. This research paper describes the process of building such a KB based on multi-disciplinary work with experts and the analysis of 21 actual proven cases.

Section 2 provides details on the background of this research from two complementary perspectives: the viewpoint of experts in the detection of neurological disorders in childhood, and application of expert systems in medicine. The next section explains the methodology used in the development of the KB. A description is provided of the process of system knowledge eduction with KBS tools, the procedures for extracting information used by experts who were part of the multidisciplinary team that developed it and the stages of construction of the ontology necessary for the reasoning model. Section 4 presents the main results of the research. Section 5 presents the principal conclusions of the work.

II. BACKGROUND

A. *Early detection systems of neurological disorders in children*

The white paper on early attention emphasizes the importance of intervention for the transitory nature of a certain development disorder [1]. Such disorders must be considered a significant deviation from the "course" of development resulting from health or relational events that compromise biological, psychological and social development. Hence, detection of possible alterations in child development is essential in early attention, as it will enable the activation of a number of mechanisms of action. Therefore, the use of smart systems in primary care becomes of greater interest as it can allow for detection of neurological disorders in children and prevention of added pathologies, achievement of functional improvements and allow for a more adaptive adjustment between a child and his or her surroundings. Narbona highlights the fact that a neurological disorder does not manifest itself solely in a delay in the correct acquisition of speech and language, although a delay in the correct acquisition of language is the first alarm sign of a future neurological disorder [3]. The paper published by Nelson et al. explains why the correct acquisition of language is of vital importance [5], while Fejerman stresses that a complete neurological and pediatric evaluation can reveal related developmental disorders, starting with detection of a language disorder [6]. Although medical procedures are available to detect a number of neurological disorders in children [2, 7], these procedures are difficult to apply in primary pediatric care,

as many require significant time and specialized knowledge. The review in this research found no solutions that exploit the potential of Information Systems in combination with artificial intelligence to provide pediatricians with efficient and effective assistance in the early detection of these disorders.

B. *Knowledge-based systems in medicine*

Medicine is one of the fields to have benefited most from the use of computers, as a pioneer in the use of knowledge-based systems. Two landmarks that showed the validity of these technologies in medicine are ELIZA and MYCIN. The laboratories of the Massachusetts Institute of Technology (MIT) developed in the 1960's ELIZA project, simulating the behavior of a psychoanalyst. This was the first application in psychiatry that offered a "credible" answer to patients' questions by constructing generated sentences to these questions by changing certain words or phrases. In the mid-1970's, the MYCIN project emerged, and over time it became one of the most influential systems in the history of KBS [8, 9]. MYCIN was a system designed for diagnostics and therapy of infectious blood diseases.

The success of such past experiences positively contributed to the use of these technologies in medicine and healthcare. Table I contains a list of applications for decision-making support on monitoring of multiple diseases, including the name of the system, a brief description of its use and the date and country of creation [10]. As we may observe, only the LISA project and the SimulConsult tool can detect health problems in children. SimulConsult can perform neurological evaluations of children. In our study of KBS in medicine, no tool was found for the early detection of language disorders in children.

TABLE I. SUPPORTING SYSTEMS FOR DECISION-MAKING IN MEDICINE

System name	Description	Date and place
HEPAXPERT I,II,III	Analyzes and interprets tests for detecting hepatitis A, B, C and D	1991 Austria
VIE-PNN	SE for nutrition of newborn children in intensive care	1993 Austria
CEMS	Supporting system for decision-making in mental health. Can be used for diagnostics and treatment of patients and monitors and gives alerts on methods and results	1993 US
Coulter® FACULTYTM	KBS used to assist in work flow as an educational tool in hematology laboratories	1996 United Kingdom
TxDENT	Follow-up and provision of recommendations for patients undergoing odontological care	1997 Canada
RetroGram	Generates medication regimes using medical history and genetic information of patients with HIV	1999 United Kingdom
Automedon	KBS for administration of mechanical respiration in intensive care	2001 France
TherapyEdge	Graphically tracks and automatically processes information (medication, condition) of patients with HIV and chronic illnesses	2001 US
ERA	Support system for interactive decision-making for identifying patients suspected of having cancer	2001 United Kingdom
ATENIA	Control of hypertension in primary care, offering recommendations for care and medication	2002 US
LISA	Assists in decision-making for children with lymphoblastic leukemia	2004 United Kingdom
SimulConsult	Support software for medical decision making allowing professionals to combine clinical and laboratory conclusions, allowing for identification of useful conclusions for a diagnosis	2008 US

III. METHODOLOGY

The acquisition and systematization of needed knowledge required for the early referral KBS of a smart system proposed in this paper is a critical aspect that determines its effective use in primary care. The process of Knowledge Acquisition (KA) is the first step for creating a KBS and it strongly influences the conditions for correct operation. The methodology for KA requires consideration of both the definition of the knowledge to be systematized and the conceptualization and formal design of the information compiled from human and materials sources in order to model the functioning of the KBS. For this reason, the main methodologies available for extracting knowledge were studied, with a comparison of GROVER, CommonKADS (CK), Methontology and IDEAL. The conclusion of this study was to use a combination of CK and Methontology, as these methods offer the greatest potential for application in certain phases of the KBS construction. The relation between the KA collected from humans' expertise and its translation into a formal ontology is successfully achieved in a cyclic way.

CommonKADS is a knowledge engineering methodology for the design and development of KBS based on knowledge extracted from human experts and its codification to allow for its processing by a system [9]. The application of CK for the system design provided a set of early detection items to be considered by the pediatrician. This structured knowledge reflects all important aspects of the KBS to be implemented and verified through a user tool. CK was used in KA meetings as it is most suitable for modeling the knowledge extracted from language specialists in the form of ontology.

Methontology is a methodology oriented towards the implementation of an ontology in the activity of conceptualization of the KBS that has been successfully used by many authors [11]. Methontology defines a set of tasks that

enabled moving from an informal specification of the domain of application, collected with the language specialists, to a semi-formal specification of the domain. This facility makes easier the understanding of the ontology for smart screening of language disorders by consulted neuropediatricians and language therapists as well as the system developer.

The following team of experts from the field of healthcare participated in the construction of the KB needed to enable knowledge inference in the KBS to perform early screening of language disorders in primary care: a neonatologist with high expertise in development disorders and child disability, former director of the neonatology department of San Carlos Hospital in Madrid, two primary care pediatricians, a neuropediatrician presently working in the Quirón Hospital of Madrid and two experts in specific language impairment who are therapists at the Language Intervention Center (LIC) at La Salle Campus (UAM) of Madrid. Acquisition and formalization processes were developed on the basis of information gathered in open meetings and then structured with the team of experts. The process of eduction of expert knowledge relied on the use of additional techniques such as questionnaires, surveys and interviews designed in accordance with the objectives to be met by the smart system, with a view to support pediatricians in primary care working in the public health system.

The first group of experts – a pediatrician, a neuropediatrician and a neonatologist (experts 1) – held five open meetings between September 2009 and May 2011 in order to fully define the problem to be solved, work on early screening of language disorders, with the construction of a smart detection system. The second group of experts – a neuropediatrician and two therapists specialized in neurological disorders in children (experts 2) – created the knowledge base (KB) through ten structured meetings. The verification of the KB was undertaken by a third group of experts (the same than experts 2) through application to 21 actual proven cases of children who had received therapy in the LIC.

Figure 1 depicts this process of detailed specification of the problem domain by a series of meetings with the group of experts.

Table II provides a summary of the LIC cases involved in the KA process, with a view to properly verifying the KB. The necessary process of verification was developed through the construction of a web tool to allow the experts to check the system's answers for the developmental items provided by the KB. The KB will be validated by primary care pediatricians (experts 4) in autumn of 2012, using an ad-hoc web interface for validation of the resulting KB.

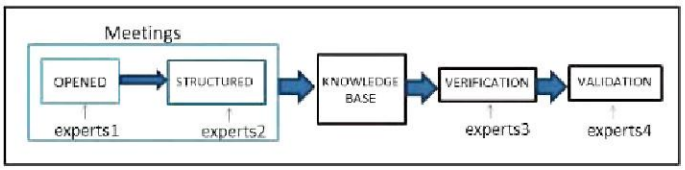

Fig. 1. Development methodology

The knowledge extracted in meetings should be represented in a well structured and comprehensible manner in a KB to make the knowledge useful and relatable for solving problems arising in the domain of execution. As noted by Torsun in 1995, representation of knowledge requires formalisms or structures representing it, either through declarative logic, mathematical formulas or concept maps [12]. The use of logical basis as a tool for knowledge representation requires adapting formal languages to models of knowledge expressed in natural language. Out of all the available models for knowledge representation, this research used description logic mainly because it is related to the development of ontologies for integration in the semantic web, as is done extensively with OWL language (Ontology Web Language) [13].

IV. RESULTS

This section contains the main results of this research. First, it describes the construction of the KB and then the process of formalization of the KB in OWL with the tool Protégé. Lastly, an example is shown of an ad-hoc web interface for verification of the resulting KB.

A. Knowledge base (KB) for language disorders

The KB has been built through an iterative process of structured meetings between September 2011 and May 2012. Only the two language therapists and the neuropediatrician participated in the meetings, using CK techniques to extract information such as structured interviews to complete the KB with questions to be asked by the primary care pediatrician children's tutor upon arriving at the care facility. The starting point for building the KB is the Denver Test, as it is extensively used in primary care [14].

With this basis, questions were sharpened to focus on language, with a view to enabling the primary care pediatrician to detect possible delays in children's development that require closer attention or immediate referral to an appropriate specialist. The process of building and purging the KB is based on the experience of LIC, who checked developmental items against the appearance of language disorders through a retrospective analysis of information on levels of language acquisition of 21 children who received therapy in the LIC [15]. Table II summarizes the data of the 21 LIC cases analyzed in the KA process, which lays the groundwork for the KB by indicating the most frequent alarm periods in detection of disorders and the resulting diagnosis.

TABLE II. ANALYSIS OF LIC CASES FOR REFINEMENT OF KB

Sex	66% boys and 29% girls
First alarm sign	66% cause alarm aged between 24 and 36 months because they do not speak
	14% cause alarm aged more than 36 months because they do not speak
	10% cause alarm owing to febrile seizure aged less than one year

Person who sounds the alarm	90.4%: parents
	4.76% of cases, alarm sounded by school
	4.76% of cases, alarm sounded by pediatrician
Diagnosis	66.6%: specific language impairment (SLI)
	28.6% delayed reading/writing
	4.76% cognitive delay

Observation of the data backs the hypothesis that diagnosis in all cases is related to the level of language acquisition, or problems deriving from it, when children begin to read and write. In most cases, it is the family which belatedly detects delays in language when a child barely speaks at the age of two years. In 2 of the 21 cases, parents did not become alarmed until the child has surpassed the age of three years. Information systematized with the LIC therapists is significant for mastering the problem [15] and enabled relating KB questions with the demand for intervention in cases of either alert or alarm. The study showed that the questions would have a negative answer in many of the cases under study and, for this reason, therapists consider them to be significant. The structuring of the final KB consists of 136 questions between month 1 and month 72 in the life of the child, and questions may be of two types:

a) Questions called Alert Milestones that imply bringing forward the visit. A negative answer to these developmental milestones means that the child makes a return visit within three months to allow for re-evaluation of the level of language acquisition.

b) Questions called Alarm Milestones that imply referral. These items could be considered as reasons for alarm and suggest referral to an appropriate specialist.

Table 3 details the KB questions for a child between the ages of 1 and 4 months. The first column indicates the child's age in months at the time of evaluation and the question type (Alert or Alarm). The second column shows the question the pediatrician asks the child's tutor to evaluate the child's state of language acquisition and the "System decision" column contains the system's answer in the event of a negative answer to the question (referral to specialist or bring forward visit).

TABLE III. SHOWS KB QUESTIONS FOR 1, 2, 3 AND 4 MONTHS

Age – Milestone	Question to be answered by pediatrician	System decision
1 month - Alarm	Reacts to a bell	Send to specialist to check hearing
1 month - Alarm	Vocalizes without crying	Send to specialist to check hearing
2 months - Alert	Emits "OOO/AAH"	Bring forward visit (three months)
2 months - Alert	Screams to interact	Bring forward visit (three months)
3 months - Alert	Turns around or reacts (closing eyes) to a clap	Bring forward visit (three months)
3 months - Alert	Turns at the sound of mother's voice	Bring forward visit (three months)
3 months – Alarm	Emits "OOO/AAH"	Check if hearing problem ruled out Refer to neuropediatrician
3 months – Alarm	Laughs in response to stimulus	Refer to neuropediatrician
4 months - Alert	Turns at the sound of mother's voice	Bring forward visit (two months)
4 months - Alert	Emits guttural sounds (AJOS)	Bring forward visit (two months)

B. Formalization of primary care KB in OWL

Construction of the ontology, according to Methontology, required categorizing the questions the pediatrician must ask according to the months of age of the child at the time of evaluation. Protégé was used in formalization of the knowledge model to create the KB and the inference engine needed to support decision making. Protégé offers an open and useful environment for the design, modeling, implementation, manipulation and viewing of ontologies. The ontology of Protégé was built with a class hierarchy for the first 6 years, including a sub-hierarchy of classes for each month corresponding to the questions to be asked by the pediatrician. The class hierarchy in AvanceSL, in each month, includes the questions for each month as classes, as shown in figure 2 for months 2 and 3 of the first year. The figure 3 shows the OWL code for the classes of months 2 and 3.

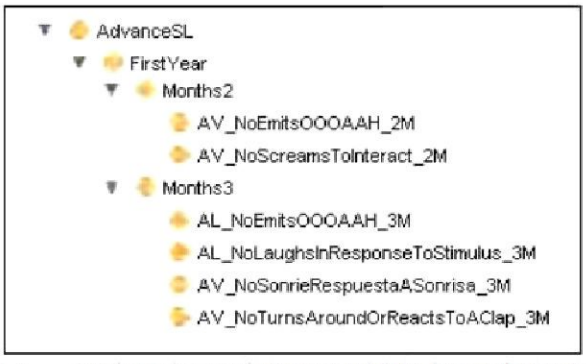

Fig. 2. Ontology of system in Protégé and OWL for months 2 and 3

```
<owl:Class rdf:ID="NextVisitInThreeMonths">
  <rdfs:subClassOf rdf:resource="#Year_1"/>
    <owl:equivalentClass>
      <owl:Restriction>
        <owl:someValuesFrom>
          <owl:Class>
           <owl:unionOf rdf:parseType="Collection">
            <owl:Class
               rdf:ID="AV_NoEmitsOOOAAH_2M"/>
            <owl:Class
           rdf:ID="AV_NoScreamsToInteract_2M"/>
           </owl:unionOf>
          </owl:Class>
        </owl:someValuesFrom>
      <owl:onProperty>
        <owl:ObjectProperty
           rdf:about="#ThereIsNegativeResponseIn"/>
      </owl:onProperty>
      </owl:Restriction>
    </owl:equivalentClass>
</owl:Class>
```

Fig. 3. Definition in OWL of the binary relationships between ontology classes

Figure 3 details the definition in OWL of the binary relationships established between ontology classes will sustain the system's reasoning process through axioms such as: If the child is 2 months old and we get a negative answer to the question: "Emits OOO/AAH" or "Screams to interact", then "To anticipate the next visit in three months" (figure 4).

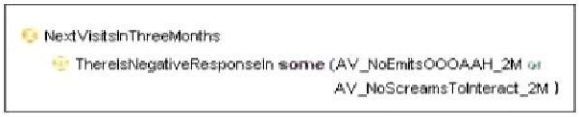

Fig. 4. Inference in Protégé for questions about 2 month-old child

Hence, the class hierarchy has been created to make inferences through the Decision System class, where this class contains motor decisions according to year and the type of milestone to which the decision belongs. Figure 5 shows the logical formulation of the correspondence of these axioms with inferences in Protégé.

C. *KB verification through a primary care tool*

The verification of the KB required an usable tool so that specialists might interact with the KBS in an efficient way. Experts should be able to evaluate whether the system proposal to refer to a specialist arising from a detection of language development disorders was correct or not. Thus, a web tool was built to facilitate the work of experts and primary care pediatricians. For this reason, the Pegasus web system was deployed, providing the final user with a easy to use verification interface as shown in Figure 6. The resulting tool is based on an internal connection of the KB, implemented in OWL, with a Pellet reasoning engine.

Fig. 5. Inference in Protégé for questions about 2 month-old child

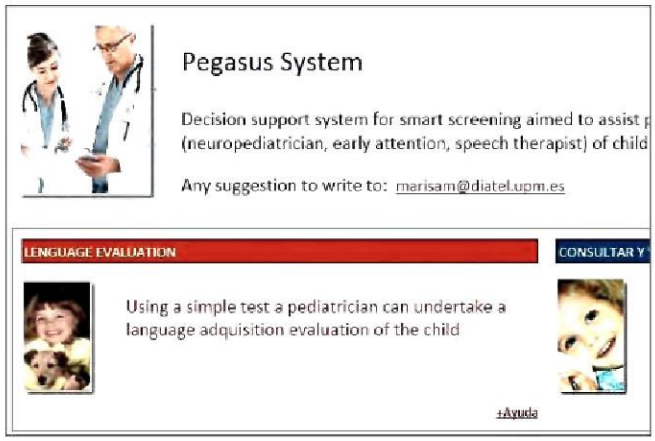

Fig. 6. Access to the early detection tool of language disorders

Five specialists have been involved in the verification stages both for usability and system performance tests along six months (Two pediatricians, two language therapists and one neuropediatrician). This process contributed in a satisfactory way to improve the graphical user interface and the reasoning rules of the KB. End users pointed out that they could use the tool by themselves in clinical routine. Once the end users logged into the system, he or she began to evaluate a child's state of language acquisition starting from the general information (Sex, Name, Date of birth, Gestation Period) as shown in Figure 7. This process was performed in a random and anonymous way with samples which are homologous to the original cases that were considered from the LIC records.

Fig. 7. Process of language evaluation

Figure 8 shows the sample of the question suggested by the system to check the language status of a two months baby: "Emits OOO/AAH" and "Screams to interact". Once the pediatrician answers to these questions, the result is provided by the early detection tool based on a negative answer to the question "Emits OOO/AAH". At this point, the system proposes bringing forward the next visit and also suggests that the pediatrician enter an opinion on the system's decision or an alternative to the proposal of the system as detailed in Figure 9.

Fig. 8. Process and result of language evaluation

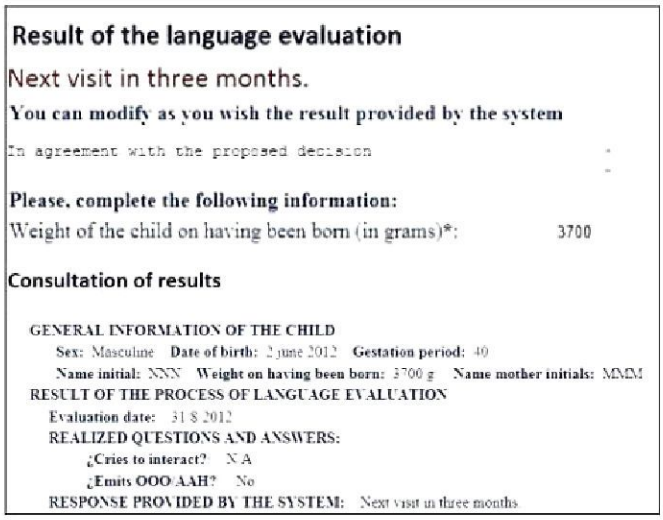

Fig. 9. Query of results of language evaluation

After the language evaluation process, the pediatrician or specialist - speech-language pathologist, neuropediatrician, early attention – to whom the child is referred did access to system to check the results of the language evaluation. Figure 9 displays the consultation of results as provided by the KBS in this step. This facility made possible the positive verification of the smart system for language disorders screening according the criteria of the five medical users involved.

V. CONCLUSIONS

This research work provides an innovative smart solution for the difficult task of detecting language disorders among children aged 0 to 6 in routine visits to pediatricians in primary care. The combined use of CommonKADS and Methontology facilitated the construction and formalization of an ontology that enables the building of an intelligent system capable of tackling the problem both in an efficient and effective way. The involvement of 4 experts in neuropediatrics, neonatology and language disorders was crucial for both defining the problem and for selecting significant 21 real, proven cases, from a base of about 60 clinical records, to refine the KB, verified thanks to available experience of 5 specialists in pediatrics care. The KB, formalized with the resulting ontology, can assist pediatricians in detecting language disorders, although deeper validation is foreseen with the 39 cases in the LIC that were not used in the formalization stage in order to fine-tune the system with the Pegasus tool. The opinion of the experts involved allows for a forthcoming start of the second stage of validation, with a view to deployment and evaluation in routine clinical practice. This controlled evaluation is scheduled for autumn 2012 in cases considered of interest by the two pediatricians already involved in the final stage. A future line of research would be the design of a model of self-learning based on unsuccessful referral information that can be compiled in the fine tuning planned for the system in the deployment phase in primary care.

ACKNOWLEDGEMENTS

Mrs. Paloma Tejeda of the Language Intervention Center (LIC) La Salle Campus Madrid, Universidad Autónoma de Madrid. Dr. José Arizcun, expert neonatologist in developmental disorders and child disability. Dra. Beatriz Chiclana and Dr. Erwin Kirchschlager, pediatricians in the Jazmin Health Center. Dra. Teresa Díaz, pediatrician in the Quirón Hospital of Madrid.

This article is part of research we're conducting in the Talisec+ project (a framework for knowledge-based management of accessible security guarantees for personal autonomy; TIN2010-20510-C04-01), supported by the Ministry of Education and Science of Spain through the National Plan for R+D+I (research, development, and innovation).

REFERENCES

[1] Grupo de atención temprana, "Libro blanco de la atención temprana", Real patronato de prevención y de atención a personas con minusvalía, 3rd ed., 2005.

[2] Council on Children With Disabilities, Section on Developmental Behavioral Pediatrics, Bright Futures Steering Committee and Medical Home Initiatives for Children With Special Needs Project Advisory Committee. "Identifying infants and young children with developmental disorders in the medical home", Pediatrics, vol. 118, 2006, pp. 405-420.

[3] J. Narbona and C. Chevrie-Muller, El lenguaje del niño. Desarrollo normal, evaluación y trastornos, Eds. Masson, 2003.

[4] M.C. Arrabal Terán and J. Arizcun Pineda, "Alteraciones del desarrollo y discapacidad", Grado de pediatría en España, Genysi, 2007.

[5] H.D. Nelson, P. Nygren, M. Walker and R. Panoscha, "Screening for speech and language delay in preschool children: Systematic evidence review for the US preventive services task force", Pediatrics, vol. 117, 2006, pp. 298-319.

[6] N. Fejerman and E. Fernández Álvarez, Neurología pediátrica, 3rd ed., Eds. Panamericana, 2007.

[7] R. Paul, Language disorders from infancy through adolescence: Assessment and intervention, 3rd ed., Eds. Mosby-Elsevier, 2007.

[8] E.H. Shortlife, "Computer–based medical consultations: MYCIN", Eds. American Elsevier, 1976.

[9] A. Alonso Betanzos, B. Guijarro Berdiñas, A. Lozano Tello, J.T. Palma Méndez and M.J. Taboada Iglesias, Ingeniería del conocimiento. Aspectos metodológicos, Eds. Pearson educacion. 2004.

[10] Inteligencia artificial en la medicina , http://galeon.com/iaenlamedicina/marco.html, 2004. [Last accessed on: May, 2012]

[11] O. Corcho, M. Fernández-López, A. Gómez-Pérez and A. López-Cima, "Building legal ontologies with METHONTOLOGY and WebODE", Springer, vol. 3369, 2005, pp. 142-57.

[12] I.S. Torsun, Foundations of Intelligent Knowledge-Based Systems, Eds. Academic Press, London, 1995.

[13] OWL web ontology language, http://www.w3.org/TR/owl-features/, 2004. [Last accessed on: May, 2012]

[14] W.K. Frankenburg, J.B. Dodds and P. Archer, "The DENVER II: A major revisión and restandardization of the denver developmental screening test", Pediatrics, 1992, pp. 89-91.

[15] M. Peñafiel Puerto, "Mejorando las habilidades en indicadores tempranos de los trastornos del lenguaje", FAPap, vol. 5, 2012.