The Dispedia Framework: A Semantic Model for Medical Information Supply

Romy Elze

Department of Computer Science
University Leipzig
Leipzig, Germany
elze@informatik.uni-leipzig.de

Klaus-Peter Fähnrich

Institute for Applied Informatics (InfAI) e.V.
Aninsitut of the University Leipzig
Leipzig, Germany
faehnrich@infai.org

Abstract—**The Dispedia Framework is an information system in the complex field of rare diseases. The goal of the system is to harmonize social care conditions and health care conditions with the focus on personalization and patient autonomy. The main task was to analyze the existing system of one rare disease and to structure and model the data in a tiered approach. On the basis of the Dispedia Model, we developed an information system that supports the information logistics between patients with rare diseases and other (all) players (e.g., doctors, therapists, and researchers).**

Keywords-Complex System; Ontology; Health Care; Information Model; Knowledge Mangement; Rare Disease; Patient Care Information System

I.   INTRODUCTION

This work describes an approach to engineer a health information system in the context of service engineering. "Service Engineering provides methods and tools for a systematic and structured development of new information-intensive service offerings and service systems [1]." The area of service modeling as a sub-category of service science includes the formal and semi-formal description of services. Standardized and reusable methodologies and models in this area are often deficient [2]. One example of a holistic model of integrated service systems is Böttcher's Metamodel [3]. This model includes the dimensions of service component, product model, process model, and resource model. It is particularly suitable for complex, varied services such as IT services. Considering comprehensive services which are provided collaboratively, the information model with a systemic view as a further dimension should be focused on. Only if the systems of medical services and information are understood as complex systems with highly interactive contexts and agents, the problems can be addressed through strategies that clarify the patterns and interrelationships of the system [4].

The aim of the paper is to show how patterns and interrelationships in a service system of complex disease can be found and how the information engineering can be realized. In Section 2 of the paper, the complex service system in case of rare diseases is described. In Section 3, the methods of analyzing and modeling of the system are demonstrated. Section 4 exposes the formalization of the system model. The summary in Section 5 identifies the most important facts of the paper, and in Section 6, the conclusions are drawn and fields of further work are tagged.

II.   PROBLEM DESCRIPTION

The field of research about rare diseases is extremely complex because of multifaceted disease progressions, the involvement of many stakeholders from different sectors (i.e., social care, health care, therapists, and aid suppliers), and the high costs incurred from treatment and care.

The knowledge of orphan diseases is primarily limited to a few specialized institutions [5]. This restriction often leads to excessive demands on the side of the medicating agents, who lack such specific knowledge. Activities are rarely coordinated centrally. Therefore, multidisciplinary care is still insufficient [6].

Due to the unpredictability of the course of a rare disease, it is impossible to preemptively provide information or define processes across the whole course. Due to a lack of quantitative and qualitative information, the confusion is also high on the side of the patients. This lack is diametrically opposed to the goal of supporting important treatment decisions at an early stage of the disease by means of an active information policy. Complex diseases like rare diseases (see Figure 1 (B)) are characterized by non-linear processes, and information and interventions subsequently depend on more contextual influences of the recipient. [4]
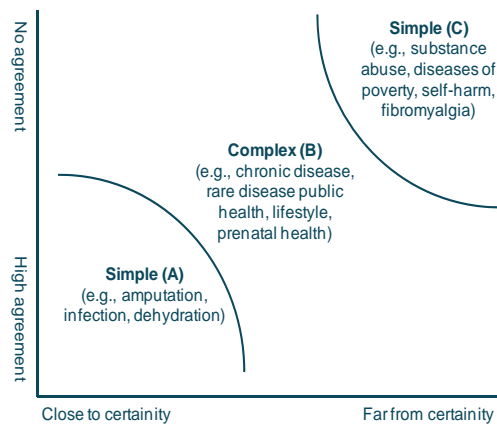

Figure 1: Complexity of diseases (according to [4])

Additionally, Berg constitutes that with respect to these approaches it is a vital aspect to consider the context of intervention- and information recipients [7]. These recipients can be both patients as well as other actors of the heterogeneous network. The challenge is to structure and to qualify the information for the information recipient [8],[9],[10]. Therefore a "…detailed consideration of the disease is inseparable from a detailed consideration of the whole person [11]". In the following, the methods and concepts for modeling such a complex topic are presented.

III. ANALYSIS AND MODELING OF RARE DISEASE

To analyze problems and processes of rare diseases, we chose to use the example of Amyotrophic Lateral Sclerosis (ALS), which is a degenerative disease of the nervous system [5]. To model the knowledge and relationships in the ALS knowledge domain, a concept map was developed as a starting point. The concept map [12] and the knowledge map [13] are procedures for the graphical representation of knowledge. Modeling knowledge with these techniques supports the qualitative and quantitative knowledge acquisition. The knowledge is thereby reduced to essential concepts which are arranged in relation to one another. The concept map was formed based on the analysis of 19 medical records of patients and their families from books and web

http://www.dgm.org/), the first model was consolidated. After the evaluation of an additional 41 expert interviews, the holistic Rare Disease Concept Model was completed (see Figure 2).

The demand for information is characteristic for the subject of ALS. Therefore, the model of the rare disease is designed with a focus on information. We distinguish between information for a patient, information about a patient, and information for stakeholders (see Figure 3).
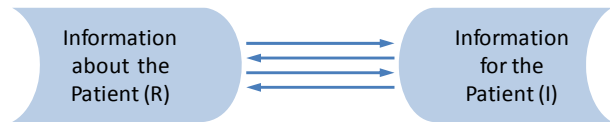

Figure 3. Information Types for Medical Information Supply

The medical field differentiates between specialized documentations of anamnesis. Especially in rare or chronic diseases and in rehabilitation, complex histories must be documented. With respect to ALS, for example, the neurologist uses the Functional Rating Scale for ALS (ALSFRS) [14]. Furthermore, therapists such as speech therapists or occupational therapists use their own
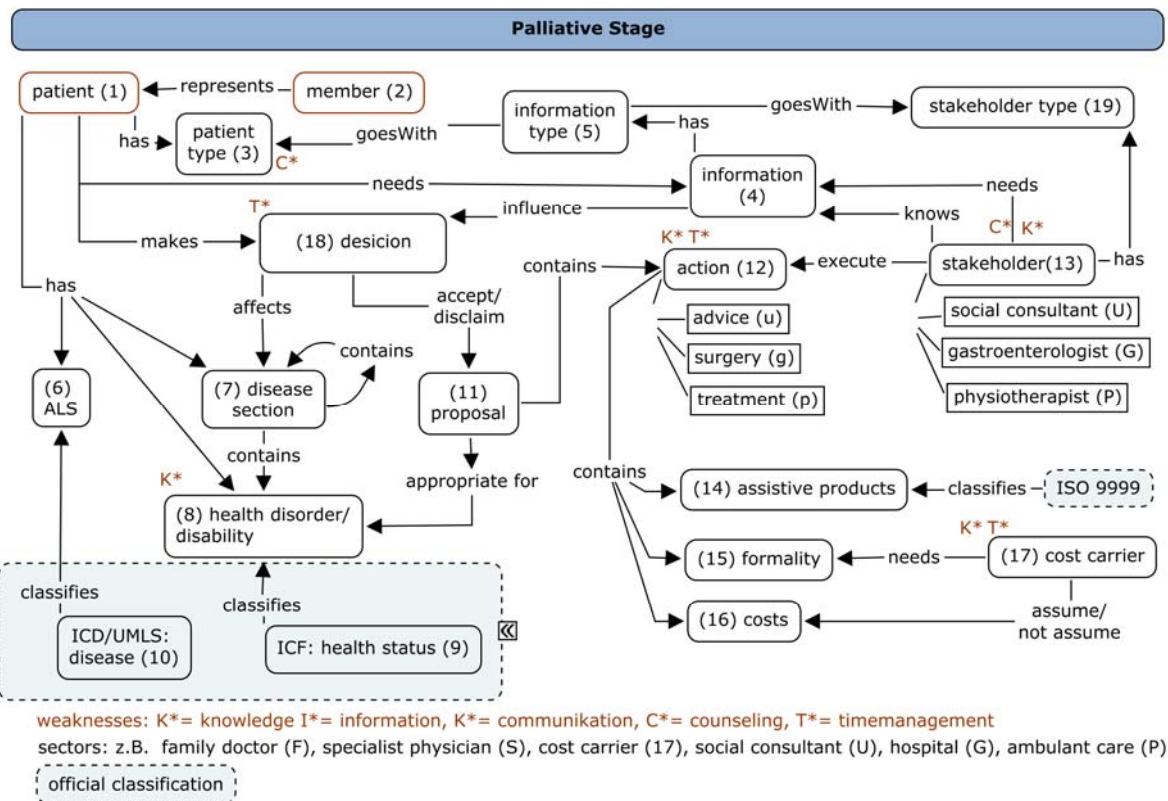

Figure 2. Concept Model of ALS Disease

sites. After discussion with experts (e.g., Charité Campus Virchow-Clinic Neurological Clinic: http://www.als-charite.de/; German Society for Muscle Diseases e.V.:

anamnesis sheets. In rehabilitation, a description concerning the International Classification of Functioning, Disability

and Health (ICF) is common. In the model, the utilization of all possible patient descriptions is intentional. Therefore, the idea of Linked Data [15] is pursued.

Modeled after personal consultations between professionals and patients, different information characters are represented. Patients not only have physical needs, but are also characterized by certain psychological traits. Patients and other parties involved in this process are therefore recipient types that must be categorized as specific information recipients. For instance, a very anxious and sensitive patient may be considered a special type, for example the patient type (3) "sensitive".

A patient who is a physician by trade would presumably hope to receive more detailed information and would be considered patient type "skilled". A patient who is very active and dedicated to his health may like to obtain information about treatment options to pursue by himself, for instance, anti-thrombosis injections. The information to be passed on to the patient in a counseling session would be matched exactly to the type of information recipient. For example, detailed information would be given to a skilled recipient who was listed under information type (5) "skilled".

As shown in Figure 2, the constructed Concept Model of ALS Disease demonstrates the relation between the concept patient (1) and the term of information (5).

IV. FORMALIZING SYSTEM MODEL

The formalization of the Disease Concept Model follows commonly used methodologies. These methodologies

include the analysis as described above, coding the knowledge, reusing existing formal ontologies, evaluation, and documentation [16]. The resulting formal knowledge representation was encoded in Web Ontology Language (OWL). "OWL is intended to be used when the information contained in documents needs to be processed by applications [17]", which allows, for instance, a modular development of complex areas and also the re-use of existing knowledge bases [18]. The employment of RDF/S [19] and OWL [17] for a representation of the Disease Model architecture has various advantages.

- The representation of information in a coherent structure without a direct connection to specific applications. This facilitates the development of applications which do not focus on the information logistics advised by this approach.
- The definite identification of the conceptual and concrete resources due to the use of dereferenceable URIs.
- The reutilization of the concepts in applied domains to generate an exchange of resources without a loss of information.
- Interlinking and reification of concrete resources.
- The application of established resources of the Linked Data Web, e.g., Dbpedia [20] and PubMed [21].

To use the formal Disease Model as a re-usable ontology, we developed the Dispedia Ontology (see Figure 4) hold under the domain www.dispedia.de. The name Dispedia is derived from the goal of the designed
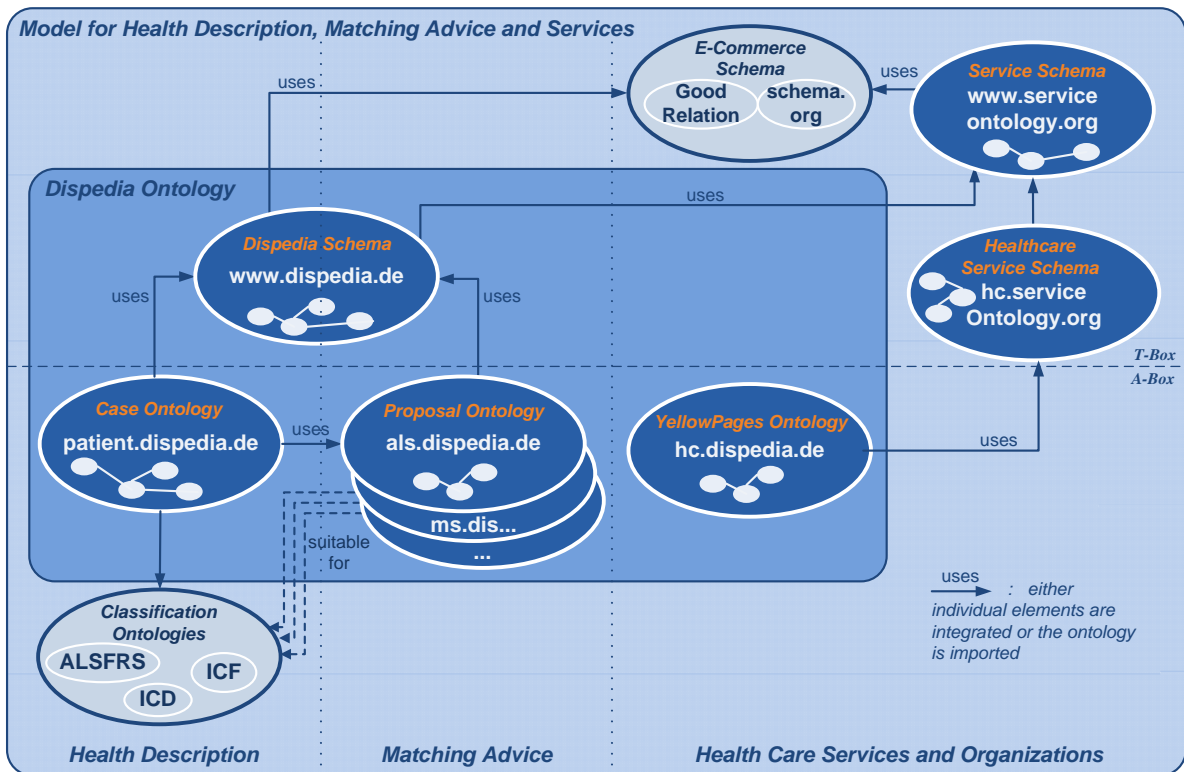

Figure 4. Dispedia Framework

knowledge base to make disease-specific information available and usable for humans and machines. The Dispedia Ontology includes the vocabulary or scheme (Core Ontology), the expert knowledge base (Proposal Ontology), the knowledge base about the patient (Case Ontology), the allocation of proposal information descriptions for patient parameters, and the concepts to interlink additional classifications (see Figure 4) as described above. The Dispedia Framework realizes the description of the whole person through subject-specific and user-dependent classifications, the modular description of disease-specific proposals, the integration of heterogeneous stakeholders, and the patient-specific allocation of information. A prototypical deployment of using the Dispedia Framework is realized on the basis of Ontowiki, a web based tool providing support for agile, distributed knowledge engineering scenarios [22].

V. SUMMARY

The semantic modeling of the medical information supply takes both the patient type and the information type into consideration. By taking advantage of the properties these types had in common, an adaptable information system could be developed. An application that uses the Dispedia Model Architecture adapted the described field of knowledge. The core ontology is the framework which includes describing classifications, domain specific knowledge, and matching. To manage information about the patient in the ALS example, the core ontology uses the classification ontologies ICF and ALSFRS. When using semantic web technologies, the modification and utilization of other classifications is feasible. At the same time, the subject of ALS information for the patient was replaced or amended with information for other recipients.

The use of the available e-commerce ontology www.schema.org furthermore allowed the information to be enriched with tangible offers and the tracking of patients, for instance, for medical aids.

VI. CONCLUSION AND FUTURE WORK

Our work focused on improving the information logistics in the existing health system regarding ALS. The actual state of our approach, which was evaluated by using select patients with ALS and the course of the disease in their specific cases, combines the following different advantages:

- The reutilization of the concepts in applied domains.
- Different information providers can operate on the same standard using the structure given by the architecture concepts and their relations.
- The ontological model of the system can be expanded by interlinking and remaining flexible with regards to accommodating new concepts.

- General aspects of the health care processes for the disease can be represented and linked to their implications for the individual stakeholder.
- Existing knowledge bases like the International Classification of Disease (ICD) [23] and the international Classification of Functioning, Disability and Health (ICF) [15] can be connected and re-used.

In the current state of development, we focus on functionalities supporting knowledge acquisition in order to allow the approach of Dispedia itself to evolve. Therefore, the integration of more usable medical classifications is targeted in a similar manner as the possibility of interlinking further proposal ontologies in addition to the actual existent ALS Ontology. We will pursue the idea of decentralizing data on the web for patients as well. After modeling the expert knowledge structure, the content of ALS hold under the domain als.dispedia.de has to be enlarged, completed and audited. In the same structure further knowledge bases about other diseases have to be developed. For the remote maintenance of patient data appropriate security concepts must be evaluated. The connectivity for cost carriers is one of the open investigation areas and so is the development of security strategies for sensitive data.

REFERENCES

[1] D. Spath, et al., "Service Engineering — A Transdisciplinary Approach in Service Research," in Services Science, B. Stauss, et al., Eds., ed: Springer Berlin Heidelberg, 2008, pp. 41-53.

[2] M. Böttcher and K.-P. Fähnrich, "Service Systems Modeling: Concepts, Formalized Meta-Model and Technical Concretion," in The Science of Service Systems, H. Demirkan, et al., Eds., ed: Springer US, 2011, pp. 131-149.

[3] M. Böttcher, "Architektur integrierter Dienstleistungssysteme - Konzeption, Metamodell und technikraumsspezifische Konkretisierung," Dr. rer. nat., Institut für Informatik, Universität Leipzig, Leipzig, 2009.

[4] C. A. Brown, "The application of complex adaptive systems theory to clinical practice in rehabilitation," Disability and rehabilitation, vol. 28, pp. 587-93, 2006.

[5] J. D. Mitchell and G. D. Borasio, "Amyotrophic lateral sclerosis," The Lancet, vol. 369, pp. 2031-2041, 2007.

[6] L. K. Ng, Fary and S. Mathers, "Multidisciplinary care for adults with amyotrophic lateral sclerosis or motor neuron disease," in Cochrane Database of Systematic Reviews, ed, 2009.

[7] M. Berg, "Patient care information systems and health care work: a sociotechnical approach," International Journal of Medical Informatics, vol. 55, pp. 87-101, 1999.

[8] P. J. K. O. S. E. Eysenbach G, "Empirical studies assessing the quality of health information for consumers on the world wide web: A systematic review," JAMA: The Journal of the American Medical Association, vol. 287, pp. 2691-2700, 2002.

[9] E. M. N. M. L. S. Berland Gk and et al., "Health information on the internet: Accessibility, quality, and readability in english and spanish," JAMA: The Journal of the American Medical Association, vol. 285, pp. 2612-2621, 2001.

[10] H. A. Taylor, et al., "Implementation of a user-centered framework in the development of a web-based health information database and call center," J Biomed Inform, vol. 44, pp. 897-908, Oct 2011.

[11] A. Miles, "Complexity in medicine and healthcare: people and systems, theory and practice," J Eval Clin Pract, vol. 15, pp. 409-10, Jun 2009.

[12]   A. J. Cañas*, et al.*, "Concept Maps: Integrating Knowledge and Information Visualization," in *Knowledge and Information Visualization*. vol. 3426, S.-O. Tergan and T. Keller, Eds., ed: Springer Berlin / Heidelberg, 2005, pp. 205-219.

[13]   A. O'Donnell*, et al.*, "Knowledge Maps as Scaffolds for Cognitive Processing," *Educational Psychology Review,* vol. 14, pp. 71-86, 2002.

[14]   J. Cedarbaum*, et al.*, "The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function," *Journal of the neurological sciences,* vol. 169, 1999.

[15]   WHO, "ICF The International Classification of Functioning, Disability and Health," 2002.

[16]   M. Uschold and M. Gruninger, *Ontologies: Principles, Methods and Applications* vol. 11, 1996.

[17]   W3C. (2004, Februar). *OWL Web Ontology Language Overview*. Available: http://www.w3.org/TR/owl-features/

[18]   R. Elze*, et al.*, "Dispedia.de – A Linked Information System for Rare Diseases," in *Information Quality in e-Health*. vol. 7058, A. Holzinger and K.-M. Simonic, Eds., ed: Springer Berlin / Heidelberg, 2011, pp. 691-701.

[19]   F. E. Manola, Miller. (2004, *Rdf primer*. Available: http://www.w3.org/TR/rdfprimer/.

[20]   C. Bizer*, et al.*, "Linked Data - The Story So Far," *Int. J. Semantic Web Inf. Syst.,* vol. 5, pp. 1-22, 2009.

[21]   A. Névéol*, et al.*, "Semi-automatic semantic annotation of PubMed queries: A study on quality, efficiency, satisfaction," *Journal of Biomedical Informatics,* vol. 44, pp. 310-318, 2011.

[22]   N. Heino*, et al.*, "Developing Semantic Web Applications with the OntoWiki Framework," in *Networked Knowledge - Networked Media*. vol. 221, T. Pellegrini*, et al.*, Eds., ed: Springer Berlin / Heidelberg, 2009, pp. 61-77.

[23]   WHO, "International Classification of Diseases (ICD)," 2006.