Child Abuse Monitor System Model: A Health Care Critical Knowledge Monitor System

Tiago Pereira(✉) and Henrique Santos

ALGORITMI Centre, Information System Department,
University of Minho, Azurém Campus, 4804-533 Guimarães, Portugal
tiago.pereira@algoritmi.uminho.pt,
hsantos@dsi.uminho.pt

Abstract. The Childhood protection is a subject with high value for the society, but, the Child Abuse cases are difficult to identify. The process from suspicious to accusation is very difficult to achieve. It must configure very strong evidences. Typically, Health Care services deal with these cases from the beginning where there are evidences based on the diagnosis, but they aren't enough to promote the accusation. Besides that, this subject it's highly sensitive because there are legal aspects to deal with such as: the patient privacy, paternity issues, medical confidentiality, among others. We propose a Child Abuses critical knowledge monitor system model that addresses this problem. This decision support system is implemented with a multiple scientific domains: to capture of tokens from clinical documents from multiple sources; a topic model approach to identify the topics of the documents; knowledge management through the use of ontologies to support the critical knowledge sensibility concepts and relations such as: symptoms, behaviors, among other evidences in order to match with the topics inferred from the clinical documents and then alert and log when clinical evidences are present. Based on these alerts clinical personnel could analyze the situation and take the appropriate procedures.

Keywords: Health care knowledge sensitivity · Health care decision support system · Ontology · Health care knowledge security · Knowledge management · Topic models · Information retrieval · Text mining

1 Introduction

The health care information systems had evolved strongly and with the recent implementations of the electronic health record these information systems interoperate together and allow us to access patient information within of multiple healthcare organizations. This aspect brings new opportunities to develop new technologies to promote more patient or health care services benefits. Since health care information is highly sensible there are important aspects that cannot be neglected, such as: patients privacy, medical confidentiality, confidentiality of the clinical procedures and legal aspects. This context is prominent when we're dealing with child abuse matters. Child abuse cases are difficult to identify since they are more pervert and dissimulated [1]. When clinical personnel identify evidences by diagnosis in order to formulate the

R. Giaffreda et al. (Eds.): IoT360 2014, Part I, LNICST 150, pp. 255–261, 2015.
DOI: 10.1007/978-3-319-19656-5_37

accusation these evidences are not strong enough to do it. On the other hand, we assume that this difficulty could lead to a substantial number of cases that aren't identified. To address this problem we've developed a child abuse critical knowledge monitor system that based on medical expert knowledge alert and log when child abuses evidences are present in clinical documents. This is done automatically by the system using information retrieval and text mining techniques which allows to transform text in tokens, from tokens we infer topics, then we match them with an ontology that supports the child abuse evidences as concepts and finally, when there are present in documents clinical evidences of child abuse the system alerts and logs these evidences for further analysis by clinical personnel. In the Sect. 2 we explain what implicit considerations we should consider in patient privacy and health care information security, in Sect. 3 we defined the health care critical knowledge, in Sect. 4 we explain the research methodology that this research is based, Sect. 5 we describe the components of the child abuse critical knowledge monitor system model, in Sect. 6 we'll present the preliminary results of this research, finally, in Sect. 7 we'll describe the contribution to health care information systems.

2 Privacy and Health Care Information Security

2.1 Information Security and Privacy

Information is valuable for any kind of organization and when there are rights involved, such as: patient privacy, medical confidentiality, and information that could stigmatize persons in social context like health care, the needs of information preservation are even higher. Information security consists in the preservation of the security properties of information and based on many standards and work published in this field, we normally identify a set of three properties there are the core of information security: confidentiality, integrity and availability of information [2, 3]. In health care its easy to recognize these properties but we have to reinforce them with the right of privacy that cannot be affected by the promotion of the information security properties. In a concise form, we'll describe each of these aspects.

Confidentiality. It means that any information that is exchanged or stored by the systems is secret and only the authorized persons can access to it. In health care this property is fundamental because in all phases of the clinical process exists information exchange and can only be accessed by authorized personnel.

Integrity. Keep the integrity of information is keeping its accuracy, i.e., there should be guaranties that some modification of the information is intentional and made by authorized personnel. In this objective, we need three actions, not allow unauthorized modifications, keep updated the authorized personnel and keep the accuracy of the information.

Availability. It means that the information should be accessible permanently at all the time. In health care the access to the patient information could be vital to him, e.g. An anesthesiologist should have access to information of the patient about allergies and other diseases since that with anesthesia could lead the patient to a complicated situation.

Privacy. Privacy, more than an information security property, it's a right. In health care we should consider the privacy of the patient, medical and non-medical personnel. For this reason no mechanism implemented to support the information security properties could affect the privacy that consists in maintaining the secrecy of personal information which force to get it free from monitoring and unauthorized access. Typically, to access to the patient information a consent declaration is needed.

3 Health Care Critical Knowledge

The development of standards of health care software is a big step to the interoperability between information systems in this area. There are, at least, six entities that have developed standards in this field: The American Society for Testing and Materials with ASTM-E31, The American National Standards Institute with ANSI-HL7, The European Committe for Standardization with CEN-TC251, The International Organization for Standardization with ISO-TC215, The Association of Electrical Equipment and Medical Imaging Manufactures with NEMA-DICOM and IEEE with multiple standards [4]. With the evolution of the health care information systems the access to patient information using Electronic Health Record Systems (EHRS) is facilitated, e.g. Urgency treatment data, health monitoring data, among others. According to the health domain analysis report from the technical committee from HL7 about security and privacy of health care information, particularly, in exchange of information between information systems and according to HL7 Security and Privacy Ontology was possible for us to identify critical knowledge concepts in health care domain [5, 6]: Substances abuse, Sexual abuse and domestic violence, Genetic disease, Sexual transmitted disease, Sickle Cell, Sexuality and Reproductive, HIV/AIDS, Psychiatry and Taboo [6]. From this, we explore the subject General Abuses with the focus on the child abuse and based on regulations and legal documentation we have constructed an ontology that maps the concepts: symptoms, behaviour and other evidences of child abuse [7], see Fig. 1.

4 Research Methodology

For the different phases and objectives of this research we have used multiple research techniques: survey literature techniques, content analysis and proof of concept in Design Research context.

4.1 Design Science Research

"Design Science research is a research paradigm in which a designer answers questions relevant to human problems via the creation of innovative artefacts, thereby contributing new knowledge to the body of scientific evidence. The designed artifacts are both useful and fundamental in understanding that problem" [8].

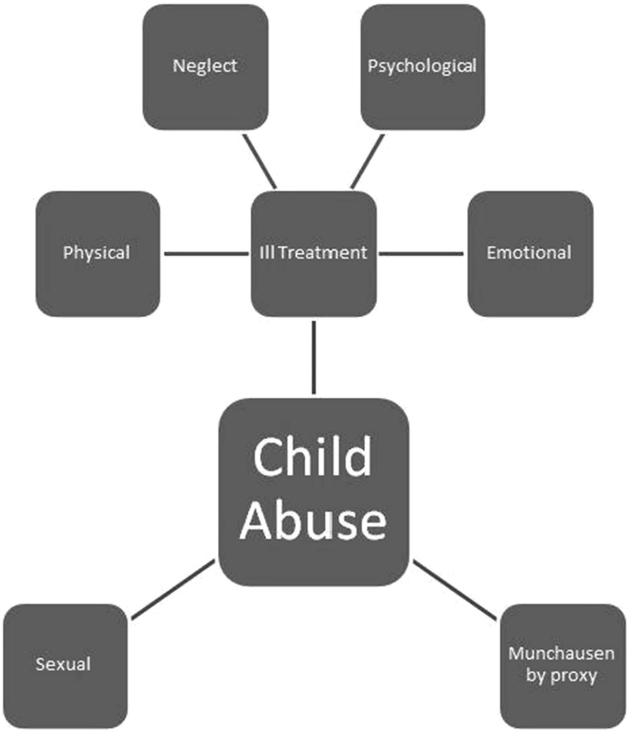

Fig. 1. Child abuse ontology

The design science research has its roots in the sciences of the artificial. Artificial as something that is created by humans that doesn't exists in Nature. Design Research is fundamentally a problem-solving paradigm. It consists in seeking innovation through ideas, practices, technical abilities and products obtained from a set of routines such as: analysis, design, implementation, and use of information systems concerning the effectiveness and efficiency achievement on organizations. The outputs forms of Design research could be: constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices), and instantiations (implemented and prototype systems) [8].

5 Child Abuse Critical Knowledge Monitoring System

The model of the system is defined by four components: the knowledge capture component, the critical knowledge ontology component; the critical knowledge repository component; and the alert and log component. Each component is based on a variety of information systems fields [10, 11].

5.1 The Knowledge Capture Component (KC)

The knowledge capture component requisites are: extracting tokens from documents in a variety of formats, such as text and audio. Additionally, the implemented system supports other formats: video (extracting sound and text within the video), webpages, among others; and transform the tokens extracted in such format that could be searchable concerning the privacy and confidentiality, integrity and availability of documents. In order to do it, we have implemented a topic model approach using two methods, latent Dirichelet [12–14] and Pachinko allocation model (PAM) [15, 16]. The PAM has been chosen because it can establish relations between topics and topic descriptors. The use of the topic model approach is fundamental because allows to driven topics from documents and ignore (because of its lower occurrence within a document) personal data (names, contacts and addresses) complying with the privacy and information security properties. To use topic models from the extracted tokens from the documents we needed to filter [17] them, essentially, tokens less than four (configurable) characters and trivial discourse tokens such as “and”, “or”, punctuation, among others. As an output of this component we get a searchable set of descriptors clustered by topics that co-occur in the document.

5.2 The Critical Knowledge Ontology Component (CKO)

The critical knowledge ontology component requisites are: allowing the editing of the critical ontology; and matching the ontology with the output of the knowledge capture component. This component uses Portégé Editor and Portégé API, see acknowledgments, for the matching procedure with the topic descriptors.

5.3 The Critical Knowledge Repository Component

The critical knowledge repository component requisites are allowing the storage of the outputs from the KC and CKO. This component uses a document management tool with the control version capability. The control versions could be useful in future implementations and we could analyse multiple diagnosis of the same patient in an historical perspective.

5.4 The Alert and Log Component

The alert and log component requisites are: alerting the user by email of the probability of the document containing evidences about a child abuse; and register the evidences identified by the system of each document for further analysis. The system should select what cases should trigger an alert or register only. In order to do this we’ll use artificial intelligence classification algorithms very well tested and implemented to assess the evidences and give value to the sensibility of the document in this context.

6 Preliminary Results

We have done tests in the laboratory only. The system is capable of identifying evidences within a document in many formats and in Portuguese and English languages, see Fig. 2. Documents with no relation to the subject of child abuse were tested, too, and they produce negative results, i.e., the system didn't identify evidences on these documents. The next step will be testing the system in real context. We'll test the system with documents from child abuse suspicion cases and confront the results with cases that was formulated an accusation. In order to evaluate the accuracy of the system by component we'll use the typical assessment models of knowledge management systems [18].

Topic 0		Topic 8	
child_young	0,14206	protection_child_young	0,02537
general_health	0,0566	health_bad	0,01202
situation_risc	0,00888	figure_algoritmo	0,01068
place_hospital	0,00666	framework_framework	0,00668
for_life_integrity	0,00555	integrity_physical_child	0,00668
physics_child_young	0,00555	nurse_medic	0,00668
parental_who	0,00555	abuse_neglect	0,00668
residence_child_young	0,00555	young_oposition	0,00534
court_family	0,00555	young_oposition_intervention	0,00534
model_report_for	0,00444	other_info	0,00534
procedure_intervention_for	0,00444	info_constant	0,00534
idea_suicidal	0,00444	intensive_when	0,00534
young_family_care	0,00444	for_fin	0,00534
health_primary	0,00333	health_for	0,00401
service_urgent_basic	0,00333	promote_health	0,00401
treatment_framework	0,00333	treat_intervention	0,00401
sindrome_munchausen	0,00333	inpatient_care_consult_external	0,00401
injuries_external_organs	0,00333	treat_care_health	0,00401
presence_sperm	0,00333	for_detection	0,00401
sperm_body	0,00333	uncomun	0,00401

Fig. 2. Topics captured from guide of detection of child abuse, the portuguese health general board document

7 Contributions of This Research

We propose a decision support system that addresses the problem of child abuse case identification. This subject is highly sensitive and valuable to the society and the community. This research has produced a child abuse ontology that could have many applications. In the information systems field, we bring to discuss the knowledge security: can we deal with knowledge security like we do with information security? The sensitivity of knowledge is fundamental when dealing with knowledge preservation?

Acknowledgement. This work is financed by FEDER funds through the Competitive Factors Operational Program – COMPETE and Portuguese national funds through FCT – Fundação para a ciência e tecnologia in project FCOMP-01-0124-FEDER-022674.

This work was conducted using the Protégé resource, which is supported by grant GM10331601 from the National Institute of General Medical Sciences of the United States National Institutes of Health.

References

1. JN: Maus tratos a crianças cada vez mais perversos e difíceis de identificar. Jornal De Notícias. Controlinveste (2013). www.jn.pt
2. ISO: ISO/IEC 27001:2013 Information technology - Security techniques - Information security management systems – Requirements. ISO (2013)
3. Oladimeji, E.A., Chung, L., Jung, H.T., Kim, J.: Managing security and privacy in ubiquitous eHealth information interchange, vol. 10, pp. 26:1–26 (2011)
4. IRMA: User-Driven Healthcare: Concepts, Methodologies, Tools, and Applications: IGI Global (2013)
5. H. S. O. WG: HL7 Version 3 Standard: Security and Privacy Ontology, Release 1, Version 3 ed: HL7 (2013)
6. S.W.g.o. HL7: Composite Security and Privacy - Domain analysis Model Report, Ballot, HL7, May 2010
7. P. M. o. Health: Despacho n.º 31292/2008, in 236, ed. Diário da república: INCM (2008)
8. Hevner, A., Chatterjee, S.: Design Science Research in Information Systems. Design Research in Information Systems, pp. 9–22. Springer, Heidelberg (2010)
9. March, S.T., Hevner, A.R., Park, J., Ram, S.: Design science in information systems research. Manage. Inf. Syst. Q. **28**, 75–105 (2004)
10. Pereira, T.R., Santos, H.: Critical knowledge monitor system model: healthcare context. In: European Conference on Knowledge Management, Cartagena, Spain (2012)
11. Pereira, T., Santos, H.: Health care critical knowledge monitor system model: health care critical knowledge ontology component. In: SHEWC2013 - XIII Safety, Health and Environment World Congress, Porto, Portugal, p. 002 (2013)
12. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
13. Hofmann, T.: Unsupervised learning by probabilistic latent semantic analysis. Mach. Learn. **42**, 177–196 (2001)
14. Steyvers, M., Griffiths, T.: Probabilistic topic models. In: Landauer, T., McNamara, D.S., Dennis, S., Kintsch, W. (eds.) Handbook of Latent Semantic Analysis, vol. 427, pp. 424–440. Erlbaum, Hillsdale (2007)
15. Mimno, D., Li, W., McCallum, A.: Mixtures of hierarchical topics with pachinko allocation. In: Proceedings of ICML, pp. 633–640 (2007)
16. Li, W., McCallum, A.: Pachinko allocation: DAG-structured mixture models of topic correlations. In: Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, Pennsylvania (2006)
17. Sharda, R., Delen, D., Turban, E.: Decision Support and Business Intelligence Systems, 9th edn. Prentice Hall Press, Upper Saddle River (2010)
18. Pereira, T.R., Santos, H.: The matrix of quality dimensions of knowledge management: knowledge management assessment models review. Knowl. Manage. Int. J **12**, 33–41 (2013)