A semantic tool for the protection of personal information act

Yahlieel Jafta[1][0000−0002−8307−5635] and Louise Leenen[1,2][0000−0002−9212−550X]

[1] Department of Computer Science, University of the Western Cape, Private Bag X17, Cape Town, Western Cape, 7535
2858132@myuwc.ac.za lleenen@uwc.ac.za
[2] Centre of Artificial Intelligence Research (CAIR)

Abstract. The increase of connectivity due to technology advances is shifting the attention of legislators in various jurisdictions to the protection of personal information and data. The focus of this paper is the protection of privacy information, specifically the POPIA within South Africa. This paper presents the development of an ontology to establish a small knowledge base for the regulations in the POPIA and how it affects organizations and individuals.

Keywords: Legal ontology · data protection · information privacy · Protection of Personal Information Act · General Data Protection Regulation

1 Introduction

Organizations that are operating in South Africa are being confronted by the authorization of the Protection of Personal Information Act 4 of 2013 (POPIA) [5]. The POPIA seeks to protect the right of privacy that applies to individuals and juristic entities (referred to as data subjects) by enforcing regulations [5]. It is with this in mind that a knowledge base for legislation of the POPIA will be valuable for assisting with the education of both the data subjects and organizations on the POPIA. The paper is arranged as follows. Section 2 describes related work Section 3 the requirements and methodology. Section 4 outlines the design, implementation followed by a summary and future work.

2 Background and related work

Semantic technologies refer to different technologies aimed at the derivation of the meaning of information. An ontology provides a shared domain vocabulary and a set of assumptions on the meanings of concepts described in the vocabulary.

An ontology for the GDPR was developed for the data protection requirements [1]. It shows the data protection prerequisites with regards to the GDPR change and introduces a methodology for incorporating it into a work process to express these necessities inside a business procedure through the ontology.

The GDPRov project [4] is an ontology concerned with the management of compliance through recognizing provenance information identified with assent and individual personal information required for consistency documentation. "It is an OWL 2 linked open data ontology" [4] that represents the provenance of assent and data lifecycle work processes for the GDPR.

3 Requirements and methodology

The requirements for the ontology are defined by a set of competency questions the ontology should answer. These questions will serve as the litmus test in the evaluation phase of the development process and will help define the scope of the ontology [2]. The OWL-based ontology editor, Protégé, was used.

Our ontology engineering methodology is a combination of METHONTOLOGY [3] and the Ontology Development 101 [2] method. METHONTOLOGY provides high–level activities for the development life cycle and the ontology development 101 method provides granular steps for the design and implementation phases.

4 Design and implementation

The design consists of listing all the key terms (chosen based on the competency questions) considered important for the knowledge base, defining an initial class hierarchy, and defining properties for classes and their features. Terms include *data subject, person, processing, personal information, has rights.*

The classes are created from a subset of the concepts listed in the terms above. *Person, Accountability, DataSubject, and ReponsibleParty* forms part of the initial set of classes. The properties are the concepts that will describe the relations between classes, as well as any additional data they might have. This includes, but not limited to, a data subject *has rights*, *the right to access* to information is a data subject right.

An initial evaluation was performed on the ontology by performing some basic queries using the DL Query plugin in the Protégé editor. The reasoner, Hermit version 1.4.3.456, was used for consistency verification. The queries included querying the rights of data subjects, describing personal information and special personal information. The evaluation is still in progress.

5 Conclusion and future work

This paper outlines the development of an ontology to provide a knowledge base on various concepts within the Act that will promote transparency and education that can aid with the inception of this Act. The development of the ontology is now in the evaluation. A set of classes and properties was created to demonstrate a working ontology. The evaluation of the ontology will be performed to assert the satisfaction of requirements followed by maintenance. The latest version of the ontology can be found at https://cs.uwc.ac.za/honours/2019/yjafta/ontology.owl.

References

1. Bartolini, C., Muthuri, R., Santos, C.: Using ontologies to model data protection requirements in workflows. In: Otake, M., Kurahashi, S., Ota, Y., Satoh, K., Bekki, D. (eds.) New Frontiers in Artificial Intelligence. pp. 233–248. Springer International Publishing (2017)
2. F. Noy, N., Mcguinness, D.: Ontology development 101: A guide to creating your first ontology. Knowledge Systems Laboratory **32** (01 2001)
3. Fernández-López, M., Gómez-Pérez, A., Juristo, N.: Methontology: From ontological art towards ontological engineering. In: Proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series. American Asociation for Artificial Intelligence (March 1997), http://oa.upm.es/5484/, ontology Engineering Group ? OEG
4. Pandit, H.J., Lewis, D.: Modelling provenance for gdpr compliance using linked open data vocabularies. In: PrivOn@ISWC (2017)
5. South African Government Gazette: Protection of personal information act (2013), http://www.justice.gov.za/legislation/acts/2013-004.pdf