Linking African Traditional Medicine Knowledge

Gossa Lô, Victor de Boer, and Stefan Schlobach

Department of Computer Science, the Network Institute,
Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
{a.g.lo, v.de.boer, k.s.schlobach}@vu.nl

Abstract. African Traditional Medicine (ATM) is widely used in Africa as the first-line of treatment thanks to its accessibility and affordability. However, the lack of formalization of this knowledge can lead to safety issues and malpractice. This paper investigates a possible contribution of the Semantic Web in realizing the formalization and integration of ATM with data on conventional medicine (CM). As a proof of concept we convert various ATM datasets and link them to CM data. This results in a Linked ATM knowledge graph. We finally give some examples with some interesting SPARQL queries and insightful results.

1 Introduction

Traditional medicine (TM) denotes a heterogeneous set of diagnostic- and therapeutic practices that uses scientifically often unvalidated knowledge, orally passed on through generations. Its aim is to prevent, diagnose and treat physical and mental illnesses. The World Health Organization notes that African Traditional Medicine (ATM) plays a crucial role in Sub-Saharan Africa where it is the first line of treatment for ca. 80% of the population [1], thanks to its accessibility, affordability and embeddedness in the traditional belief systems [9].

As the use and availability of plants differs per region, lack of formalization not only raises safety concerns, but also risks the loss of critical knowledge [3]. Formalizing TM knowledge can benefit the determination of medical effects of plants in terms of diseases and symptoms [5]. Furthermore, bridging the gap between ATM and CM by integrating the former into formal health care systems can result in a deeper understanding of pathology and medical knowledge.

This paper presents a proof of concept on how Semantic Web technologies can contribute to the preservation and formalization of ATM, and its integration with CM. This is demonstrated by converting ATM datasets into machine-readable data and by connecting them to Linked Data knowledge graphs on the Web. Finally, the benefits of the presented concept will be shown by validating the use cases formulated below with the use of queries.

An illustration of how Web technologies can contribute to the representation of TM is the Traditional Chinese Medicine (TCM) Database System[1]. It has been maintained since 1984 by the Institute of Information on TCM, and contains over

[1] http://cowork.cintcm.com/engine/windex1.jsp, accessed 2017-07-5

1,100,000 items of data, divided over 40 categories. Furthermore, a semantic eScience infrastructure was established, in response to the sheer volume and diversity of TCM information and services that affect its interoperability [2].

The Linking Open Drug Data (LODD) project aggregated and added biomedical data (e.g. about drugs, TCM, diseases, etc.) to the Linked Data Cloud. Their focus is to facilitate the obtaining of new insights and finding unforeseen associations between entities. The data sets contain over 8.4 million RDF triples and 388,000 RDF links to external data sources [4,8].

2 Case study and methods

We investigate the potential of Semantic Web technologies for the preservation and formalization of ATM knowledge through two use cases. The first concerns the difference in treatment in two regions in Madagascar. Both the availability of plant species in a region and the knowledge about treatment methods that is passed on in a community influence treatment practice. Formalizing these methods prevents loss of knowledge and can help ATM practitioners in rural regions to exchange knowledge and find new treatments. The second use case focuses on investigating differences between ATM practice in Senegal and Madagascar. This exchange of knowledge not only benefits ATM practitioners, but could also be a valuable contribution to drug discovery and pharmacology on a global scale.

Keur Massar Traditional Hospital. One dataset used was provided by Hôpital Traditionnel de Keur Massar[2] in Senegal founded in 1980 by Professor Yvette Parès[3], who had years of experience working with traditional practitioners from different ethnic groups in Senegal. This hospital has a strong focus on phytotherapy along ATM traditions and produces medicinal plant preparations (about 1,000 products and recipes) from their botanical garden.

Datasets. Besides this dataset, two datasets on ATM practice in Madagascar were used. The first contains data on medicinal plants used to treat the six most frequent diseases in the Ambalabe rural community in Madagascar [6]. The second dataset contains data on the most used medicinal plants by communities in Mahaboboka, Amboronabo, Mikoboka, in Southern Madagascar [7]. The most relevant attributes that were stored are: the hierarchy of plant names, plant parts, in addition to diseases, ailments and preparation & administration modes.

We then linked the plant- and disease data in our datasets to resources on BioPortal[4] and DBpedia[5]. The BioPortal REST API was used to search for URIs for terms across ontologies, such as the Human Disease Ontology [6], the

[2] http://www.hopitalkeurmassar.com/
[3] Head of the department of Plant Biology at Cheikh Anta Diop University in Dakar
[4] http://bioportal.bioontology.org/
[5] http://www.dbpedia.org
[6] https://bioportal.bioontology.org/ontologies/DOID

Symptom Ontology [7] and SNOMEDCT [8]. If no URI is found, a synonym of the term is entered or linked to a proper resource on DBpedia.

Dataset conversion and linking. Our three datasets (.CSV) were translated into RDF and linked to the target datasets. First, columns from the original tabular data are stored in a dictionary as values, and assigned to a new key object name:
'set2_99': [u'Emilia humifusa DC. Rakotoarivelo', u'Asteraceae', u'Infected wound', u'L', u'Angea', u'Cataplasm on wound', u'400', 0.02, u'ID', 25L]
A new data model was created that does not use existing ontologies, with converted triples of the form `atmData:set2_99 atmVocab:familyName dbpg:Asteraceae`.

Results. The result of this conversion is a total of 13,028 triples, of 672 plant types that treat 1,799 health conditions. The data is available at a public GIT repository at https://github.com/biktorrr/linkedatmdata. For live browsing and querying, we provide a triple store endpoint at http://semanticweb.cs.vu.nl/linkedatm/home. The 672 set objects all have the same structure as the below RDF Turtle snippet shows.

```
atmData:set2_99 atmVocab:ailment atmData:Infected_wound ;
  atmVocab:binomialName atmData:Emilia humifusa DC. Rakotoarivelo ;
  atmVocab:familyName dbpg:Asteraceae ;
  atmVocab:plantParts "L"^^xsd:string ;
  atmVocab:preparationAdministration "Cataplasm on wound"@en .
```

New and existing prefixes are used to simplify SPARQL queries with:

```
prefix dbr: <http://dbpedia.org/resource/>
prefix dbo: <http://dbpedia.org/ontology/>
prefix atmData: <http://semanticweb.cs.vu.nl/linkedatm/resource>
prefix atmVocab: <http://semanticweb.cs.vu.nl/linkedatm/vocab/>
prefix prl: <http://purl.obolibrary.org/obo/>
prefix snmd: <http://purl.bioontology.org/ontology/SNOMEDCT/>
```

3 Validating

By querying possible outcomes of the use cases, the proof of concept is validated. Consider the following use case:
How does Malaria treatment differs in two different regions in Madagascar?

```
SELECT ?ailment ?family ?binomialName ?parts ?preparation WHERE {
  ?s atmVocab:familyName ?family .
  ?s atmVocab:binomialName ?binomialName .
  ?binomialName rdfs:label ?binomial .
  ?s atmVocab:genusType ?genusType .
  ?s atmVocab:ailment|atmVocab:diseaseTreated ?ailment .
  ?ailment rdfs:label ?ailmentLabel .
  ?s atmVocab:plantParts ?parts .
  ?s atmVocab:preparation|atmVocab:preparationAdministration ?preparation .
  FILTER(?ailmentLabel = "Malaria"@en) }
GROUP BY ?ailment ?family ?binomialName ?parts ?preparation
LIMIT 6
```

[7] https://bioportal.bioontology.org/ontologies/SYMP
[8] https://bioportal.bioontology.org/ontologies/SNOMEDCT

As Fig. 1 shows, differences in plant use and treatment of malaria occur even between neighboring regions. While the Ambalabe community uses several parts, the other regions predominantly use the leaves. Thus, LinkedATM could both contribute to collaborative knowledge sharing between communities, and preserve critical knowledge for future generations.

ailment	family	binomialName	parts	preparation
Malaria	Amaranthaceae	Chenopodium ambrosioides L. Sussman	"L"^^xsd:string	"Decoction drunk and bathe"@en
Malaria	Amaranthaceae	Henonia scoparia Moq. Randrianarivony	"L"^^xsd:string	"Decoction steam inhale"@en
Malaria	Amaranthaceae	Henonia scoparia Moq. Randrianarivony	"R"^^xsd:string	"Powder mix with water and forehead mask"@en
Malaria	Anacardiaceae	Sclerocarya birrea subsp. caffra (sond.) Kokwaro Rakotoarivelo	"L"^^xsd:string	"Decoction, steam inhale"@en
Malaria	Apocynaceae	Catharanthus roseus (L.) G. Don	"Leaves"^^xsd:string	"Decoction"@en
Malaria	Apocynaceae	Mascarenhasia lisianthiflora A. DC. Rakotoarivelo Randrianarivony	"L"^^xsd:string	"Decoction drunk"@en

Fig. 1. Difference in treatment of malaria in distinct communities in Madagascar

The second case focuses on the difference in use of a specific plant species between Senegal and Madagascar.The Sclerocarya birrea tree grows in both Southern and West Africa. However, the treatment purposes differ per country. The following query shows its use in Senegal:

```
SELECT DISTINCT ?family ?genusType ?binomialName ?indication WHERE {
  ?set atmVocab:nomGenre ?genusType .
  SERVICE <http://dbpedia.org/sparql> { ?genusType dbo:family ?family } .
  ?set atmVocab:nomBinomial ?binomialName .
  ?set atmVocab:indication ?indication .
  ?indication rdfs:label ?indicationLabel .
  FILTER(?genusType = dbr:Sclerocarya) .
  FILTER(?binomialName = atmData:Sclerocarya_birrea) .}
GROUP BY ?family ?genusType ?binomialName ?indication
```

The family of the tree (i.e. Anacardiaceae) is unknown and is identified by linking the known genus type to DBpedia. Fig. 2 shows that the tree is used to treat hypoglycemia, as an anti-infective, an antivenin or an astringent for the skin.

family	genusType	binomialName	indication
Anacardiaceae	Sclerocarya	Sclerocarya birrea	Antinfectieux
Anacardiaceae	Sclerocarya	Sclerocarya birrea	Antivenumeux
Anacardiaceae	Sclerocarya	Sclerocarya birrea	Astringent
Anacardiaceae	Sclerocarya	Sclerocarya birrea	Hypoglycemiant

Fig. 2. Sclerocarya birrea in Senegal - query result

The aforementioned communities in Southern Madagascar on the other hand, use this tree to treat malaria, in prenatal care, postpartum recovery, dizziness during pregnancy, fever and caries, as shown in Fig. 3.

family	genusType	binomialName	ailment
Anacardiaceae	Sclerocarya	Sclerocarya birrea subsp. caffra (sond.) Kokwaro Rakotoarivelo	Dizziness During Pregnancy
Anacardiaceae	Sclerocarya	Sclerocarya birrea subsp. caffra (sond.) Kokwaro Rakotoarivelo	Post Partum Recovery
Anacardiaceae	Sclerocarya	Sclerocarya birrea subsp. caffra (sond.) Kokwaro Rakotoarivelo	Prenatal Care
Anacardiaceae	Sclerocarya	Sclerocarya birrea subsp. caffra (sond.) Kokwaro Rakotoarivelo	Fever
Anacardiaceae	Sclerocarya	Sclerocarya birrea subsp. caffra (sond.) Kokwaro Rakotoarivelo	Caries
Anacardiaceae	Sclerocarya	Sclerocarya birrea subsp. caffra (sond.) Kokwaro Rakotoarivelo	Malaria

Fig. 3. Sclerocarya birrea in Madagascar - query result

4 Conclusion

This paper stipulates that formalizing ATM, and linking it to conventional medicine, can yield innovative knowledge benefiting both ATM practitioners and researchers on the usage of plant components in conventional pharmacology. Our proof of concept focuses on two concrete use cases, using data from an ATM hospital in Senegal and the results from earlier use cases in Madagascar. This has resulted in a dataset of 13,028 RDF triples, which describe 672 plant types and 1,799 health conditions. The data has been linked to knowledge on the Web (BioPortal & DBpedia), and shown to be of interest in the use cases. Further steps are required, including obtaining more comprehensive and varied data, linked to pharmaceutical ontologies for comparison with conventional medicinal substances. Making the data accessible to ATM practitioners in rural areas in Africa would require the development of an offline and voice-based version. Semantic Web techniques can, without a doubt, contribute to Linked ATM. To achieve this, there are challenges to overcome. However, as was demonstrated in this proof on concept paper, the first step has been taken.

Acknowledgements. We thank Djibril Bâ and Geneviève Baumann for their input. This work is supported by the W4RA initiative (http://w4ra.org).

References

1. Antwi-Baffour, S.S., Bello, A.I., Adjei, D., Mahmood, S.A., Ayeh-Kumi, P.F.: The place of traditional medicine in the african society. AJHR 2, 49–54 (2014)
2. Chen, H., Mao, Y., Zheng, X., Cui, M., Feng, Y., Deng, S., Yin, A., Zhou, C., Tang, J., Jiang, X., et al.: Towards semantic e-science for traditional chinese medicine. BMC bioinformatics 8(3), S6 (2007)
3. Firenzuoli, F., Gori, L.: Herbal medicine today: clinical and research issues. Evidence-Based Complementary and Alternative Medicine 4(S1), 37–40 (2007)
4. Jentzsch, A., Zhao, J., Hassanzadeh, O., Cheung, K.H., Samwald, M., Andersson, B.: Linking open drug data. In: I-SEMANTICS (2009)
5. Kamsu-Foguem, B., Diallo, G., Foguem, C.: Conceptual graph-based knowledge representation for supporting reasoning in african traditional medicine. Engineering Applications of Artificial Intelligence 26(4), 1348–1365 (2013)
6. Rakotoarivelo, N.H., Rakotoarivony, F., Ramarosandratana, A.V., Jeannoda, V.H., Kuhlman, A.R., Randrianasolo, A., Bussmann, R.W.: Medicinal plants used to treat the most frequent diseases encountered in ambalabe rural community, eastern madagascar. Journal of ethnobiology and ethnomedicine 11(1), 68 (2015)
7. Randrianarivony, T.N., Ramarosandratana, A.V., Andriamihajarivo, T.H., Rakotoarivony, F., Jeannoda, V.H., Randrianasolo, A., Bussmann, R.W.: The most used medicinal plants by communities in mahaboboka, amboronabo, mikoboka, southwestern madagascar. Journal of ethnobiology&ethnomedicine 13(1), 19 (2017)
8. Samwald, M., Jentzsch, A., Bouton, C., Kallesøe, C.S., Willighagen, E., Hajagos, J., Marshall, M.S., Prud'hommeaux, E., Hassanzadeh, O., Pichler, E., et al.: Linked open drug data for pharmaceutical research and development. Journal of cheminformatics 3(1), 19 (2011)
9. World Health Organization: Traditional medicine strategy 2002-2005. http://www.wpro.who.int/health_technology/book_who_traditional_medicine_strategy_2002_2005.pdf (2002)