Abstract—Approximately four hundred indigenous communities in Indonesia originally utilize their traditional knowledge for supporting their daily life. Because of many benefits of that knowledge, many stakeholders have started to collect and write it into a digital report. However, the digital report was still documented in the different format of metadata because there is no specific metadata schema for describing digital data of traditional knowledge. Moreover, the differences of metadata schema will make the difficult process of documenting, managing and disseminating this traditional knowledge. To overcome this problem, this work attempted to design specific metadata schema for a domain of traditional knowledge by utilizing metadata development methods, i.e., domain analysis, derivation analysis, system-centric analysis, user-centric analysis and resource-centric analysis. The selection of those methods based on literature review result toward research articles that presented about metadata development. As a result, this paper proposed metadata schema of traditional knowledge that consists of 37 metadata elements which are categorized into 6 metadata sections, i.e., supporting data, material, supporting tool, success story, knowledge source, and knowledge engineer.

Keywords—metadata schema; traditional knowledge; information management

I. INTRODUCTION

Today, traditional knowledge as cultural heritage is not only important and utilized by indigenous people but also urban people for supporting their daily life. For example, traditional knowledge for health is still used by rural areas population in developing countries [1], [2]. However, the traditional knowledge rapidly disappears and largely undocumented because that knowledge is owned by elder and accessibility is restricted only to a member of within indigenous community itself.

In Indonesia, traditional knowledge has begun to be preserved [3] by many stakeholders through digital documentation to avoid those knowledge being disappearing [4]–[6]. However, the digital data was still documented in different metadata schema because there is no specific metadata schema for traditional knowledge. The different format metadata will impact to reuse, manage, and disseminate data of traditional knowledge [7], [8].

The previous researches have been proposed metadata schema for traditional knowledge. For examples, metadata proposed by [9] that were applied to Chinese Medicine Digital Library (CMDL), metadata developed by [10] for Chinese medicine literature metadata (TCMLM, for short) and so forth. However, metadata cannot be ‘one size fits all’ resources because it depends on the purpose of development [11].

Furthermore, a good metadata schema should cover all needs of many parties that will use metadata schema, such as end users, the connected system and so forth. Based on [12] metadata schema can be developed through several methods such as domain analysis, derivation analysis, system-centric analysis, user-centric analysis and resource-centric analysis.

In reference to research background above, this research attempted specific metadata schema for the domain of traditional knowledge by employing metadata development methods. This paper will organize into introduction, literature review, methodology, result and discussion, conclusion, acknowledgment, and references.

A. Metadata Schema

The definition of metadata schema is commonly described as specifications for representing metadata element in order to present information structure about data or dataset [13], [14]. The information structure consists of element or attribute values that are referred to the description of the data, such as attribute value about data owner, data format, and so
forth. The main aim of metadata schema development is to ease in understanding, organizing and storing the data [15], [16].

White (2005) [15] identifies four categories of metadata: structural metadata, content metadata, descriptive metadata and administrative metadata. Structural metadata is commonly described information architecture of the document, for example, title, summary, image, and so forth. Content metadata provides information about subjects associated with a particular document. Descriptive metadata presents information to ease in searching document based on its format. Then, administrative metadata delivers detailed information about date created or modified, owner of the document, and so forth.

B. Traditional Knowledge

Traditional knowledge as intangible cultural heritage is commonly defined as local, intangible and unique knowledge based on experimental that is originated from a particular local community and delivered through oral tradition [18]. The example of traditional knowledge is the knowledge how to indigenous people manage their tradition [18]. The example of traditional knowledge is the particular local community and delivered through oral knowledge based on experimental that is originated from a commonly defined as local, intangible and unique knowledge how to indigenous people manage their ecological relations of society and nature and knowledge how to adapt to environmental or social changes [19]. Based on [20], traditional knowledge is categorized into eight fields: beliefs, medicine, knowledge technology, education, communication, agriculture, food technology, and arts and crafts.

C. Method of Metadata Development

1) Domain Analysis: is used as the preliminary study to develop metadata which is adapted from a field of Information Science [21] through interpreting domain of information resources and defining its scope [22]. The particular aim of the method is to identify a domain of text resource, type of text resource, end user group and their activity in using the text resource [23].

2) Derivation Analysis: is a method to derive metadata element by reviewing the related existing metadata standard or schema [24]. For instance, metadata MARC-XML is the result of derivation analysis from metadata MARC.

3) System-Centric Analysis: is a method to derive metadata element by reviewing information architecture of an existing system that is related to proposed metadata [8]. The example of this method utilization is to develop metadata schema of the phyknome project [25].

4) User-Centric Analysis: is a method to identify metadata element by reviewing information needs of end user [26] for achieving their specific purpose [27]. The example of user-centric analysis utilization in metadata development is presented on research by [28] by observing information needs of health practitioner in a library of a large pharmaceutical company. As the result of the user-centric analysis, they found that drugs, diseases, genes, companies, methods, authors, geographic regions, and drug sales were new elements of metadata schema for the library of a pharmaceutical company.

5) Resource-Centric Analysis is a method to identify metadata element-based available information on the resource that will be described by metadata schema [7]. The example of resource-centric analysis utilization in metadata development is presented on Chao (2015) [27]. As the result of the resource-centric analysis, Chao (2015) found nine mandatory metadata elements for journal articles of soil science, including method descriptive name, method type/sub-category, brief method summary, method number/identifier, method source, source citation, media name, method official name, and instrumentation.

D. Related Work

In recent years, there has been an increasing amount of metadata schema for traditional knowledge that is developed for different purposes by using various methods because a metadata cannot be ‘one size fits all’ [11]. In 2000, metadata schema for traditional knowledge of Chinese medicine had been proposed by Yang and Chan [9] that were applied to Chinese Medicine Digital Library (CMDL). Yang and Chan divided their metadata schema into three sections, i.e., herbs, proprietary, and recipes [9]. The development of metadata schema used a user-centric method by asking information needs to end user and system-centric method by reviewing information architecture of previous Chinese medicine system. Then, eleven years later, [30] proposed metadata schema for cultural heritage documentation which consisted of metadata sections: people, restoration, management, basic, history, building, and publishing.

In 2012, [31] proposed metadata for the documentation of archaeological assets (objects, ancient buildings, and archaeological sites) in Archaeology Research Centre (STARC). The development of metadata schema used derivation analysis, system-centric analysis, user-centric analysis, and resource-centric analysis that defined metadata elements and grouped them into project information, cultural heritage asset, digital resource provenance, and activities.

In 2013, [32] proposed metadata about Chinese intangible cultural heritage that was developed by using derivation analysis to Dublin Core and related metadata schema and user-centric analysis by reviewing end-user needs (government). The result of [32] work is the 67 metadata elements that were grouped into 14 sections. In the same year, metadata schema for traditional knowledge of Chinese living epic traditions has been proposed by [33]. It was developed by using derivation analysis, system-centric analysis, user-centric analysis, and resource-centric analysis. As many 104 metadata elements have been identified and grouped into 19 metadata element sections.

In 2014, [10] proposed traditional Chinese medicine literature metadata (TCMLM, for short). It was developed by using derivation analysis to metadata Dublin Core and ISO 13119 Health Informatics and user-centric analysis for defining specific metadata element of traditional Chinese medicine. TCMLM consists of 24 metadata elements which are grouped into 7 sections: identification, content, distribution, quality, constraint, maintenance, and relation. Then, all related works have been summarized in Table 1 below.
TABLE I
RELATED WORKS OF METADATA SCHEMA FOR TRADITIONAL KNOWLEDGE

| Author          | Total of element | Total of section | Domain of data                                      | System applied metadata                                      | Method used                                      | Metadata Schema |
|-----------------|------------------|------------------|----------------------------------------------------|----------------------------------------------------------------|------------------------------------------------|-----------------|
| Yang and Chan (2000) [9] | 35               | 3                | Traditional Chinese medicine                       | Chinese Medicine Digital Library (CMDL)                         | User-centric method and system-centric method   | Appendix A      |
| Ning et al. (2011) [30] | 78               | 7                | Cultural Heritage Documentation                    | Integrative Management System of Taiwan Cultural Heritage      | System-centric method                           | Appendix B      |
| Ronzino et al. (2012) [31] | 78               | 4                | Archaeological assets (objects, ancient buildings, and archaeological sites) | Science and Technology in Archaeology Research Centre (STARC) repository | Derivation method, resource-centric method, and user-centric method | Appendix C      |
| Ye & Zhou (2013) [34] | 68               | 14               | Chinese intangible cultural heritage               | Not mentioned                                                  | Derivation method, resource-centric method, and user-centric method | Appendix D      |
| Qobumo et al. (2013) [33] | 106              | 19               | Living Epic Traditions in China                    | Institute of Ethnic Literature (IEL), Chinese Academy of Social Sciences (CASS) | Derivation method, resource-centric method, and user-centric method | Appendix E      |
| Yu et al. (2014) [10] | 24               | 7                | Traditional Chinese Medicine Literature Metadata   | Not mentioned                                                  | User-centric analysis method and derivation analysis method | Appendix F      |

II. MATERIAL AND METHOD

There are eight phases of research methodology (Fig. 1). We first conduct phase named literature review to get the understanding of metadata schema and its current research. Since the goal of this study is to obtain metadata element to the better document, manage, and disseminate traditional knowledge, we adopted the metadata schema development methods (i.e., domain analysis, derivation analysis, system-centric analysis, user-centric analysis and resource-centric analysis) as part of research methodology. Then, we finalize metadata schema and write-down conclusion.

![Methodology phase diagram](image)

The first phase is the literature review. In this phase, we limit our research to find literature about metadata schema, traditional knowledge, and cultural heritage. After the application of these criteria, we found selected articles for metadata schema definition, traditional knowledge definition, metadata schema development methods and related works.

The second phase is domain analysis. This step is to identify the domain of resource, type of resource, end user group and their activity in using the resource. Domain analysis has been done by using document analysis and interview to related stakeholder to get insight related to traditional knowledge.

The third phase is derivation analysis. This phase is to identify candidate of elements from related existing metadata. Related metadata schema for derivation analysis was selected based on domain and purpose of metadata.

The fourth phase is system-centric analysis. In this phase, we review existing system which related to our proposed metadata schema. The next phase is user-centric analysis. This phase has been done through interviews have been conducted on individuals and communities from the public owners of traditional knowledge in the several cities in Indonesia.

The sixth phase is resource-centric analysis. We collected data about traditional knowledge in several formats. We studied the provided information of collected data that will be described by metadata schema. The seventh phase is metadata finalization. All identified elements were mapped and grouped into several sections. Then, the last phase is the conclusion.

III. RESULT AND DISCUSSION

In domain analysis phase, we collected and analysed secondary data and primary data. Secondary data are the related documents with traditional knowledge included national policy, report of traditional knowledge research, etc. Primary data are collected by interviewing experts in
traditional knowledge. In this research, we interviewed two experts in traditional knowledge field who have research experiences over 10 years. The purpose of domain analysis is to clearly determine domain, end users and their related activities in utilizing digital resources of traditional knowledge. The summary of domain analysis was presented in Table 1 below.

| Key Point          | Description                                                      |
|--------------------|-------------------------------------------------------------------|
| Domain             | Traditional knowledge; cultural heritage; indigenous knowledge   |
| Type of document   | Articles                                                          |
| End-user group     | Traditional knowledge researcher group, public, government, traditional knowledge owner |
| Activity of end-user group | Government: looking for traditional knowledge situation based on research data for making government policy to support traditional knowledge preservation. |
|                     | Traditional knowledge researcher group: conducting research to gather information regarding traditional knowledge in targeted society or ethnic group  |
|                     | Public: looking for and learn traditional knowledge               |
|                     | Traditional knowledge owner: providing traditional knowledge based on elder indigenous people and self-experiences. |

Based on the result of domain analysis before, we identified primary data and secondary data for derivation analysis, system-centric analysis, user-centric analysis and resource-centric analysis. The brief result information for each phase is depicted in Fig. 2 below.

In derivation analysis, we reviewed related metadata schema that has been developed by Yang and Chan (2000), Ning et al. (2011), Ronzino et al. (2012), Ye & Zhou (2013), Qobumo et al. (2013) and identified 389 metadata element candidates. In the system-centric analysis, we reviewed systems named Plant Resources South East Asia (PROSEA), Perlindungan Tanaman Obat dan Pengobatan Tradisional or Medicinal Plant and Traditional Medication Protection (LINSTRAD), Indonesian Scientific Journal Database (ISJD), Indonesian Science and Technology Digital Library (ISTDL) and identified 57 metadata element candidates. In the user-centric analysis, we interviewed individuals and communities of traditional knowledge owners in the several cities in Indonesia that are Bengkulu, Gresik, Padang, and Bandung and identified 30 metadata element candidates. Then, in the resource-centric analysis, we reviewed documentation of traditional knowledge that has gathered from PROSEA, LINSTRAD, ISJD, and ISTDL and identified 50 metadata element candidates. As a final metadata schema, we found 37 metadata elements that related to traditional knowledge domain as depicted in Fig. 3.
Based on Fig. 3 above, there are 37 metadata elements that divided into six metadata sections. The detailed about six metadata sections are elaborated below.

1) **Knowledge Engineer:** describes people who conducted survey and interview with traditional knowledge owner.

2) **Knowledge Source:** describes people who owned traditional knowledge.

3) **Material:** describes the needed material that involves conducting guidance from traditional knowledge.

4) **Supporting Tool:** describes the needed tool that use to process material in order to achieve the goal of traditional knowledge.

5) **Success Story:** describes the experience of people who successfully tried the traditional knowledge.

6) **Supporting Data:** describes information of data source that use to explain traditional knowledge in more detail.

Then, every single metadata section consists of several metadata elements. The detail of metadata element can be seen in Table 2 below.

| Section          | Field                     | Description                                                                 |
|------------------|---------------------------|-----------------------------------------------------------------------------|
| Supporting Data  | Supporting data ID        | Identity number of supporting data of traditional knowledge                  |
|                  | Video source              | Call number of the video version of data                                    |
|                  | Audio source              | Call number of the video version of data                                    |
|                  | Transcript source         | Call number of the interview transcript as result of the traditional knowledge acquisition |
|                  | Date created              | Date created of supporting data collection                                 |
|                  | Date modified             | Date modified of supporting data collection                                 |
|                  | Collector                 | Users who collected supporting data collection                             |
| Material         | Material ID               | Identity number of material collection                                      |
|                  | Regional name             | Regional name of material                                                   |
|                  | Scientific name           | Scientific name of material                                                 |
|                  | Image source              | Call number of the photo version of material                                |
| Supporting Tool  | Supporting tool ID        | Identity number of supporting tool                                           |
|                  | Tool name                 | Name of supporting tool                                                     |
|                  | Tool specification        | Specification of supporting tool                                            |
|                  | Tool dimension            | Dimension of supporting tool (length, width, height)                        |
| Knowledge Source | Knowledge source ID       | Identity number of knowledge source                                         |
|                  | Name of traditional knowledge | Common name of traditional knowledge                                   |
|                  | Knowledge know-how        | Description of how to practice of traditional knowledge                    |
|                  | Knowledge owner           | Owner detail of traditional knowledge                                       |
|                  | Photo of knowledge owner  | Photos of owner of traditional knowledge                                    |
|                  | Location of knowledge owner | Latitude and longitude coordinate                                           |
|                  | History of knowledge source | How the owners of traditional knowledge gain traditional knowledge      |
|                  | Date of knowledge surveyed | Date of survey conducted by knowledge engineer                           |
|                  | Assignment letter number  | Number of assignment letter conducted survey (if applicable)               |
|                  | Date issued               | Data of knowledge saved to system                                           |
|                  | Date modified             | Date of digital knowledge modified                                          |
|                  | Modified by               | Users who modified digital knowledge                                        |
| Knowledge Engineer | Surveyor ID              | Identity number of surveyor                                                 |
|                  | Name of surveyor          | Real name of surveyor                                                       |
|                  | Phone number              | Telephone number                                                            |
|                  | Brief personal information | Information of surveyor                                                     |

### IV. CONCLUSION

This research gives contributions in developing metadata element set which can be used as basic information architecture to document, manage, and disseminate traditional knowledge.

After conducting domain-analysis, derivation analysis, system-centric analysis, user-centric analysis, and resource-centric analysis for traditional knowledge domain, a metadata element set was proposed. This new schema is designed for the description of metadata for traditional knowledge domain. The purposed metadata as many 37 metadata elements are categorized into 6 metadata sections,
i.e., supporting data, material, supporting tool, success story, knowledge source, and knowledge engineer.

As future research, the obtained metadata element set will implement to the system named I-Grest (Indonesian Genesic Resources and Traditional Knowledge).

ACKNOWLEDGMENT

The authors thank University of Indonesia and Ministry of Research Technology and Higher Education for facilitating research with a research grant of university excellent in 2015-2016.

REFERENCE

[1] T. Gedif and H. J. Hahn. “The use of medicinal plants in self-care in rural central Ethiopia,” J. Ethnopharmacol., vol. 87, no. 2–3, pp. 155–161, 2003.
[2] P. M. Shrestha and S. S. Dhillion. “Medicinal plant diversity and use in the highlands of Dolakha district, Nepal,” J. Ethnopharmacol., vol. 86, no. 1, pp. 81–96, 2003.
[3] H. Noprisson, E. Hidayat, and N. Zulkarnain. “A Preliminary Study of Modelling Interconnected Systems Initiatives for Preserving Indigenous Knowledge in Indonesia,” in 2015 International Conference on Information Technology Systems and Innovation (ICITSI), 2015, pp. 1–6.
[4] E. K. Damayanti, E. A. Zuhud, H. M. Sangat, and T. Permanasari. “Pemanfaatan Dokumentasi Pengetahuan Lokal Tumbuhan Obat Untuk Mewujudkan Masyarakat Mandiri Kesehatan (Utilizing Local Knowledge of Medicinal Plant Documentation toward Health-Independent Community),” in Seminar Nasional Etnobotani IV (National Conference on Ethnobotany IV), 2009, pp. 239–244.
[5] I. Nurlahida, R. Manurung, and A. M. Arymurthy. “Performance comparison analysis features extraction methods for batik recognition,” in International Conference on Advanced Computer Science and Information Systems (ICACISIS), 2012.
[6] D. Fitriasah, A. N. Hidayanto, R. A. Zen, and A. M. Arymurthy. “APDATI: E-Fishing Logbook for Integrated Tuna Fishing Data Management,” J. Theor. Appl. Inf. Technol., vol. 75, no. 2, 2015.
[7] T. C. Chao, “Identifying Description Indicators for Research Data from Scientific Journal Publications,” iConference 2014 Proc., 2014.
[8] X. Mu, J. Jordan, W. Jeong, and J. Plessis. “Towards a Combined Approach to Metadata in Corporate Knowledge Management: A Case Study of GE Healthcare,” J. Inf. Knowl. Manag., vol. 8, no. 3, pp. 189–200, 2009.
[9] C. C. Yang and W. W. Chan. “Metadata Design for Chinese Medicine Digital Library Using XML,” in Proceedings of the 33rd Hawaii International Conference on System Sciences, 2000.
[10] T. Yu, M. Cui, H. Li, Y. Zhao, and Z. Zhu, “Traditional Chinese medicine literature metadata: A draft technical specification developed by the International Organization for Standardization,” in Frontier and Future Development of Information Technology in Medicine and Education, Lecture Notes in Electrical Engineering 269, Springer Science+Business Media Dordrecht, 2014, pp. 413–420.
[11] J. Greenberg and T. Severiens, “DCMI-Tools: Ontologies for Digital Application Description,” in Openness in Digital Publishing: Awareness, Discovery and Access - Proceedings of the 11th International Conference on Electronic Publishing, 2007, pp. 437–444.
[12] H. Noprison, D. I. Sensuse, Y. G. Sucarylne, and Lukman. “Metadata Development for Ethnophytomedicine Research Using Metadata Analysis Approach,” in The 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE 2016), 2016.
[13] K. Lang and M. Burnett. “XML, metadata and efficient knowledge discovery,” Knowledge-Based Syst., vol. 13, no. 5, pp. 321–331, 2000.
[14] S. Higgins, “What are Metadata Standards,” Digital Curation Centre, University of Edinburgh, 2007. [Online]. Available: http://www.dcc.ac.uk/resources/briefing-papers/standards-watch-papers/what-are-metadata-standards. [Accessed: 12-Dec-2015].
[15] K. M. Hüner, B. Otto, and H. Osterle, “Collaborative management of business metadata,” Int. J. Inf. Manage., vol. 31, pp. 366–373, 2011.
[16] C. Y. Lee. “A knowledge management scheme for meta-data: an information structure graph,” Decis. Support Syst., vol. 36, pp. 341–354, 2004.
[17] M. White. “Metadata - Just for Librarians Anymore,” EContent, vol. 28, no. 5, 2005.
[18] R. C. Ryser. Indigenous Peoples and Traditional Knowledge. Berkshire Publishing Company, 2011.
[19] L. Hwaskasi, E. Luna, Syamsulidik, and R. Shaw. Local & indigenous knowledge for community resilience: Hydro-meteorological disaster risk reduction and climate change adaptation in coastal and small island communities. Jakarta: UNESCO, 2014.
[20] A. M. Kaniki and M. E. K. Mphahlele, “Indigenous knowledge for the benefit of all: can knowledge management principles be used effectively?,” African J. Libr. Inf. Sci., vol. 68, no. 1, pp. 1–15, 2002.
[21] B. Hjørland and H. Albrechtsen. “Toward A New Horizon in Information-Science - Domain-Analysis,” J. Am. Soc. Inf. Sci., vol. 46, no. 6, pp. 400–425, 1995.
[22] E. Khazraee. “Domain Specific Considerations of Metadata for Cultural Heritage,” in International Conference on Dublin Core and Metadata Applications, 2011, pp. 189–191.
[23] B. Hjørland and J. Hartelt. “Afterword: Ontological, epistemological and sociological dimensions of domains,” Knowl. Organ., vol. 30, no. 3–4, pp. 239–245, 2003.
[24] L. M. Chan and M. L. Zeng. “Metadata Interoperability and Standardization – A Study of Methodology Part I: Achieving Interoperability at the Schema Level,” D-Lib Mag. 12, 2006.
[25] S. Mohd Talkah, L. Zulkiflee, and M. Shahir Shamsir. “Development and deployment of an ethno botanical and phytochemical knowledge database of Malaysia,” J. Teknol. (Sciences Eng.), vol. 64, no. 2, pp. 79–84, 2013.
[26] P. Johnston. “Metadata and Interoperability in a Complex World,” 2003. [Online]. Available: http://www.ariadne.ac.uk/issue37/dc-2003-rt. [Accessed: 17-Dec-2015].
[27] DELOS Working Group on Registries. “Principles of metadata registries: A White Paper of the DELOS Working Group on Registries,” 2002.
[28] P. Roberts and W. Hayes. “Information Needs and the Role of Text Mining in Drug Development,” in Pacific Symposium on Biocomputing, 2008, pp. 592–603.
[29] T. C. Chao, “Mapping Methods Metadata for Research Data,” Int. J. Digit. Curation, vol. 10, no. 1, pp. 82–94, 2015.
[30] Y. Y. Ning, W. K. Hua, C. H. Ming, and H. W. Shan. “The Standard of Management and Application of Cultural Heritage Documentation,” Geo-Informationes FICE CIV, vol. 6, no. 1, pp. 354–363, 2011.
[31] P. Ronzino, S. Hermon, and F. Niccolucci, “A Metadata Schema for Cultural Heritage Documentation,” in Electronic Imaging & the Visual Arts: EVA, Firenze : Firenze University Press, 2012, pp. 36–41.
[32] Y. Ye, P., & Zhou, “The Framework of Chinese Intangible Cultural Heritages Metadata Standards,” in International Workshop on Computer Science in Sports (IWCSS), 2013.
[33] B. Qobumo, G. Cuixiao, and Y. Hubin, “Customizing Discipline-based Metadata Standards for Digital Preservation of Living Epic Traditions in China,” in Digital Heritage International Congress (DigitalHeritage), 2013, pp. 145–151.
[34] P. Ye and Y. Zhou, “The Framework of Chinese Intangible Cultural Heritages Metadata Standards,” in International Workshop on Computer Science in Sports (IWCSS), 2013.
Appendix A: Metadata Schema for Chinese Medicine Digital Library [9]
Appendix B: Metadata Schema for Cultural Heritage Documentation [30]
Appendix C: Metadata Schema for Cultural Heritage Documentation [31]
Appendix D: Metadata Schema for Chinese Intangible Cultural Heritages [34]
Appendix E: Metadata Schema for Living Epic Traditions in China [33]
Appendix F: Metadata Schema for Traditional Chinese Medicine Literature [10]