

Analysis and Design of Law and Human Rights Research Information Systems for Evidence-Based Policy Making

Machyudhie¹, Rudi Trisno Yuwono²

¹Master Student, Faculty of Technology and Engineering, Business Information System, University of Gunadarma, Indonesia

²Faculty of Technology and Engineering, University of Gunadarma, Indonesia

*Corresponding author: macyudh @ gmail.com

Abstract—Evidence-based policy making is an important factor for an organization to achieve its objectives. However, in practice there are still policies produced that have not been fully based on research, a strong knowledge and database, due to the difficulty of mapping all community needs and the absence of an adequate database and information to identify issues. This situation causes the policies taken can leave several problems and recurring cases. Current developments in information technology with the phenomenon of big data, information that can be widely obtained through the internet using the method of data scraping and extraction of information that will be stored in a database, it is considered important to develop a Law and Human Rights Research Information System by using the System Development Life Cycle (SDLC) methodology approach and Unified Modeling Language (UML) so that the analysis and design of the system can provide a more accurate description and by needs, namely supporting evidence-based policy making in the field of law and human rights (evidence-based policy making), support the improvement of research quality with adequate data; and provide information about law and human rights issues and public services.

Keywords— System, Analysis, Design, SDLC, UML, Scraper, NER.

I. INTRODUCTION

Evidence-based policy making is an important factor for an organization to achieve its objectives, so that the resulting policy can be based on a strong knowledge and database [1]. The problems of Indonesia's vast territory with a scattered population make it difficult for policy makers to map out all the community needs. In addition, it was found that the results of the research had not represented an update of the problems in the community, so that an adequate database and information was needed to identify and inventory existing law and human rights issues.

The Agency for Law dan Human Rights Research and Development seeks to enhance its role to provide an adequate database and information to identify and inventory law and human rights issues. Evidence based data can include expert knowledge, published research or the data from the internet [2]. The database must be relevant, representative and valid requirements [3].

With the development of information technology and the phenomenon of big data, information can be obtained through the internet, using web scraping methods to get raw data from various sources of information, especially from online media [4]. Named Entity Recognition (NER) method is used to identify entities and classify entities so that news or events can

be categorized according to the compilation of issues datasets. [5].

Based on these needs, it is important to develop a law and Human Rights Research Information System into a tool to inventory and identify issues in the field of law and human rights and public services more comprehensively. This system was developed as a database service with the aim of supporting evidence-based policy making in the field of law and human rights and improving the quality of research with adequate data.

II. METHODS

A. Research Methods

The method used in this research was descriptive qualitative method. The following is an overview of the stages of the research conducted.

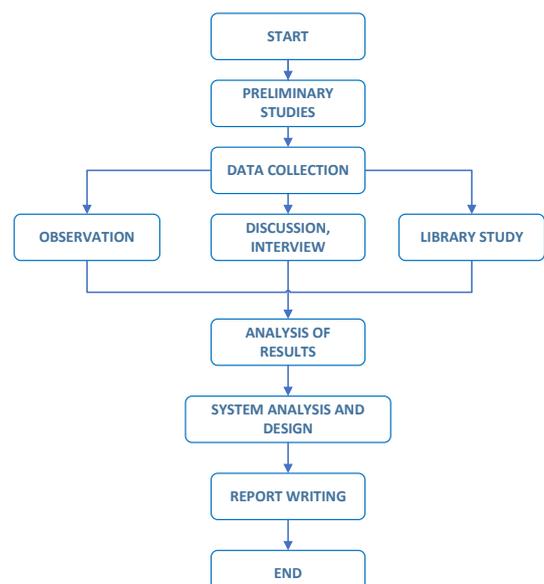


Fig. 1. Research Stages

The research was begun by identifying existing problems. Data collection was obtained through observation, discussion and interviews, and literature study. Then the results were analyzed, followed by system analysis and design using the System Development Life Cycle (SDLC) methodology and Unified Modeling Language (UML).

B. System Design Methodology

Comprehensive system design methodology with structured analysis using SDLC to plan and manage the system development process. In the waterfall model, the results of each phase are called products that can be sent, or the final product, which is continued to the next phase [6].

The SDLC model usually includes five steps, including systems planning, systems analyst, system design, system implementation, systems support and security.

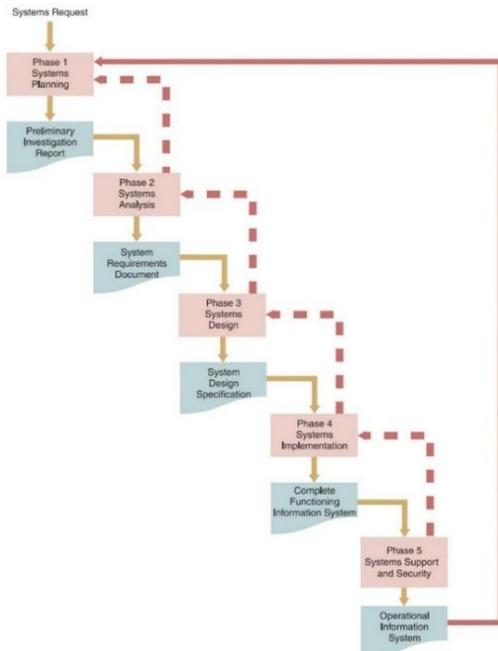


Fig. 2. The phases and results of SDLC are shown in the waterfall model [6]

Systems modeling using UML is a standard approach for modeling information systems using graphics, symbols, and notations to improve communication and clarity [7].

III. RESULT AND DISCUSSION

A. Problem Analysis

Policy making that has not been based on a strong knowledge and data can be caused by :

- The results of the research have not represented an update issues in the community and have not yet met the needs of stakeholders.
- Indonesia's geographical condition with its wide population distribution makes it difficult to map out all the needs of the community.
- The support element of the policy making process that uses the results of research, case studies, field experiments, and surveys will require a longer time and longer stages to produce a recommendation for policy making.
- Limited availability of relevant, representative and valid databases and information on law and human rights issues and public services.
- evidence-based policy making has not yet utilized information technology optimally

B. Opportunities in Evidence-Based Policy Making with the Utilization of Information Technology

Evidence-based policy making can be optimized by using information technology with web scraping and NER methods.

The database must be relevant, representative and valid requirements. First, it is relevant which means the database must be related to stakeholder functions.

Second, being representative means being able to represent real conditions in the community, updating and sourced from what is happening in the community.

Third, the data must be valid, which means that online and social media data sources are registered and can be trusted.

C. System Requirements Analysis

The system requirements are:

- The need to know, collect and analyze data on law and human rights issues for evidence-based policy making, where the source of the data is obtained through:
 - Online media monitoring and scraping/crawling.
 - Exchange data with other organizations, both Government Agencies or NGOs engaged in Law and Human Rights.
 - Direct data input by officers.
- The need to find out, collect, and analyze online media coverage, social media, reports related to public services, where data sources are obtained by:
 - Monitoring and scraping/crawling of social media and online media
 - Exchange data and information with other systems.
 - Direct data input by officers.

D. System Design

1. System Architecture

According to the system requirements analysis, system architecture can be designed as follows:

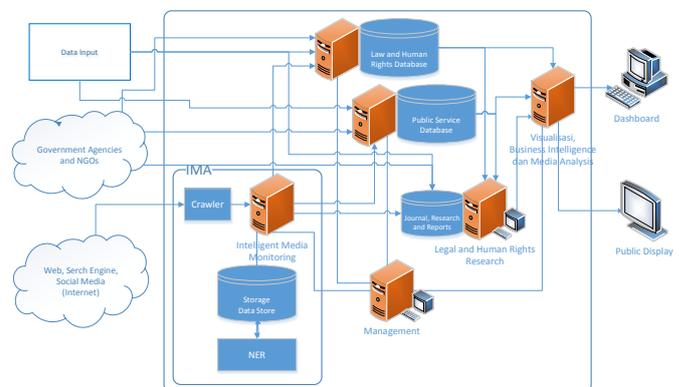


Fig. 3. System Architecture

The system architecture of Figure 3 has the following main components and processes:

- Data Source:** having data sources that include input data manually, provided for special conditions such as conflicts or disasters. Collecting content from government agencies

and non-governmental organizations, the web, search engines, and social media.

- Intelligent Media Monitoring (Scraping and NER):
 - Web scraping collects all information content that is on the web and social media automatically. Then extract the information needed based on the extraction of online media content into database storage, including URL address, news title, publication time, author, and news content.
 - NER is used to make an introduction to the entities of each word in the headline and news content. This introduction will make it easier for news or event to be categorized according to a dataset of law and human rights issues and public services. A series of name entity processes are also carried out for tokenization functions, sentence separation, word-level feature extraction, POS tagging, sentence-level feature extraction, search list, and classification.
- Data Storage: It has 3 databases which include the extraction of online media content, a database of law and human rights, and a database of public services.
- Data Analysis (Data Visualization): Analysis of the data presented in the form of a dashboard that is useful in supporting a policy.

2. Use Case Diagrams

Figure 4 illustrates the design of a use case model various types of user interactions, processes and systems

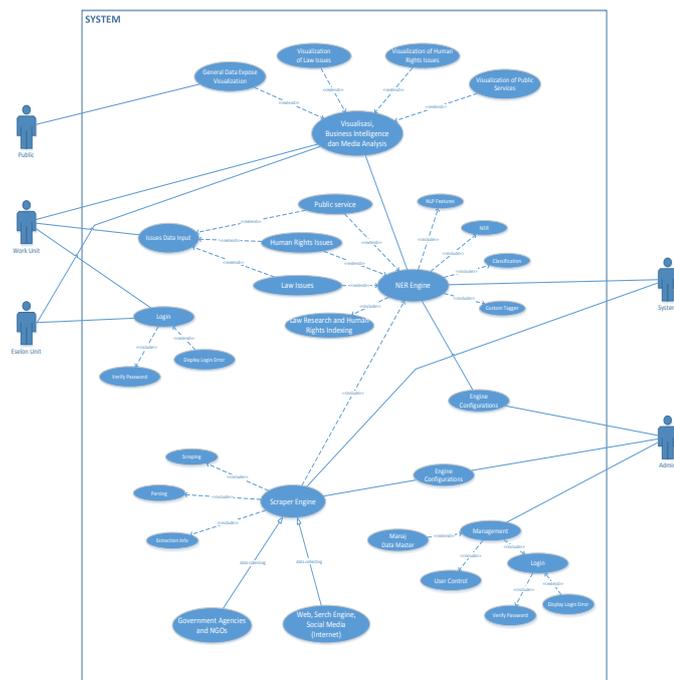


Fig. 4. Use Case System

3. Activity Diagram

Activity diagram modeling illustrates the flow of an activity to other activities in a system. The following is an activity diagram on the system which includes Scrapper Engine activity, NER Engine activity, and visual data.

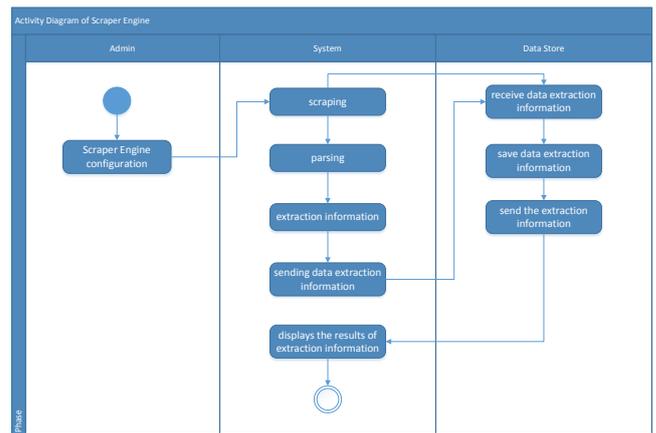


Fig. 5. Activity Diagram of Scrapper Engine

Figure 5 explains the Scaper Engine activities doing web scraping collecting information web content on the web and social media that results in the extraction of online media content information. The process starts with the admin configuring a scraper engine that discusses media configuration, parsing configuration, scraping period, and other features. Then the system carries out web scraping, parsing, extraction information, and sending it to the database. The database receives and stores the results of the extraction and the system displays the results of the extraction of information.

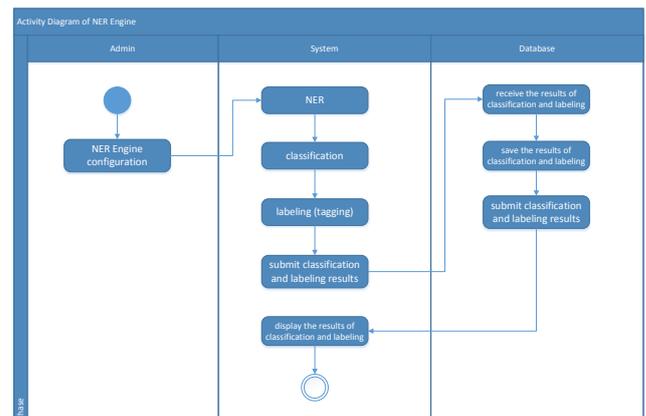


Fig. 6. Activity Diagram of NER Engine

Figure 6 explains the activity of the NER Engine which introduces the entities detected in the news media and classifies according to the dataset that has been compiled. The process starts with the admin configuring the NER engine including the NER data configuration and classification configuration. Then the system introduces NER entities, classifies data, and embeds tagging on online media data. The database receives and stores classification and labeling results and the system displays the results.

REFERENCES

- [1] A. Ramdhani and M. A. Ramdhani, "Konsep Umum Pelaksanaan Kebijakan Publik," *J. Publik J. Ilm. Bid. Ilmu Adm. Negara*, vol. 11, no. 1, pp. 1–12, 2017.
- [2] G. Marston and R. Watts, "Tampering with the Evidence: A Critical Appraisal of Evidence Based Policy," *Draw. Board an Aust. Rev. Public Aff.*, vol. 3, no. 3, pp. 143–163, 2003.
- [3] P. P. Bachtiar, "Bridging Research and Policy Through Evidence-Based Policy Advocacy," no. 32, 2011.
- [4] A. V. Saurkar, K. G. Pathare, and S. A. Gode, "An Overview On Web Scraping Techniques And Tools," *Int. J. Futur. Revolut. Comput. Sci. Commun. Eng.*, pp. 363–367, 2018.
- [5] Z. S. Abdallah, M. Carman, and G. Haffari, "Multi-domain evaluation framework for named entity recognition tools," *Comput. Speech Lang.*, vol. 43, pp. 34–55, 2017.
- [6] G. B. Shelly and H. J. Rosenblatt, *Systems Analysis and Design, Ninth Edition*. 2012.
- [7] P. Wallace, *Introduction to Information Systems, 2nd Edition*. New Jersey: Pearson Education, Inc, 2015.