PERANCANGAN SISTEM INFORMASI PENGGELOLAAN DATA PENELITIAN

DESIGN OF INFORMATION SYSTEM FOR RESEARCH DATA MANAGEMENT

[1]Onny Rafizan, [2]Vidyantina Heppy Anandhita

[1]Puslitbang Aptika-IKP, Badan Litbang SDM Kementerian Kominfo
[2] Puslitbang SDPPPI, Badan litbang SDM kementerian Kominfo
[1]*onny.rafizan@kominfo.go.id*

Abstrak

Pengelolaan data penelitian merupakan hal yang penting bagi setiap lembaga penelitian. Tidak adanya aktivitas pengelolaan data dan penyimpanan data terpusat dalam lembaga penelitian berpotensi membuat data hasil penelitian tersebut hilang maupun tidak dapat digunakan kembali. Sistem informasi pengelolaan data dibutuhkan oleh lembaga penelitian untuk dapat mengelola data secara terpusat dengan memperhatikan proses bisnis penelitian dan kebutuhan terhadap proses penyimpanan dan sharing data penelitian. Pengumpulan data primer melalui wawancara kepada narasumber teknis yang memiliki kepakaran dan pengalaman langsung berkaitan pada pengelolaan data penelitian, serta validasi kredibilitas data dilakukan berjenjang ke jabatan peneliti yang lebih tinggi (madya dan utama) serta kepada pejabat struktural yang memfasilitasi kegiatan penelitian untuk mendapatkan pemodelan bisnis dan user requirement. Pemodelan sistem dilakukan dengan UML *(Unified Modeling Language).* Hasil dari penelitian ini adalah rancangan sistem informasi pengelolaan data yang selaras dengan kebutuhan proses bisnis Badan Penelitian dan Pengembangan SDM Kementerian Kominfo untuk memfasilitasi aktivitas penyimpanan, pelestarian dan berbagi akses data penelitian

Kata kunci: Pengelolaan data penelitian, Pemodelan bisnis, Perancangan sistem informasi

Abstract

Management of research data is important for every research institute. The absence of activity for centralized data management and data storage within the research institute has the potential to make the research data disappear or not reusable. The data management information system is needed by the research institute to manage data centrally by paying attention to the business process of research and requirement to the process of storage and sharing of research data. Primary data collection using interviews to the researcher and statistician with expertise and direct experience are related to the management of research data, as well as validation of data credibility are tiered to higher researcher positions (Madya and main researcher) and to structural officials who facilitate research activities to obtain business modelling and user requirement. System modelling in this study use UML (Unified Modeling Language). The result of this study is the design of information systems for research data management that align with business process in ICT Research and HR Development Agency Ministry of Communication and Information Technology (MCIT), to facilitate the activities of storage, preserve and sharing access the research data.

Key words: *Management research data, Business modelling, Design information system*

143

INTRODUCTION

Research data is the output of systematic research that involves the process of observing, experimenting, and testing hypothesis (Pryor, 2012 quoted from Koltay, 2015). Research data are valuable information assets that generated or gathered by the scientific method. However, datasets are potentially fragile, susceptible to storage failure and obsolete technology and may also sensitive, which had private information for example, so it needs to be managed with the correct steps of security (Pinfield, Cox, & Smith, 2014).

The absence of activity for centralized data management and data storage within research institution has the potential to make the valuable data disappear or cannot be found. This situation can cause many dark data spread in the research institute. Dark data is the data that cannot be indexed and carefully stored, so it makes data almost invincible by researcher and other potential users, because of that the data tend to least utilized and disappeared at the end (Heidorn, 2008). That is why the activities of centralized research data management are essential to implement into each of research institution, such as ICT Research and HR Development Agency Ministry of Communication and Information Technology (MCIT).

Higman and Pinfield (2015) mentioned that management of research data is a relationship between data sharing on the one hand with the development of policy data management research and practice at the institutional level on the other side. Tammaro and Casarosa, (2014) describe research data management as data lifecycle management that covers spectrum activities. The activities starting from planning the research data management in the early stage of the project, to gathering data as part of the process of research; identifying, processing, and accessing from a set of data; and finally, for maintaining archives and sharing data within a correct repository. .

Surkis and Read (2015) explain that research data lifecycle consist of six phases that adapted from UK Research's phase. Among that six phase, five of phase formed cycle, whereas the phase of using data are intersecting from giving access to data directly into processing data.
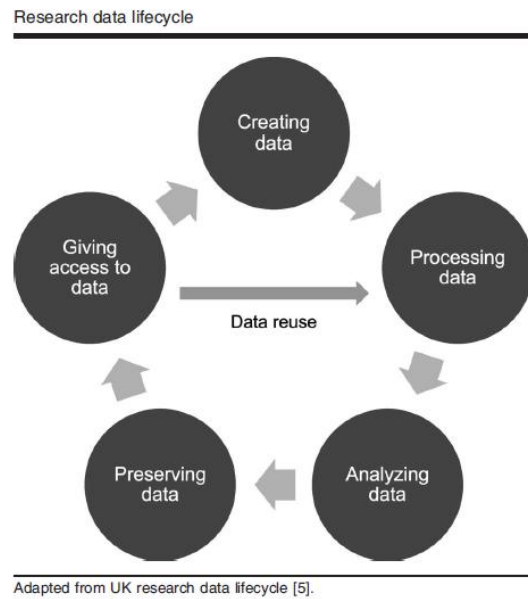

Figure 1. Research data life cycle

Source: Surkis and Read (2015)

The first three phase in the lifecycle which creating or gathering data, processing data from a raw format into another form that can be analyzed, the phase of analyzing data so the results can be distributed into another form of academic output such as an article of the journal. The three steps of that phase required management of data and transform it from raw data to the process for analyzing data and make sure that the data are describes with the correct method that can be understand. If the data can be understand, it can be reuse by another researcher that will use it for validity test from the original results or analyze it again from original data with different method. Whereas, the last three steps of data lifecycle include maintaining data after research has completed, giving access from data to another person and finally reusing the data to do new research or to check reproducibility from original results.

Sharing data are one of the important parts of the activities of research data management. Data of research results are important information assets that can be reused again for academic purpose as well for future research. Therefore, how the data are managed and shared became an important thing. According to Eynden (2013), the benefits for sharing research data are:

a. For the public: the benefit for sharing data for the public is the production of high-quality research with social value.

b. For the funders: by sharing with public, the benefits for funders is to make optimal use of

publicly funded research and avoid duplication of data collection

c. For the scientific community: sharing data with the scientific community maximizes appropriate transparency, improves the quality of verification, replication and trust, as a valuable resource for teaching, for example, methodology as well as promoting innovation - unintentional, new uses of data.

4. For the research member: sharing data with research members enable the maximum use of data/information from their contributions and minimizing unreachable data collection (sick members, data held only by elite members etc.).

The benefits from sharing data can be achieved with a good data management so the integrity, source, even the ownership of data can be accounted for. Therefore, as a research institution, The ICT Research and HR Development Agency MCIT needs some application system that can support management and storage of research data so it can be shared online.

Several previous studies related to the design of information systems data management has been done, although not specific to the management of research data. Alparisi and Bunyamin (2015) undertook the development of web-based pupil's data management information system at the Pondok Pesantren Ash-Shofi with the USDP (Unified Software Development Process) with UML (Unified Modeling Language). Kartikasari (2012) undertook the design of mail management information system at kindergarten and primary school of Pringkuku Sub-district to increase work productivity because the process of recording and reporting will be fast and accurate. Another study was conducted by Basri and Devitra (2017) to design a web-based archive information management system in General Election Commission (KPU) of Tebo District. Basri and Devitra (2017) use object-based modelling techniques, namely use case diagrams, class diagrams, and activity diagrams. The output of this research is a prototype design of web-based archive information management system at Tebo District General Election Commission (KPU) that can provide online archiving service.

Based on that problem identification and relevant studies, the research question that is submitted is how the design for information management system of research data in ICT Research and HR Development Agency Ministry of Communication and Information Technology (MCIT)? The scope of this research is designing system information with regards to the business process of research and the needs for storing and sharing of research data in the of ICT Research and HR Development Agency - MCIT.

METHOD

In the cycle of SDLC (Software Development Life Cycle), the research that the authors do are the first two steps of systems modelling, which are Business Modelling and Requirement. To decide the business model that needed in the design for Information Management System of research data, gathering primary data with a qualitative approach by using interview and observation. The in-depth interview is done with the technical source person that has expertise and experience that directly involved with the management of research data, along with validating the credibility of gathering and processing results of data that tiered to higher researcher positions (Madya and Utama researcher) as well as to structural officials who facilitate research activities. Stages of qualitative data analysis in this study follow the flow of steps of qualitative data analysis prepared by Creswell (2014).

To bridge the needs of business processes with the design of information systems that will be made then modelled business processes with BPMN (Business Process Modeling Notation). This modelling aims to develop information systems that align with business processes and institutional needs.

The design of information management system of research data used unified modelling language (UML). Seidl & Scholz (2015) mention Unified Modeling Language (UML) is a consolidation of best practices that have been established for many years in the use of modelling languages. UML allows us to present very diverse aspects of the software system. Use case diagram and class diagram used in this research. Use case diagrams are used to document the functionality of existing systems and to record user-made use of functions. The contents of the use case diagram express expectations of the ownership of the system to be developed. The diagram documented the requirements that the system must meet and become important for detailed technical design. In particular, the use case diagram is intended to answer the question: What is depicted? (System), who interacts with the system? (Actors) and what can the actors do? (Use case). Class diagrams are

used to model the static structure of a system to describe the elements of the system and the
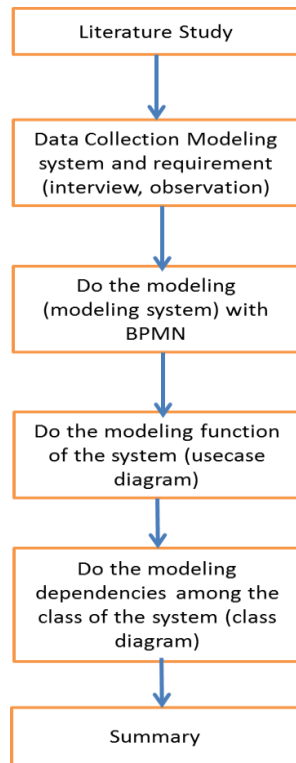

Figure 2. Research flow

ANALYSIS AND DISCUSSION

Information and important findings from interview results are processed for making business process modelling and requirement for management of research data in ICT Research & HR Development Agency. Some of the findings are in the research data lifecycle (Surkis & Read, 2015) about preservation activity (storage and preservation) centralized data of research results are not done

relationships between them. Briefly, the flow of research can be seen in the following figure:

internally yet in the ICT Research and HR Development Agency -MCIT. So that the data that had been gathered and generated throughout the research still disperse among the research team. Similarly, the activities in the phases of internal data reused during this time are done in non-formal steps. Therefore with the design of information management system of research data is expected to facilitate centralized storage activities with designated responsibility as well as documenting each request data both internal and external for data reused.

In the implementation of management of research data, it is necessary to involve coordination and role from many sides in the organization of ICT Research & HR Development Agency of Ministry of Communication and Information Technology. According to the concept of Surkis and Read, (2015), he phases of gathering research data, processing and analyzing research data are the responsibility and authority of researcher and research team; therefore those activities are on the outside of the information system that had been designed. On the preservation phase (storage and preservation of data) and giving access to the data for reused are within the requirement for this information management system of research data.

Based on the results of in-depth interviews can be obtained functional requirements that show the behaviour of the system designed (table 1)

Table 1. List of Functional Requirement

No.	Functional Requirement	Priority
1.	The system allows the data managers to perform data confirmation and control the data request process.	High
2.	The system allows the supplier of data to store data according to its classification.	High
3.	The system allows the process of data requests from internal users and external.	High
4.	The system can display the menu according to the role of each user.	Medium
5.	System capable of displaying data and graphic of research results that can be download.	Medium
6.	The system can search stored data with keywords	Low

Business Process Modeling Notation

According to the results of gathering data from the interview for the functional requirement, the business process modelling for the information management system of research data can be described in the BPMN flow with the help of Bizagi tools.

To understand the process of the system, the detail of the process can be seen in figures below.
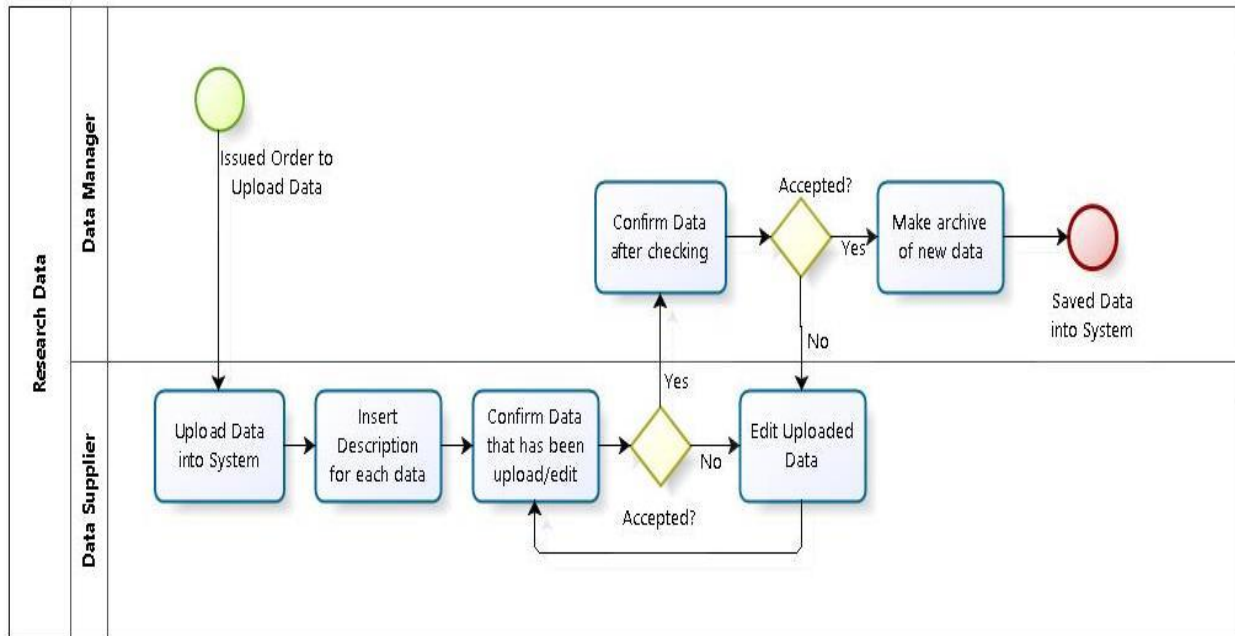

Figure 3 : Business Process for Saving Data

The process in figure 3 is the Business Process for Saving Data. At this process, it can only be started when Data Manager gives an order to Data Supplier to upload data. The role of Data supplier in this process is to upload the research data, give the description, and finally give confirmation so the data can be saved by the Data Manager. The role of Data Manager in this process is to give second confirmation before the data that had been uploaded to be saved into the system. This two-steps confirmation are used to ensure the integrity, accuracy, and quality of the data, so before the data is saved into the system, it needs to be checked twice by two different people.

After the research data has been uploaded and saved into the system, the next process it to manage the data when there is a request. At this process for requesting data, it is divided into two categories according to who requested it. It can be divided into two categories, which are **Internal** and **External**.
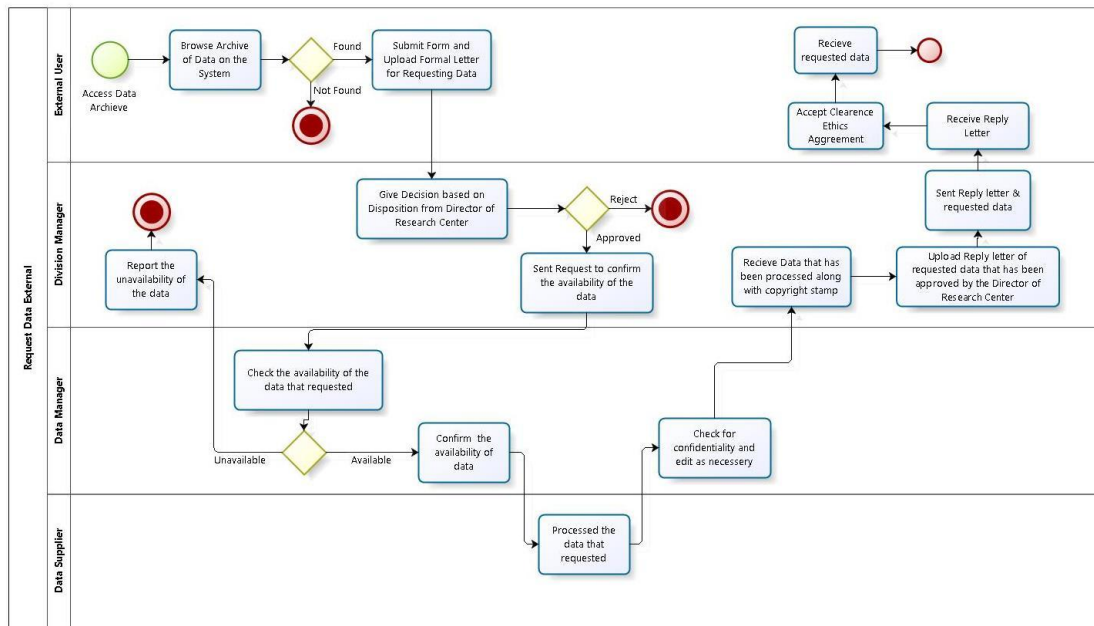
Figure 4: Business Process Request Data for Internal

The process of Requesting Research Data for the internal is start when the users were browsing the list of data that available on the archive. If the user found the data that he/she needed, they can request the data by submitting the form that available on the system. When the form had been submitted, it will be sent to Manager of the Division that owns the data that had been requested. After the Division Manager had confirmed and given permission, the system will notify the Data Manager. At this point, the Data Manager will check the requested data by confirming the order to ensure the quality and integrity of the data.

After there is confirmation from Data Manager, the Data Supplier in the division that owns the data will also give confirmation and processed the data that had been requested, since it is not allowed to send raw data. After the requested data had been processed, it will be sent to Data Manager, and he/she will put a copyright stamp into the data that had been processed and sent it to the user who requested the data. The user who requested the data can only download the data after he/she had accepted the Clearance Ethics Agreement that will be shown when they wanted to download the data.

Figure 5 : Business Process Request Data for External

The process of Requesting Data for External is a bit different with Internal. Both users can still browse the data of research that available on the archive; but when they wanted to request the data, they need to upload and submit a formal letter to the research centre that owns the data. The division manager will give a decision whether to accept or reject the request based on disposition. After the division manager gives its acceptance, the system will notify the Data Manager to check whether the data that had been requested is available or not. If the data is available, Data Manager will confirm so the system will give a task to Data Supplier (the one who upload the data). At this point, Data Supplier will process the requested data since it is not allowed to send raw data. After that, the processed data will be sent to Data Manager, and the data will be check if there is data that need to be confidential, the Data Manager will edit it as necessary. Finally, the data that had been processed will be sent to the Division Manager, and he/she will apply copyright stamp to the data. At this part, the Division Manager will also upload a reply letter that had been signed by Director of Research Center and will send it along with the data that had been requested. The same with the process on Internal, the user who requested the data can only download the data after he/she had accepted the Clearance Ethics Agreement that will be shown when they wanted to download the data.

Usecase Diagram
The functional requirements of Management System for research data can be divided into two parts, saving data and requesting data. Saving data is the process to store the research data into the system. And Request Data is the process where the Data Manager sent the research data to those who request the data. The activity for the two processes can be seen as follow.
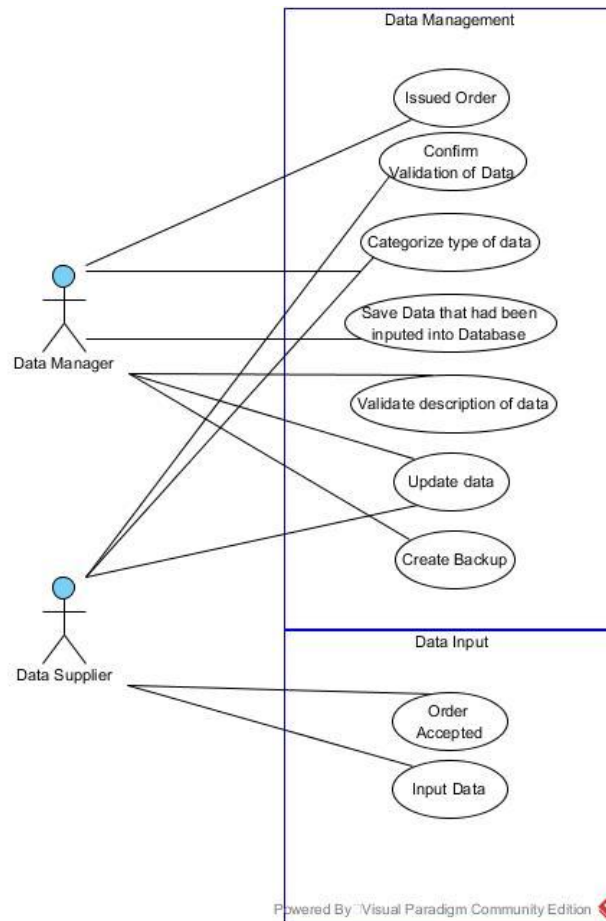
Figure 6 : Activity Diagram for Saving Data

The activities for the process to save the data into the System are done by two actors. First is the Data Supplier, he/she is the one who had been appointed by the Director of Research Center to upload the research data into the system. It is necessary for Data Supplier to be the one who also coordinates the research itself because he/she was the one who understands most of his/her research data. The Data Supplier will upload his/her research data after they had received an order from the system to upload the data. They responsible for ensuring the quality, integrity, and availability of their own data when they upload it. However, they need second confirmation before the data can be saved into the System's Database. The second actor is Data Manager, who responsible to manage the data that has or going to be saved into the system. Data Manager also needs to coordinate with other actors such as Data Supplier to ensure the quality, integrity, and availability of the data. In the process of saving data, Data Manager is the one who can save the data that had been upload into the System's Database by Data Supplier. This two-stage confirmation is necessary to ensure that second people in charge can check the data that had been upload before it has been saved into the system.
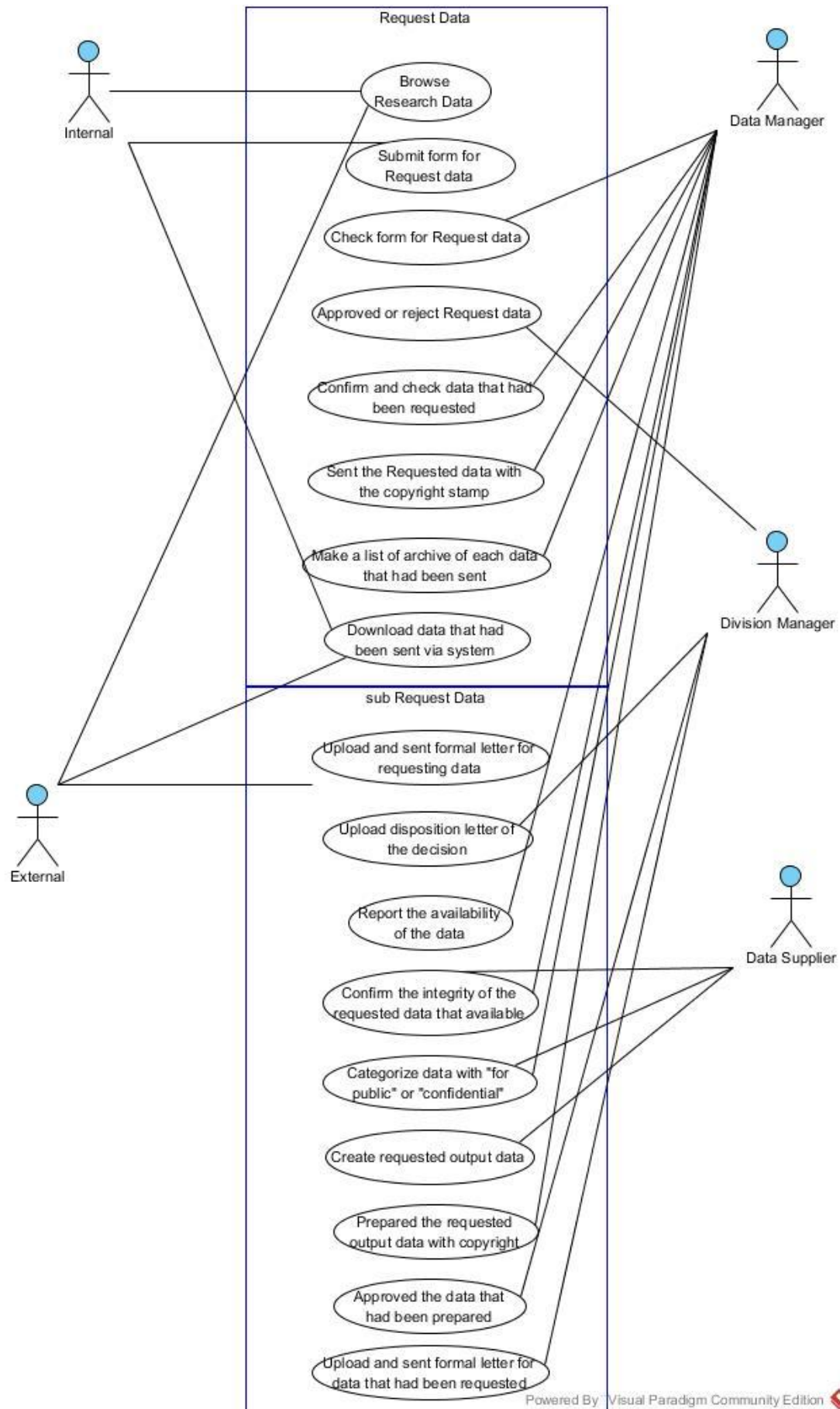
Figure 7: Activity Diagram for Request Data

Data that had been upload into the system can also be reused by other people, and this is also one part of the function of Information System Management for Research Data. The activities for the process to request data had more actors, such as Internal, External, Data Manager, Division Manager, and Data Supplier. The internal actors are the staffs or researchers inside the institution who would like to request data. They can browse the archive to see what kind of research data that available and submit form request it if they want it. The external actor is the person from outside of the Institution. They can also browse the archive to see what kind of research data that available, and however if they want to request data, they need to submit a formal letter from their institution. Data Manager in this process also the same with the previous process, someone who manage the data in the system, at this part they also need to coordinate with Data Supplier to process the requested data. If it a request from Internal, Data Manager can communicate and sent the requested data via the system. Division Manager is the person who directly

under the Director of Research Center. At this process, he/she is the one who will give the decision about the data that requested from external, based on disposition from his/her Director of Research. Division Manager also responsible for communication and sending the data via the system to the external, this process needs to be handled by a person on level Division Manager because he/she will also represent the Research Institution. The last actor is the Data Supplier. The data supplier in this process is the one who responsible to upload the requested data in the activity diagram on Figure 6. At this process, he/she will coordinate with Data Manager to process the data that had been requested before it can be sent to the Division Manager.

Class Diagram

When modelling a system, we need to understand the relationships and source code dependencies among classes in the system to better understanding what operation will affect another operation or actor. Therefore, the class diagram was design as follows:
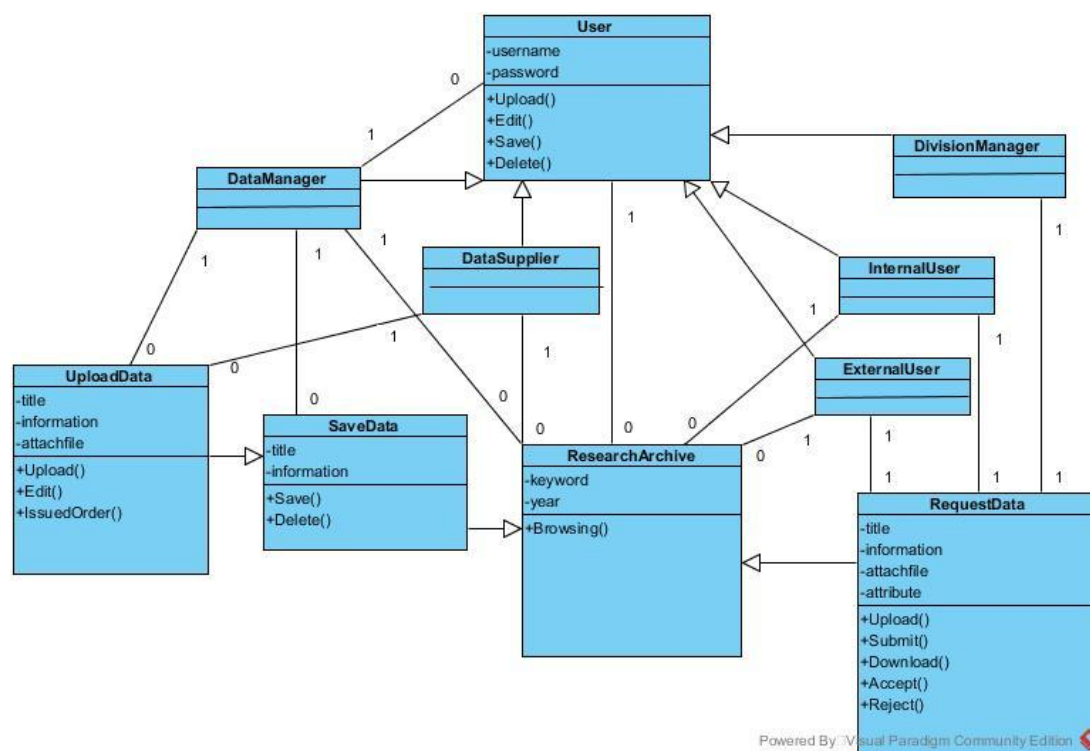

Figure 8: Class Diagram for Information Management System

The core of the system is on the Research Archive, on its own, its only function to browse the data that had been saved on the system. Therefore, this class will have dependencies with other class such as SaveData and UploadData, where both of those classes are used to adding new research data on the system which later can be viewed on ResearchArchive. As mentioned in the activity diagram previously, the function for saving data will be handled by Data Manager, while Data Supplier will be responsible for uploading the data into the system. For requesting data, it will be managed on the class of RequestData; this is where both of external and internal user will submit a request and download data (for the external user, the request data will need permission from Division Manager before it can be processed).

Prototype Information System.

Below is the view of the prototype of the designed system with main features as follow:

Figure 9. Upload Files

For the researcher in ICT Research and HR Development Agency, they can add new data such as a graph, table, pie chart, or another type of information that had been processed into the system by login with their own user account. The researchers who had been assign task from Data Manager will be able to upload the files or information on the research that they had been given task for.

Figure 10. Browse Information

The information that available on the system can be browsed from the archive based on the year that the research had been done. After the user had chosen the year that he/she want, they can click on the title of the research that they had interested in. On this steps, this list can be viewed both by internal or external users.
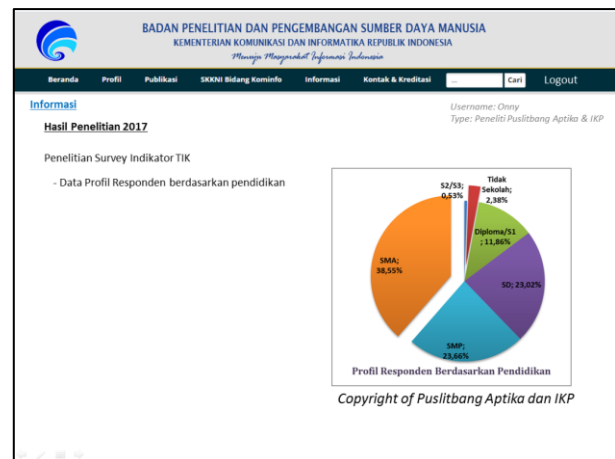

Figure 11. View Data

After the user had chosen the title of the research that he/she had interested in, they choose the data or information available that they want to see. After they choose the information they want to see, the system will present the information, at this phase this information can be seen by both internal and external users, since the data that upload into the system are for public users. The data that present in this phase will have information about who own the data. Usually, the copyright of the data belongs to the Research Division that doing the research itself.

CONCLUSION

In the process of developing the information system, the design phase is the important process that ensures the information system that will be implemented will support the business process and align with user requirements. This research intends to give design prototype of information management system of research data in ICT Research and HR Development Agency, MCIT by accommodating the requirements from its user and paying attention to the business process of the research. The data management information system is designed in order to preserve research data centrally and can also display data based on the classification of research data (public/ non-public). The system is also able to accommodate requests for research data from internal and external by online. So far, the process of application of research data is done manually so that with this system any data that requested can be recorded and processed online.

ACKNOWLEDGMENT

BIBLIOGRAPHY

Alparisi, R., Algoritma, J., Tinggi, S., Garut, T., Booch, G., Santri, P., … Language, U. M. (2015). Pengembangan Sistem Informasi Pengelolaan Data Santri di Pondok Pesantren Ash - Shofi Berbasis Web, 1–6.

Basri, B., & Devitra, J. (2017). Analisis Dan Perancangan Sistem Informasi Pengelolaan Arsip Berbasis Web (Studi Kasus: Pada Komisi Pemilihan Umum (KPU) Kabupaten Tebo). *Jurnal Manajemen Sistem Informasi*, *2*(1), 227–243.

Creswell, J. W. (2014). Research Design. *Sage*, 183–190. https://doi.org/10.4135/9781473907676

Eynden, V. V. a N. D. E. N., Team, M., & Archive, U. K. D. (2013). Data Life Cycle & Data Benefits of Managing and, (April), 24–25.

Higman, R., & Pinfield, S. (2015). Program : electronic library and information systems Article information : *Program: Electronic Library and Information Systems*, *Vol. 49*(No. 4), 364–381.

Kartikasari, S. W. (2012). Sistem Informasi Pengelolaan Surat Masuk dan Surat Keluar. *Jurnal Speed 13 FTI UNSA*, *9*(2), 82–88.

Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science*, *49*(1), 3–14. https://doi.org/10.1177/0961000615616450

P. Bryan Heidorn. (2008). Shedding Light on the Dark Data in the Long Tail of Science. *Library Trends*, *57*(2), 280–299. https://doi.org/10.1353/lib.0.0036

Pinfield, S., Cox, A. M., & Smith, J. (2014). Research data management and libraries: Relationships, activities, drivers and influences. *PLoS ONE*, *9*(12), 1–29. https://doi.org/10.1371/journal.pone.0114734

Seidl, M., & Scholz, M. (2015). *UML @ Classroom*.

Surkis, A., & Read, K. (2015). OF INTEREST * Research data management, *103*(July), 154–156. https://doi.org/10.3163/1536-5050.103.3.011

Tammaro, A. M., & Casarosa, V. (2014). Research Data Management in the curriculum: An interdisciplinary approach. In *Procedia Computer Science* (Vol. 38, pp. 138–142).https://doi.org/10.1016/j.procs.2014.10.023