A Study on the Establishment of Open Source Website for Open Smart Broadcast Platform

Sang Kwon Kim*, Deokkyu Jung*, Sang-Yun Lee*, Sun-Joong Kim*
* ETRI (Electronics and Telecommunications Research Institute), Korea
ksk@etri.re.kr, jdk30621@etri.re.kr, syllee@etri.re.kr, kimsj@etri.re.kr

***Abstract* — As the number of successful global open source software have been increased recently, a research and development software funded by the government has been directed to distribute open source. In this way, we want to represent open source website for Open Smart Broadcast Platform as part of the open smart broadcast research. This website is easy to utilize for those who want to develop smart media platform. Through the website, users can easily download the source and feedback the result.**

***Keywords* - SNS Crawler, Open Source, Open API, Open Smart Broadcast Platform, convergence type metadata**

I. INTRODUCTION

Recently, the government has proceeded to provide open source of the main research and development (R&D) software technology to face open source based R&D trend [4]. A R&D software of performance task funded by the government is activated for national public software R&D community. [6] Therefore, the government is trying to expand the private sector R&D software accomplishment and enhance the R&D technology in the software industry.

Accordingly, joint researches between ETRI and private enterprise have developed Open Smart Broadcast Platform that one of the alternative national public software R&D shared platform since 2015. Starting from R&D to strict quality assurance of software product, the Open Smart Broadcast Platform is available for the user to provide as distribution of open source.

Open Smart Broadcast Platform provides VoD clip service [3] of the scene unit Semantic clustering, Open API of convergence type metadata from script, subtitle, crowdsourcing information. The user utilizes convergence type metadata provided by Open API and contributes to Media Commerce [2] or Context Awareness [5] of advertisement service to construct an open media ecosystem. In respect of business operator or service user, Open Smart Broadcast Platform improves smart media service that creates benefit.

In terms of open smart broadcast technology, we developed SNS Crawler to extract many kinds of keyword that generates convergence type metadata from subtitle, script or non-formal data. Web crawler [7] reflected SNS Crawler. Generally, the web crawler is a kind of computer program to search automatically and systematically through the web, which is also called web indexing. In the open smart media based SNS Crawler, the web crawler is collected by social network service (SNS), blogs or news portal website. SNS Crawler extracts keyword from the search with the URL, where these are connected in original website to collect directly.

In this paper, in order to activate R&D accomplishment of open source website, we introduce SNS Crawler to distribute the open source for user to download easily, which consists construction of a website, main functions, database schema, and statistics.

II. SNS CRAWLER

SNS Crawler [1] is optional collection technique that extracts from SNS social platform based on Open Smart Broadcast Platform. This platform supports a variety of basis analysis from the big data information. For user who want to collect social big data, the SNS Crawler improves statistical or analytic method. To prove the point even further, we described basic process of SNS Crawler in Figure 1.

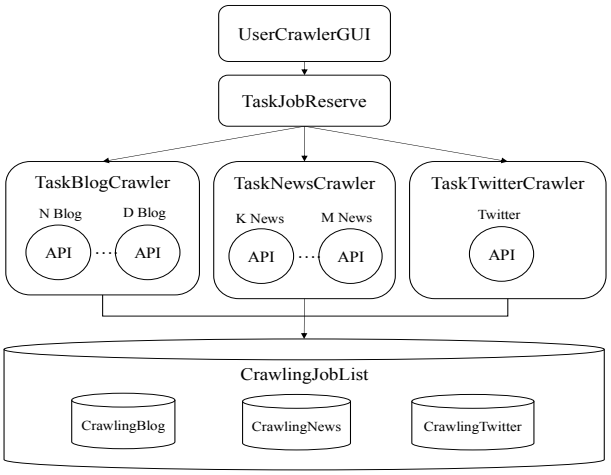

Figure 1. Job Scheduler of SNS Crawler

When user visits one of web pages and execute specific function, threads have been made by user request. These threads are working as a job reservation, which is relevant to SNS module. Crawlers have ability to schedule job threads automatically as a user collection management. SNS Crawler is made up of job scheduler. The job scheduler consists several modules inside of SNS Crawler.

- UserCrawlerGUI is practically managed by user's activation.
- TaskJobResearve is the main job scheduler controlled by automatic function from UserCrawlerGUI, which commands to three crawlers. This module constantly functions as a time scheduler.

- TaskCrawler consists three modules: TaskBlogCrawler, TaskNewsCrawler and TaskTwitterCrawler. In addition, each crawler has own APIs of portal website per SNS module.
- CrawlingJobList gathers particular data from three crawlers as a SNS collection. In addition, this module manages reserved information for collection policy including, CrawlingBlog, CrawlingNews, and CrawlingTwitter.

III. PROPOSED DESIGN

This section represents the system development environment and established open source website where everyone can access the download of SNS Crawler open source, which purposes the opening of open source and the application of R&D accomplishment.

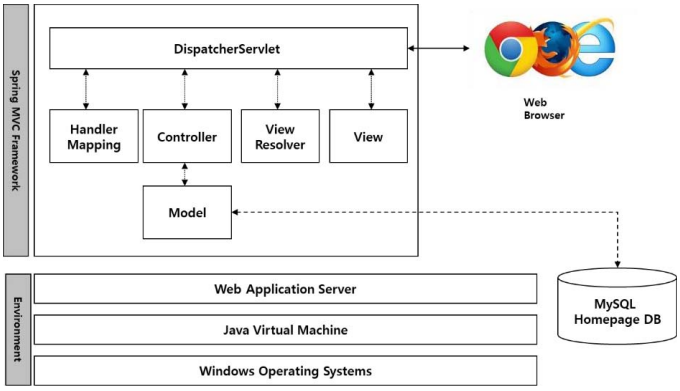

Figure 2. Development Environment of SNS Crawler Home Page

Figure 2 shows development environment for creating website to download SNS Crawler. This environment is basic open source homepage administration module connected to server and Open API in Open Smart Broadcast Platform. The main modules of web server connected with users as follows:

- DispatcherServerlet provides Controller with user request, makes adequate requests by the Controller as return values to the Viewer
- HandlerMapping decides which Controller deals request URL by users.
- Controller processes user request and send to DispatchServlet as a result. The Controller also processes Model module to utilize MySQL Homepage DB.
- ViewerResolver decides which View has to show the processing result of the Controller.
- View makes the result by ViewerResolver.

A. Overview

SNS Crawler is available to collect social big data automatically from different kinds of social media sources such as blog, twitter, news, which provides SNS RAW DATA of the various analyses, keyword based flexible collection policy and schedule management function.

Thus, we built open source website for those who want to utilize SNS Crawler developed by Open Smart Broadcast Platform, where is depicted in Figure 3. The purposes of SNS Crawler for users are the opening of open source, the comprehension of installation procedure, and the offering of system utilization.

Figure 3. Main Screen of SNS Crawler Home Page

In Figure 4, SNS Crawler registers user to process open source downloads. While the user requests download process, SNS Crawler responses the download process to the user. In view of SNS Crawler user, we designed simple procedure of open source download. Without complicated condition, the user who signed up in SNS Crawler is available to utilize the open source inside of open source download homepage.

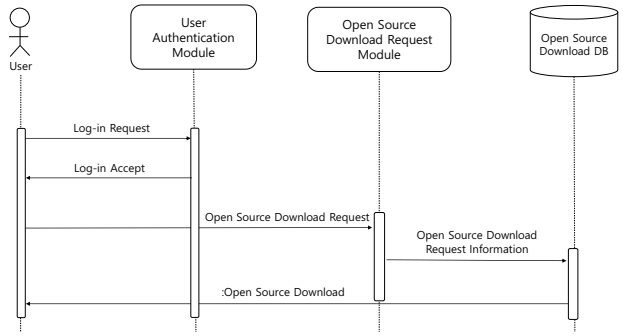

Figure 4. Flow Diagram of User Management and Open Source download Management

B. Features

SNS Crawler provides four functions of the open source website establishment as follows:

1) User Management Function

To make it easier download access about open source, we minimize the user information as ID, password, job and e-mail. The open source download is accessed only for the user login.

We also added inputs of simplified purpose of utilization for the user as Poll-up to analyze statistical data.

2) *Open Source Management Function*

Website administrator can access this function to register, modify and delete the open source. The administrator compresses open source as a zip file to distribute for the user, and maintains open source version control. The HTTP based open source management system provides user statistical data to administrator by daily, weekly, monthly and yearly.

3) *Survey and Feedback Management Function*

In order to survey the satisfaction for the user who downloaded the open source, we produce statistical function to the administrator. This function was designed as simplified poll based survey method to minimize the user access. Moreover, in the matter of feedback function, we provide the posting Q&A of user's open source on the website. Regarding the convenience of frequent questions, we classified FAQ and Q&A.

4) *Help Desk Function*

To consider users who download SNS Crawler open source in the open source home page, including user manual and short video, we define Help Desk Function to install the open source providing convenient installation environment.

C. *Database Schema*

Figure 5 shows the construction of database for open source download to establish SNS Crawler website. The relationship diagram of tables is defined as the information, which includes system administrator, users, permission to access code, management of open source & bulletin board, and access statistics from users. These Tables are also described as database structure and data, including the information of survey participation and administration through the users.

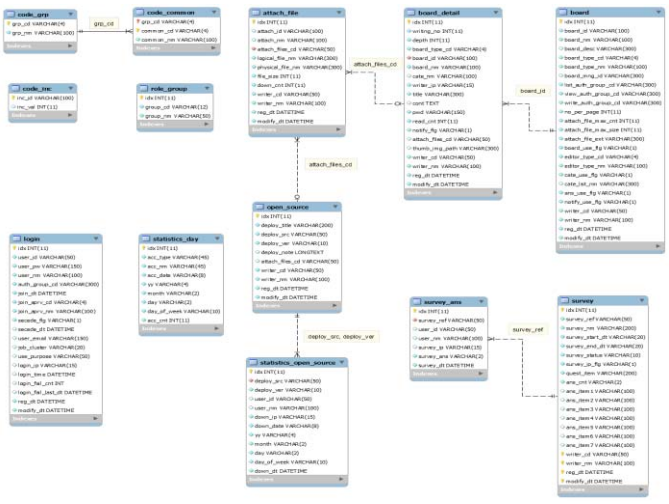

Figure 5. Website construction DB table relationship diagram

D. *Distribution and Access Statistics*

The SNS Website provides the distribution of open source and user statistics. The distribution statistics for open source shows users' functions that categorized by person's job, propose, and activity. As the survey of open source usage satisfaction and user access statistics is provided by manager, we are available to issue the user trend utilized by open source website. Additionally, web page statistics for users in various of homepage functions is provided as daily, weekly, monthly, and yearly. Figure 6 presents an example that functions page usage statistics.

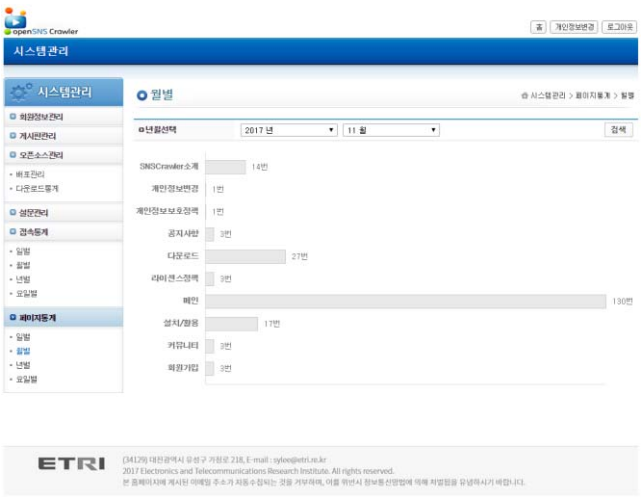

Figure 6. Monthly access statistics for SNS Crawler Home Page features

IV. CONCLUSIONS

As the number of the successful example of open source based global software has been increased, we are trying to make a better process to develop by utilizing the open source on the matter of R&D software funded by the government. As part of that, this paper purposes establishment of open source website in Open Smart Broadcast Platform. If more and more open sources are provided, it will be a great chance to develop smart broadcast media for those who eager to research smart media technology. We also anticipate that the small businesses can facilitate the open smart broadcast platform.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government. (MSIP) [No. 2015-0-00219, Development of smart broadcast service platform based on semantic cluster to build an open-media ecosystem]

REFERENCES

[1] S. Y. Lee, J. W. Son, W. J, Park and S. J. Kim, "SNS Crawler Engine for Topic Expansion," *in ICTC2016*, pp. 837–839, October 2016.

[2] Sang-Yun Lee, Jeong-Woo Son, and Sun-Joong Kim, "Design and Implementation of Media Commerce Service based on Topic", International Conference on Convergence Content 2015, pp. 249-250, 2015.12

[3] Lee, Sang-Yun, et al. "Proposal of VoD clip service based on real-time broadcasting content recognition." *Consumer Electronics-Asia (ICCE-Asia), IEEE International Conference on*. IEEE, 2016.

[4] Simon, Kimberly D. "The value of open standards and open-source software in government environments." *IBM Systems Journal* 44.2 (2005): 227-238.
[5] Raj, Lima, and Linda Sara Mathew. "Context aware multimedia crawler for dynamic encyclopaedia construction." *Computer Communication and Informatics (ICCCI), 2016 International Conference on*. IEEE, 2016.
[6] Von Krogh, Georg, Sebastian Spaeth, and Karim R. Lakhani. "Community, joining, and specialization in open source software innovation: a case study." *Research Policy* 32.7 (2003): 1217-1241.
[7] Web crawler, https://en.wikipedia.org/wiki/Web_crawler

Sang Kwon Kim received his B.S. and M.S. degrees in electronic engineering from Sungkyunkwan University, Seoul, Korea in 1983 and 1986 respectively. During 1986-1992, he had worked for Hyundai Electronics Co. Ltd., Korea, where he was participated in the development of PABX. Since 1992, he had been with the research of Transmission, Telecommunications and Broadcasting Media at ETRI where he is currently a principal researcher in Smart Media Platform Research Section. His current research interests lie in the development of smart broadcast service platform based on semantic cluster, QC, and QA in software.

Deokkyu Jung received his B.S in Computer Engineering from Korea Polytechnic University, Republic of Korea in 2014. He received his M.S. in Computer Science and Engineering from UNIST (Ulsan National Science and Technology), Republic of Korea in 2017. He joined ETRI, Republic of Korea in 2017, where he is currently research intern. His research areas are deep learning, media platform, and distributed system.

Sang-Yun Lee received his MS and Ph.D. degrees in electronics and telecommunications engineering from Hanyang University, Seoul, and Rep. of Korea in 1996 and 2007 respectively. Since 2011, he has been with Broadcasting & Telecommunications Media Research Lab., Electronics and Telecommunications Research Institute (ETRI), Daejeon, Rep. of Korea. He focuses on embedded software, smart TV software platform, and Web technology.

Sun-Joong Kim received her BS degree in computational statistics and her MS degree in computer science from Chungnam National University, Daejeon, Rep. of Korea, in 1989 and 2000 respectively. In February 1989, she joined ETRI, Daejeon, Rep. of Korea, where she is currently principal researcher and director. Her research interests include convergence service control, smart TV, content knowledge mining.