

## 데이터 수집 자동화 기반 자연재해 통계 사이트 개발

최효현\*, 김동건<sup>o</sup>

\*인하공업전문대학 컴퓨터정보과,

<sup>o</sup>인하공업전문대학 컴퓨터정보과

e-mail: hchoi@inhac.ac.kr\*, ehdrjs4502@gmail.com<sup>o</sup>

## Data Collection Automation based on Natural Disaster Statistics Site Development

Hyo Hyun Choi\*, Dong Geon Kim<sup>o</sup>

\*Dept. of Computer Science, Inha Technical College,

<sup>o</sup>Dept. of Computer Science, Inha Technical College

### ● 요약 ●

본 논문에서는 데이터 수집 자동화를 기반으로 한 자연재해 피해 통계 사이트를 구현한다. 자연재해 피해액 데이터는 국민 재난 안전 포털 사이트[1]에 자연재난 상황 통계 데이터를 사용하였다. Python 라이브러리인 Selenium을 활용하여 웹 브라우저를 제어해 자연재해 데이터를 쿼리문으로 변환 후 데이터베이스에 저장하는 자동화 프로그램을 사용한다. 또한 ncron을 활용해 6개월마다 한 번씩 자동으로 프로그램이 실행 되도록 한다. 자연재해 종류에 따른 연도별, 시도 구역 자연재해 피해액을 웹사이트로 시각화하여 보여준다. 웹사이트 구현은 React.js와 Node.js를 활용하여 구현하였다.

**키워드:** 셀레늄(Selenium), 자동화(Automation), 자연재해(Natural Disaster), 파이썬(Python)

### I. Introduction

최근 수년간 전 세계적으로 기후 이상으로 인한 자연재해 발생 빈도가 증가하면서 공공과 민간을 가리지 않고 재난 안전에 대한 관심이 높아지고 있다. 세계기상기구(WMO)가 벨기에 루비안가톨릭대 보건대 안에 있는 재난역학연구센터(CRED)의 재난통계자료를 바탕으로 분석한 최근 보고서 내용을 보면, 기후 변화로 지난 50년 동안 전 세계 재해 발생 수는 그림1과 같이 4-5배 증가한 것으로 나타났다. [2]

자연재해 피해를 예방하고자 자연재해로 인한 피해액을 사람들에게 시각화하여 보여주는 웹페이지를 구현해 연도별 자연재해 종류에 따라 어느 시도 구역에서 어떤 자연재해로 인해 피해액이 많이 발생하였는지 한 눈에 쉽게 알아볼 수 있게 보여주어 추후 발생할 자연재해 피해를 예방할 수 있는 기대효과를 가질 수 있다.

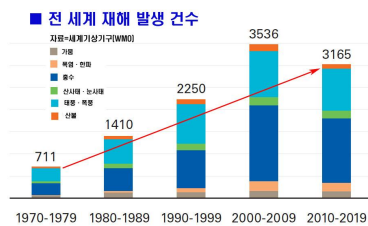


Fig. 1. Global Natural Disaster Statistics

### II. Design

본 논문에서는 Python 라이브러리인 Selenium을 활용하여 웹 브라우저를 제어해 국민 재난 안전 포털 사이트에서 자연재해 데이터를 수집하여 자동으로 DB에 저장하고 연도에 따른 자연재해 종류별 피해액을 시각화하여 보여주는 웹 사이트를 설계한다. 웹사이트는 구현은 React.js와 Node.js를 활용하여 구현한다. 구성도는 그림2과 같다.

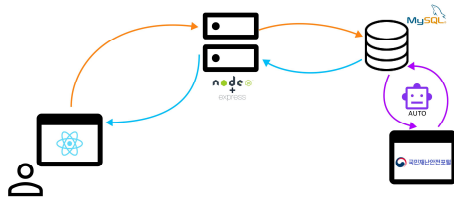


Fig. 2. System Architecture

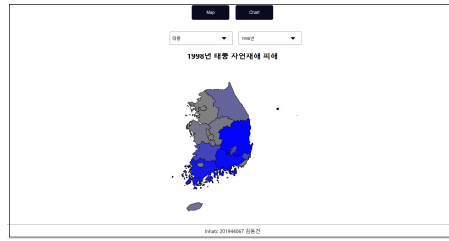


Fig. 4. Map Format Web Page

### III. Development

#### 1. The Automation based Data Collection Program

본 논문에서는 Selenium을 활용하여 데이터 수집을 자동화하여 DB에 저장하는 프로그램을 구현한다. Selenium을 채택한 이유는 국민 재난 안전 포털 사이트에서 자연재해 피해액 데이터를 수집할 때 동적인 기능을 수행하여야만 데이터를 수집할 수 있어서 Selenium을 선택하게 되었다. 자연재해 피해액 데이터를 수집하기 위해 먼저 국민 재난 안전 포털 사이트에서 엑셀 형식으로 제공하는 자연재해 피해액 데이터를 다운로드. 그 후 파일 형식을 변환 해주는 사이트들을 활용하여 최종적으로 쿼리문으로 만들어 DB에 저장되도록 한다. 자동화 기반 데이터 수집 프로그램에 Flow Chart는 그림3과 같다.

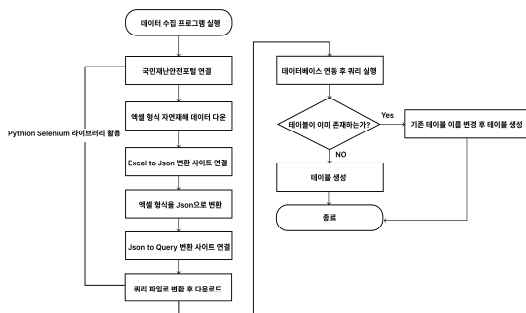


Fig. 3. Data Collection Program Flow Chart

#### 2. The Natural Disaster Damage Statistics Web Site

본 논문에서는 React.js와 Node.js를 활용하여 사용자가 선택한 연도별 자연재해 종류에 따른 시도 구역 자연재해 피해액을 DB에서 불러와 사용자들에게 시각화하여 보여준다. 시각화하여 보여주는 방식에는 지도 형식과 차트 형식이 있다. 지도 형식으로는 피해액이 높은 시도 구역은 채도를 높게 설정하여 보여주고, 피해액이 낮은 시도 구역은 채도를 낮게 설정하여 어느 시도 구역이 피해액이 높은지 한눈에 알아볼 수 있게 구현한다. 지도 형식 웹 페이지는 그림4와 같다. 차트형식으로는 피해액이 가장 높은 시도 구역부터 내림차순 정렬을 하여 피해액이 높은 시도 구역을 한눈에 알아보기 쉽도록 구현한다. 차트 형식 웹 페이지는 그림5와 같다.

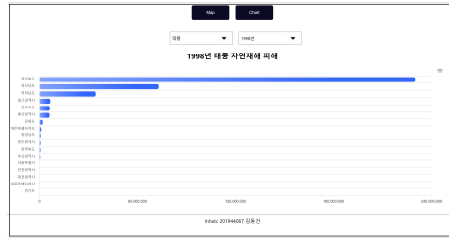


Fig. 5. Chart Format Web Page

### IV. Conclusions

본 논문에서는 데이터 수집 자동화를 기반으로 한 자연재해 피해액 통계 사이트를 구현한다. 필요한 데이터는 Python 라이브러리인 Selenium을 활용하여 웹 브라우저를 제어해 자연재해 데이터를 웹에서 수집하여 데이터베이스에 자동으로 저장하는 자동화 프로그램을 사용한다. 자연재해 종류에 따른 연도별 시도 구역 자연재해 피해액을 웹사이트로 시각화하여 보여준다. 웹사이트 구현은 React.js와 Node.js를 활용하여 구현하였다.

### REFERENCES

- [1] 국민재난안전포털 자연재난상황통계, <https://www.safekorea.go.kr/idsiSFK/neo/sfk/cs/sfc/tot/toteaiList.jsp?emgPage=Y&menuSeq=111>
- [2] "지난 50년 태풍 가뭄 등 자연재해 4~5배, 사망자수는 3배" 한겨레, 2022년 7월 20일