

# PR9 - Package & Web Scraping

김서준

2023년 11월 10일

## 1. package 기본

- R에서 기본적으로 제공하는 함수 외에 다른 프로그래머들이 필요하다고 생각해서 만들어 놓은 함수들의 모음
- 이것을 보통 library 또는 API라고 칭함
- R에서는 보통 CRAN에서 R재단이 인정한 패키지들을 공유
- 비공식적이지만 개인이 만들어서 공유한 패키지들이 있고 보통 github에 업로드 돼 있음

### 1.1. 설치

- 설치 명령어: `install.packages()`
- 다음 시간 사용할 크롤링 관련 패키지 설치

```
getwd()
```

```
## [1] "C:/UnivStudy/UnivLectures/23-2/R-programming/Works/PR9(Package & Web Scraping)"
```

```
#install.packages("rvest")  
#install.packages("httr")  
#install.packages("RSelenium")
```

### 1.2. 불러오기

- `library(package 이름)` : 설치되어 있는 패키지를 불러옴
- `require(package 이름)` : 불러오려는 시도를 하고 logical한 return 값을 반환
- `require()` 도 결과적으로 `library()` 와 동일한 기능을 하지만, 결과값에서 차이가 나게된다.

```
x <- library(rvest);x
```

```
## Warning: 패키지 'rvest'는 R 버전 4.3.2에서 작성되었습니다
```

```
## [1] "rvest"      "stats"      "graphics"   "grDevices"  "utils"      "datasets"  
## [7] "methods"   "base"
```

```
x <- require(rvest);x
```

```
## [1] TRUE
```

```
#해당 패키지가 없다고 에러가 뜨며 결과 knits 에러가 발생할 수 있는 부분.
#library(chorn)

#해당 패키지가 없더라도 에러메시지가 뜨지 않고 knits를 정상적으로 진행할 수 있다.
require(chorn)
```

```
## 필요한 패키지를 로딩중입니다: chorn
```

```
## Warning in library(package, lib.loc = lib.loc, character.only = TRUE,
## logical.return = TRUE, : 'chorn'이라고 불리는 패키지가 없습니다
```

## 1.3. 패키지 업데이트

- 패키지 개발자가 자신이 만든 패키지의 기능을 보완하거나, R버전이 업데이트 되어 호환이 되게 수정하거나 오류사항을 수정
- 명령어는 `update.packages(패키지 이름)`
- R Studio 를 사용한다면 우측 하단에 `packages` 탭이 보이는데 해당 탭 바로 밑에 `update`라는 항목이 보인다.
- 이걸 누르고 원하는 패키지 혹은 전체를 선택하고 업데이트를 누르면 됨

```
#update.packages("rvest")
```

## 1.4. 불러온 패키지 사용중지

```
#detach("packages:ggplot2", unload = T)
```

# 2. package 고급

## 2.1. package에서 함수 가져오기

- 다양한 패키지들을 불러오다 보면, 동일한 함수명을 사용해서 서로 충돌할 때가 있다.
- 그럴때는 `패키지명::함수명` 으로 어떤 패키지에서 해당 함수를 사용할 것인지 명시해서 문제를 해결할 수 있다.

```
#require(plyr)
#require(Hmisc)
#require(chron)
#require(tseries)

#plyr::summarize
#Hmisc::summarize
#chorn::is.weekend()
#tseries::is.weekend()
```

- 실습:

```
library(readxl)
```

```
## Warning: 패키지 'readxl'는 R 버전 4.3.2에서 작성되었습니다
```

```
readxl::read_xlsx("student.xlsx")
```

```
## # A tibble: 3 × 4
##   이름   나이 전공          학번
##   <chr> <dbl> <chr>      <dbl>
## 1 김서준   23 e-비즈니스학과 201921487
## 2 김도현   23 기계공학과    201921488
## 3 최민락   23 경영학과      201921489
```

## 2.3. 편리한 사용자 정의 함수

- 해당 함수는 패키지명으로 구성된 문자형 벡터를 입력으로 받아서, 설치 및 불러오기를 동시에 진행하는 함수입니다.
- 만약 해당 패키지가 이미 설치되어 있는 패키지라면 바로 설치과정을 생략하고 바로 불러오기를 진행하게 됩니다.

```
take <- function(x) {
  for (i in x) {
    if (!is.element(i, .packages(all.available = TRUE))) {
      install.packages(i)
    }
    library(i, character.only = TRUE)
  }
}
take(c("rvest", "httr"))
```

## 3. devtools

- devtools 의 주 목적은 패키지 개발에 필요한 많은 작업들을 단순화시키는 것이다. 뿐만 아니라 비공식 패키지 버전 관리의 측면에서는 상당히 중요한 패키지 이다. ## 3.1. 비공식 패키지 설치
- CRAN에서 제공하는 패키지가 아닌, github에 공개되어 있는 다양한 패키지들과 설치하여 사용할 수 있다.

```
#install.packages("devtools")
library(devtools)
```

```
## 필요한 패키지를 로딩중입니다: usethis
```

```
#install_github("패키지")
```

## 3.2. 원하는 버전 설치

- R의 패키지들이 최신 버전에 맞춰서 업데이트가 되지 않았다면, 이전 버전의 패키지를 사용해야만 하는 경우들이 있으며 이때 사용하는 것이 devtools 의 install\_version() 입니다.
- install\_version("패키지명", version = "버전명", repos = "https://cran.us.r-project.org")

- 패키지명과 어떤 버전명을 지정해주면 되며, **repos**는 어떤 서버에서 한 ◦ 내용을 다운받을지 지정해주는 것이다.

```
#library(devtools)
#remove.packages("ggplot2")
#install_version("ggplot2", version = "0.9.1", repos = "https://cran.us.r-project.org")
#packageVersion("ggplot2")
```

- 이 과정을 실습해 봤다면 ggplot2 를 지운 후 밑의 tidyverse 로 설치를 진행하면 된다.

```
# remove.packages("ggplot2")
```

## 4. tidyverse

- tidyverse 는 dplyr , tidyr , ggplot2` 등, R 프로그래밍의 핵심 패키지들을 한번에 설치 및 관리해주는 패키지이다.

```
#install.packages("tidyverse")
#install.packages("glue")
library(tidyverse)
```

```
## —— Attaching core tidyverse packages —— tidyverse 2.0.0 ——
## ✓ dplyr      1.1.3      ✓ readr      2.1.4
## ✓ forcats   1.0.0      ✓ stringr    1.5.0
## ✓ ggplot2    3.4.3      ✓ tibble     3.2.1
## ✓ lubridate 1.9.3      ✓ tidyr      1.3.0
## ✓ purrr     1.0.2
## —— Conflicts ——
——— tidyverse_conflicts() ——
## ✗ dplyr::filter()      masks stats::filter()
## ✗ readr::guess_encoding() masks rvest::guess_encoding()
## ✗ dplyr::lag()          masks stats::lag()
## ⓘ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

## 패키지 연습문제

### github에 개인 패키지 만들기

```
#install_github("selfish234/selfish234R")
library(selfish234R)
```

```
##
## 다음의 패키지를 부착합니다: 'selfish234R'
```

```
## The following object is masked from 'package:base':
##
##      double
```

```
selffish234R::double(3)
```

```
## [1] 6
```

## rvest를 이용하여 아주대 공지사항 크롤링

```
library(rvest) # rvest 패키지 로드

title <- NULL
dept <- NULL
date <- NULL
# 데이터 저장을 위한 빈 벡터 초기화

# 1, 2 페이지를 크롤링하기 위한 for 반복문
for (i in 1:2) {
  url <- "https://www.ajou.ac.kr/kr/ajou/notice.do?mode=list&&articleLimit=10&article.offset="
  #크롤링할 페이지(몇페이지인지를 나타내는 부분은 현재 없음)
  urls <- paste0(url, (i-1)*10) #몇페이지인지를 나타내주는 부분을 이전 url에 붙임
  html_source <- read_html(urls) # HTML 읽어오기

  # 공지사항 페이지에서 13번째부터 22번째 까지의 내용을 반복문을 통해 순회
  for (i in 13:22) {
    #제목 크롤링 코드
    T.selector <- paste0("#cms-content > div > div > div.bn-list-common02.type01.bn-common-cate
> table > tbody > tr:nth-child(", i, ") > td.b-td-left > div > a")
    title.nodes <- html_nodes(html_source, T.selector)
    T.title <- html_text(title.nodes)
    title <- c(title, T.title)

    # 담당부서 크롤링 코드
    D.selector <- paste0("#cms-content > div > div > div.bn-list-common02.type01.bn-common-cate
> table > tbody > tr:nth-child(", i, ") > td:nth-child(5)")
    dept.nodes <- html_nodes(html_source, D.selector)
    T.dept <- html_text(dept.nodes)
    dept <- c(dept, T.dept)

    #날짜 크롤링 코드
    date.selector <- paste0("#cms-content > div > div > div.bn-list-common02.type01.bn-common-
cate > table > tbody > tr:nth-child(", i, ") > td:nth-child(6)")
    date.nodes <- html_nodes(html_source, date.selector)
    T.date <- html_text(date.nodes)
    date <- c(date, T.date)

  }
}
```

```
#위에서 title, date, dept에 벡터로 저장한 데이터들로 ajou.notice라는 이름의 데이터 프레임 생성
ajou.notice <- data.frame(title, date, dept)

ajou.notice[,1] <- gsub("Wn", "", ajou.notice[,1])#줄바꿈 제거(공백으로 대체)
ajou.notice[,1] <- gsub("Wt", "", ajou.notice[,1])# 탭 제거(공백으로 대체)
ajou.notice[,1] <- gsub("Wr", "", ajou.notice[,1])#문자열에서 줄을 바꿀 때 사용되는 제어 문자
중 하나인 캐리지 리턴 제거(공백으로 대체)

head(ajou.notice)
```

```
##
title
## 1 [일자리+센터] 선택의
기회, 2023 산업별 핵심이슈 특강
## 2 (재공지) [메이커스페이스] 23.11월 기초교육 교육생 모집안내 (CNC조각기, 프랑스자수, 조향,
(추가)캔들&위빙코스터 기초교육 순)
## 3 'Colorful My Emotion!' 정서조절을 위한
인식과 표현(5차) 특강 참여자 모집
## 4 [일자리+센터] 로스쿨? 변호사? 여러분의 진
로 고민을 해결해드립니다! (재공지)
## 5 [메이커스페이스] 레이저커터 장비운용 기초
교육생 모집(11월 1회차_ '23.11.14)
## 6 [학부/학사과정] (23.11.09.) 2023-동계 계
절학기 국내대학 학점교류 신청 안내
## date dept
## 1 2023-11-10 대학일자리플러스센터
## 2 2023-11-10 창업지원팀
## 3 2023-11-10 인권센터 학생상담소
## 4 2023-11-10 대학일자리플러스센터
## 5 2023-11-09 창업지원팀
## 6 2023-11-09 교무팀
```

## Rselenium을 이용한 ‘관광’ 분야의 연구개발 성과 크롤링코드

```
# RSelenium 과 필요한 패키지들 로드
library(RSelenium)
```

```
## Warning: 패키지 'RSelenium'는 R 버전 4.3.2에서 작성되었습니다
```

```
library(httr)
library(dplyr)
# 원격 드라이버를 설정하며 크롬드라이버를 이용한다.
remDR <- remoteDriver(remoteServerAddr="localhost", port=4445L, browserName="chrome")
remDR$open() #원격드라이버를 실행한다.
```

```
## [1] "Connecting to remote server"
## $acceptInsecureCerts
## [1] FALSE
##
## $browserName
## [1] "chrome"
##
## $browserVersion
## [1] "119.0.6045.124"
##
## $chrome
## $chrome$chromedriverVersion
## [1] "119.0.6045.105 (38c72552c5e15ba9b3117c0967a0fd105072d7c6-refs/branch-heads/6045@{#1103})"
##
## $chrome$userDataDir
## [1] "C:\\Users\\WWgroup\\AppData\\Local\\Temp\\WWscoped_dir26052_798948873"
##
##
## $`fedcm:accounts`
## [1] TRUE
##
## $`goog:chromeOptions`
## $`goog:chromeOptions`$debuggerAddress
## [1] "localhost:63787"
##
##
## $networkConnectionEnabled
## [1] FALSE
##
## $pageLoadStrategy
## [1] "normal"
##
## $platformName
## [1] "windows"
##
## $proxy
## named list()
##
## $setWindowRect
## [1] TRUE
##
## $strictFileInteractability
## [1] FALSE
##
## $timeouts
## $timeouts$implicit
## [1] 0
##
## $timeouts$pageLoad
## [1] 300000
##
## $timeouts$script
## [1] 30000
##
```

```
##  
## $unhandledPromptBehavior  
## [1] "dismiss and notify"  
##  
## $`webauthn:extension:credBlob`  
## [1] TRUE  
##  
## $`webauthn:extension:largeBlob`  
## [1] TRUE  
##  
## $`webauthn:extension:minPinLength`  
## [1] TRUE  
##  
## $`webauthn:extension:prf`  
## [1] TRUE  
##  
## $`webauthn:virtualAuthenticators`  
## [1] TRUE  
##  
## $webdriver.remote.sessionid  
## [1] "529796517f1037ad1403a7d8cc8ea299"  
##  
## $id  
## [1] "529796517f1037ad1403a7d8cc8ea299"
```



```

# navigate함수를 통해 해당 웹사이트로 이동
remDR$navigate('https://www.ntis.go.kr/outcomes/popup/srchRstList.do')

find_ <- remDR$findElement(using="css", value="#searchKey") #searchKey라는 아이디를 가진 요소를
찾는 코드
find_$clickElement() #찾은 요소를 클릭한다.
find_$sendKeysToElement(list("관광")) #클릭한 부분에 "관광"이라는 값을 보낸다(입력한다.)

find_click <- remDR$findElement(using="css", value="#btnHeaderSearch") #btnHeaderSearch라는 id
를 가진 요소를 찾는다.
Sys.sleep(1) #잠시 대기하고
find_click$clickElement() #찾은 요소를 클릭한다.

Sys.sleep(1)
research_ <- remDR$findElement(using='xpath', value='/html/body/div[2]/div[3]/form/div[5]/div/div[4]/div[1]/table/tbody/tr[6]/td/ul/li[1]/input')
# 연구개발성과부분에서 "논문"을 찾고
research_$clickElement() #클릭하여 체크표시를 하고
filter_click <- remDR$findElement(using='xpath', value='/html/body/div[2]/div[3]/form/div[5]/div/button')
#적용하기 버튼을 찾고
filter_click$clickElement()#클릭을 한다.
research_ <- remDR$findElement(using = 'css', value = paste0('#content > div.layoutBox > div.list_box > div:nth-child(2) > div > a.rstTitle'))#필터를 적용해서 나온 결과중 첫번째 논문의 제목
을 찾고
research_$clickElement()#클릭한다.
###
myswitch <- function (remDr, windowId)
{
  qpath <- sprintf("%s/session/%s/window", remDr$serverURL,
                    remDr$sessionInfo[["id"]])
  remDr$queryRD(qpath, "POST", qdata = list(handle = windowId))
}
check_handle <- FALSE
count <- 0
while(!check_handle || count > 20){
  count <- count + 1
  windows_handles <- remDR$getWindowHandles()
  if(length(windows_handles) < 2){
    Sys.sleep(1)
  }else{
    check_handle <- TRUE
  }
}

myswitch(remDR, windows_handles[[2]])
#### 여기까지.

research_title <- remDR$findElement(using='css', value='#content > div.po_rel > div.outcomesheader_wrap > dl > dd > span.head')#논문의 제목을 찾고
research_title$getElementText()[[1]] #스칼라형태로 가져온다.

```

```
## [1] "농촌 치유관광 운영자의 농촌성 인식과 활용자원 분석"
```

```
research_date <- remDR$findElement(using='css', value='#content > div.po_rel > div.defaultInfo_
wrap > div.defaultInfo > table > tbody > tr:nth-child(4) > td:nth-child(4)') #논문이 작성된 날
짜를 찾고
research_date$getText()[1] # 스칼라 형태로 가져온다.
```

```
## [1] "2023-09-30"
```

```
research_abstract <- remDR$findElement(using='css', value='#content > div.po_rel > div.defaultI
nfo_wrap > div.defaultInfo > dl > dd') #논문의 초록을 찾고
research_abstract$getText()[1] #스칼라 형태로 가져온다.
```

```
## [1] "Along with rural healing tourism aimed at improving health and restoring mind and body
in nature, rural areas are attracting attention where abundant nature and rural sentiment can u
tilize their healing properties for visitors. How the operators running rural healing tourism b
usinesses perceive "rurality" and utilize rural resources are important for the sustainable d
evelopment of rural healing tourism. This study aims to analyze the types of rurality and the u
tilization resources recognized by rural healing tourism operators. The results indicated that
operators' perception of rurality as rural healing areas was categorized into excellent natura
l environment, nostalgic for hometown, warm affection of villagers, relaxation not seen in the
city, and rural experience. Rural resources that can be used for healing tourism were categoriz
ed into natural resources, historical resources, facility resources, tourism resources, communi
ty activity resources, and other resources. The findings suggest that rural areas have the pote
ntial to be used as an attractive resource with functions such as comfort, emotional care, and
mental and physical stability, as well as abundant natural environment. In addition, in order f
or a rural healing tourism village to establish itself as a competitive healing environment, ef
fective resource utilization strategies are needed so that village resources with healing eleme
nts can create unique healing values."
```

```
remDR$closeWindow()#창을 닫는다.
myswitch(remDR, windows_handles[[1]])
```

## 문제1

## 문제2

- 첫번째 반복문은 공지사항의 몇페이지부터 몇페이지까지의 내용을 가져올때 그 페이지를 지정하기 위함이다. 맨 마지막 부분 offset=뒤에 0이 오면 1페이지 10이오면 2페이지가 오는 구조이기에 이 반복문은 작동한다.
- 두번째 반복문은 한페이지에 있는 22개의 공지사항 중 몇번째부터 몇번째까지를 선택할지를 가능하게 하는 반복문이다. 해당 웹페이지는 12번째 까지는 어느페이지에서든 노출되는 [공지]이기 때문에 i를 임의로 13번째부터 22번째로 조정하였다.

## 문제 3

- selenium을 이용해야만 크롤링을 할 수 있는 사이트와 크롤링하려는 이유를 작성하여라. selenium을 사용하지 않는다면 사용자의 요구에 따라 동적으로 작동하는 페이지가 아닌 멈춰있는 페이지인 정적인 웹 페이지에서만 정보를 크롤링 할 수있다. 현대 사회에 있어서 SNS는 현대인에게 빠질 수 없는 필수적인 소통의 채널이다. 실시간으로 동적인 움직임을 보이며 댓글이 쌓여가는 sns사이트나 실시간으로 정보가 변하는 금융관련 사이트에서는 selenium을 사용해야만 동적인 사이트를 크롤링 할 수 있다. 개인적으로는 실시간 스포츠 스트리밍 사이트의 실시간 댓글을 크롤링하여 그 정보들을 분석해보고 싶다.

## 문제 4

10개의 항목을 크롤링하려고하면 반복된 동작 수행으로 인해 reCAPTCHA로 크롤링이 제대로 안돼서 크롤링 항목수를 5개로 임의로 수정하였습니다. 원래의 반복문은 for (i in 0:9) 입니다.

```
library(RSelenium)
library(httr)
library(dplyr)
remDR <- remoteDriver(remoteServerAddr="localhost", port=4445L, browserName="chrome")
remDR$open()
```

```
## [1] "Connecting to remote server"
## $acceptInsecureCerts
## [1] FALSE
##
## $browserName
## [1] "chrome"
##
## $browserVersion
## [1] "119.0.6045.124"
##
## $chrome
## $chrome$chromedriverVersion
## [1] "119.0.6045.105 (38c72552c5e15ba9b3117c0967a0fd105072d7c6-refs/branch-heads/6045@{#1103})"
##
## $chrome$userDataDir
## [1] "C:\\Users\\WWgru\\AppData\\Local\\Temp\\WWscoped_dir29580_1314088648"
##
##
## $`fedcm:accounts`
## [1] TRUE
##
## $`goog:chromeOptions`
## $`goog:chromeOptions`$debuggerAddress
## [1] "localhost:63834"
##
##
## $networkConnectionEnabled
## [1] FALSE
##
## $pageLoadStrategy
## [1] "normal"
##
## $platformName
## [1] "windows"
##
## $proxy
## named list()
##
## $setWindowRect
## [1] TRUE
##
## $strictFileInteractability
## [1] FALSE
##
## $timeouts
## $timeouts$implicit
## [1] 0
##
## $timeouts$pageLoad
## [1] 300000
##
## $timeouts$script
## [1] 30000
##
```

```
##
## $unhandledPromptBehavior
## [1] "dismiss and notify"
##
## $`webauthn:extension:credBlob`
## [1] TRUE
##
## $`webauthn:extension:largeBlob`
## [1] TRUE
##
## $`webauthn:extension:minPinLength`
## [1] TRUE
##
## $`webauthn:extension:prf`
## [1] TRUE
##
## $`webauthn:virtualAuthenticators`
## [1] TRUE
##
## $webdriver.remote.sessionid
## [1] "993fc54ae7189bd50603658c06a0877d"
##
## $id
## [1] "993fc54ae7189bd50603658c06a0877d"
```

```

remDR$navigate('https://www.ntis.go.kr/outcomes/popup/srchRstList.do')

find_ <- remDR$findElement(using="css", value="#searchKey")
find_$clickElement()
find_$sendKeysToElement(list("관광"))

find_click <- remDR$findElement(using="css", value="#btnHeaderSearch")
find_click$clickElement()

Sys.sleep(1)
research_ <- remDR$findElement(using='xpath', value='/html/body/div[2]/div[3]/form/div[5]/div/div[4]/div[1]/table/tbody/tr[6]/td/ul/li[1]/input')
research_$clickElement()
filter_click <- remDR$findElement(using='xpath', value='/html/body/div[2]/div[3]/form/div[5]/div/button')
filter_click$clickElement()

research_data <- data.frame(
  Title = character(),
  Date = character(),
  Abstract = character(),
  stringsAsFactors = FALSE
)

for (i in 0:4) {
  research_ <- remDR$findElement(using = 'css', value = paste0('#content > div.layoutBox > div.list_box > div:nth-child(', i + 2, ') > div > a.rstTitle'))
  research_$clickElement()

  myswitch <- function (remDr, windowId) {
    qpath <- sprintf("%s/session/%s/window", remDr$serverURL,
                     remDr$sessionInfo[["id"]])
    remDr$queryRD(qpath, "POST", qdata = list(handle = windowId))
  }

  check_handle <- FALSE
  count <- 0
  while (!check_handle || count > 20) {
    count <- count + 1
    windows_handles <- remDR$getWindowHandles()
    if (length(windows_handles) < 2) {
      Sys.sleep(1)
    } else {
      check_handle <- TRUE
    }
  }

  myswitch(remDR, windows_handles[[2]])
  Sys.sleep(1)
  # 여기서부터 데이터 추출
  research_title <- remDR$findElement(using='css', value='#content > div.po_rel > div.outcomesh
eader_wrap > dl > dd > span.head')$getElementText()[[1]]
  research_date <- remDR$findElement(using='css', value='#content > div.po_rel > div.defaultInf
o_wrap > div.defaultInfo > table > tbody > tr:nth-child(4) > td:nth-child(4)')$getElementText()
[[1]]

```

```
research_abstract <- remDR$findElement(using='css', value='#content > div.po_rel > div.defaultInfo_wrap > div.defaultInfo > dl > dd')$getText()[[1]]

# 데이터 프레임에 추가
new_data <- data.frame(
  Title = research_title,
  Date = research_date,
  Abstract = research_abstract,
  stringsAsFactors = FALSE
)
research_data <- rbind(research_data, new_data)

remDR$closeWindow()
myswitch(remDR, windows_handles[[1]])
}

# 창 닫기
remDR$close()

print(research_data)
```

```
##
Title
## 1 농촌 치유관광 운영자의 농촌
성 인식과 활용자원 분석
## 2 농촌성 인식, 치유관광 태도, 행동
의도 간의 구조관계 분석
## 3 객체 인식 모델을 활용한 적재
불량 화물차 탐지 시스템
## 4 A street-view-based method to detect urban growth and decline: A case study of Midtown in
Detroit, Michigan, USA
## 5 Fuzzy probabilistic evaluation of soil-structure interaction effects on the soft-
first-story structures
## Date
## 1 2023-09-30
## 2 2023-03-31
## 3 2022-12-31
## 4 2022-02-08
## 5 2022-07-01
##
```

## Abstract

## 1 Along with rural healing tourism aimed at improving health and restoring mind and body in nature, rural areas are attracting attention where abundant nature and rural sentiment can utilize their healing properties for visitors. How the operators running rural healing tourism businesses perceive “rurality” and utilize rural resources are important for the sustainable development of rural healing tourism. This study aims to analyze the types of rurality and the utilization resources recognized by rural healing tourism operators. The results indicated that operators’ perception of rurality as rural healing areas was categorized into excellent natural environment, nostalgic for hometown, warm affection of villagers, relaxation not seen in the city, and rural experience. Rural resources that can be used for healing tourism were categorized into natural resources, historical resources, facility resources, tourism resources, community activity resources, and other resources. The findings suggest that rural areas have the potential to be used as an attractive resource with functions such as comfort, emotional care, and mental and physical stability, as well as abundant natural environment. In addition, in order for a rural healing tourism village to establish itself as a competitive healing environment, effective resource utilization strategies are needed so that village resources with healing elements can create unique healing values.

## ## 2

This study aims to examine the structural relationships among rurality perception, tourism attitude, and behavioral intention in rural healing tourism. Particularly, this study focused on the mediation effect of attitude. Data were collected from 658 rural healing tourists by online survey. SPSS 28.0 and AMOS 28.0 were used for frequency, exploratory factor, confirmatory factor, and structural equation modeling analyses. The results of structural equation modeling indicate perception of rurality had a positive effect on attitude and behavioral intention. In addition, attitude mediated the relationships between perception of rurality and behavioral intention. The results of this study suggest the importance of understanding the interests and pursuits of rural healing tourists and providing them with the values they want. Since rurality has a direct influence on attracting healing tourists, it is necessary to establish a positive image of rural healing tourism.

## ## 3

최근 증가하고 있는 도로 위 적재 불량 화물차는 비정상적인 무게 중심으로 인해 물체 낙하, 도로 파손, 연쇄 추돌 등 교통안전에 위해가 되고 한번 사고가 발생하면 큰 피해가 유발할 수 있다. 하지만 이러한 비정상적인 무게 중심은 적재 불량 차량 인식을 위한 주행 중 축중 시스템으로는 검출이 불가능하다는 한계점이 있다. 본 논문에서는 이러한 사회 문제를 야기하는 적재 불량 차량을 관리하기 위한 객체 인식 기반 AI 모델을 구축하고자 한다. 또한 AI-Hub에 공개된 약 40만 장의 데이터셋을 비교 분석하여 전처리를 통해 적재 불량 차량 검지 AI 모델의 성능을 향상시키는 방법을 제시한다. 또한 객



체 추적을 통해 실시간 검지를 수행하는 방법을 제안한다. 이를 통해, 원시 데이터를 활용한 학습 성능 대비 약 23% 향상된 적재 불량 차량의 검출 성능을 나타냄을 보였다. 본 연구 결과를 통해 공개 빅 데이터를 보다 효율적으로 활용하여, 객체 인식 기반 적재 불량 차량 탐지 모델 개발에 적용할 수 있을 것으로 기대된다.

## 4

Urban growth and decline occur every year and show changes in urban areas. Although various approaches to detect urban changes have been developed, they mainly use large scale satellite imagery and socioeconomic factors in urban areas, which provides an overview of urban changes. However, since people explore places and notice changes daily at the street level, it would be useful to develop a method to identify urban changes at the street level and demonstrate whether urban growth or decline occurs there. Thus, this study seeks to use street-level panoramic images from Google Street View to identify urban changes and to develop a new way to evaluate the growth and decline of an urban area. After collecting Google Street View images year by year, we trained and developed a deep-learning model of an object detection process using the open-source software TensorFlow. By scoring objects and changes detected on a street from year to year, a map of urban growth and decline was generated for Midtown in Detroit, Michigan, USA. By comparing socioeconomic changes and the situations of objects and changes in Midtown, the proposed method is shown to be helpful for analyzing urban growth and decline by using year-by-year streetview images.

## 5

Even though the soil-structure interaction (SSI) has been investigated for long, less attention has been paid to the probabilistic framework. The present study used a fuzzy probabilistic approach to evaluate the effect of SSI on the nonlinear behavior of a soft-first-story structure with uncertain soil types. Sensitivity analysis was carried out to identify the relative importance of uncertainty parameters on the structural responses of the model structure. The analysis results showed that consideration of SSI in structures with relatively weak soil conditions could enhance the probability of irreparable damages in the structures. It was also observed that the variations in the weight density and shear wave velocity of soil were the most influential parameters in the structural responses.

## 도전문제

다음 페이지로 가는 코드

```
next_page <- remDR$findElement(using = "css", value = 'a[onclick*="srchList(2)"]')
next_page$clickElement() 추가
```

```
library(RSelenium)
library(httr)
library(dplyr)
remDR <- remoteDriver(remoteServerAddr="localhost", port=4445L, browserName="chrome")
remDR$open()
```

```
## [1] "Connecting to remote server"
## $acceptInsecureCerts
## [1] FALSE
##
## $browserName
## [1] "chrome"
##
## $browserVersion
## [1] "119.0.6045.124"
##
## $chrome
## $chrome$chromedriverVersion
## [1] "119.0.6045.105 (38c72552c5e15ba9b3117c0967a0fd105072d7c6-refs/branch-heads/6045@{#1103})"
##
## $chrome$userDataDir
## [1] "C:\\Users\\WWgru\\AppData\\Local\\Temp\\WWscoped_dir37760_1675364027"
##
## $`fedcm:accounts`
## [1] TRUE
##
## $`goog:chromeOptions`
## $`goog:chromeOptions`$debuggerAddress
## [1] "localhost:63902"
##
##
## $networkConnectionEnabled
## [1] FALSE
##
## $pageLoadStrategy
## [1] "normal"
##
## $platformName
## [1] "windows"
##
## $proxy
## named list()
##
## $setWindowRect
## [1] TRUE
##
## $strictFileInteractability
## [1] FALSE
##
## $timeouts
## $timeouts$implicit
## [1] 0
##
## $timeouts$pageLoad
## [1] 300000
##
## $timeouts$script
## [1] 30000
##
```

```
##  
## $unhandledPromptBehavior  
## [1] "dismiss and notify"  
##  
## $`webauthn:extension:credBlob`  
## [1] TRUE  
##  
## $`webauthn:extension:largeBlob`  
## [1] TRUE  
##  
## $`webauthn:extension:minPinLength`  
## [1] TRUE  
##  
## $`webauthn:extension:prf`  
## [1] TRUE  
##  
## $`webauthn:virtualAuthenticators`  
## [1] TRUE  
##  
## $webdriver.remote.sessionid  
## [1] "faa41e2303cb11fc875607d5854f09c4"  
##  
## $id  
## [1] "faa41e2303cb11fc875607d5854f09c4"
```

```

remDR$navigate('https://www.ntis.go.kr/outcomes/popup/srchRstList.do')

find_ <- remDR$findElement(using="css", value="#searchKey")
find_$clickElement()
find_$sendKeysToElement(list("관광"))

find_click <- remDR$findElement(using="css", value="#btnHeaderSearch")
find_click$clickElement()

Sys.sleep(1)
research_ <- remDR$findElement(using='xpath', value='/html/body/div[2]/div[3]/form/div[5]/div/div[4]/div[1]/table/tbody/tr[6]/td/ul/li[1]/input')
research_$clickElement()
filter_click <- remDR$findElement(using='xpath', value='/html/body/div[2]/div[3]/form/div[5]/div/button')
filter_click$clickElement()

research_data2 <- data.frame(
  Title = character(),
  Date = character(),
  Abstract = character(),
  stringsAsFactors = FALSE
)

next_page <- remDR$findElement(using = "css", value = 'a[onclick*="srchList(2)"]')

next_page$clickElement()

for (i in 0:4) {
  research_ <- remDR$findElement(using = 'css', value = paste0('#content > div.layoutBox > div.list_box > div:nth-child(', i + 2, ') > div > a.rstTitle'))
  research_$clickElement()

  myswitch <- function (remDr, windowId) {
    qpath <- sprintf("%s/session/%s/window", remDr$serverURL,
                     remDr$sessionInfo[["id"]])
    remDr$queryRD(qpath, "POST", qdata = list(handle = windowId))
  }

  check_handle <- FALSE
  count <- 0
  while (!check_handle || count > 20) {
    count <- count + 1
    windows_handles <- remDR$getWindowHandles()
    if (length(windows_handles) < 2) {
      Sys.sleep(1)
    } else {
      check_handle <- TRUE
    }
  }

  myswitch(remDR, windows_handles[[2]])
  Sys.sleep(1)
  # 여기서부터 데이터 추출
  research_title <- remDR$findElement(using='css', value='#content > div.po_rel > div.outcomesh

```

```
header_wrap > dl > dd > span.head')$getElementText()[[1]]
  research_date <- remDR$findElement(using='css', value='#content > div.po_rel > div.defaultInfo_wrap > div.defaultInfo > table > tbody > tr:nth-child(4) > td:nth-child(4)')$getElementText()[[1]]
  research_abstract <- remDR$findElement(using='css', value='#content > div.po_rel > div.defaultInfo_wrap > div.defaultInfo > dl > dd')$getElementText()[[1]]

# 데이터 프레임에 추가
new_data <- data.frame(
  Title = research_title,
  Date = research_date,
  Abstract = research_abstract,
  stringsAsFactors = FALSE
)
research_data2 <- rbind(research_data2, new_data)

remDR$closeWindow()
myswitch(remDR, windows_handles[[1]])
}

# 창 닫기
remDR$close()
print(research_data2)
```

```
##
Title
## 1          Compatible coupling of discrete elements and finite elements using Delaunay-Voronoi dual tessellations
## 2          Runoff Reduction Effects at Installation of LID Facilities under Different Climate Change Scenarios
## 3          A simplified ductility-based design procedure for seismic retrofit of structures using hysteretic devices
## 4          A Study on Mega-Shelter Layout Planning Based on User Behavior
## 5 Performance Evaluation of Feature Matching Techniques for Detecting Reinforced Soil Retaining Wall Displacement
##          Date
## 1 2022-03-28
## 2 2022-04-16
## 3 2022-10-01
## 4 2022-10-08
## 5 2022-03-31
##
```

## Abstract

## 1

The Voronoi cell lattice model (VCLM) is a discrete approach for simulating the behavior of solids and structures, based on a Voronoi cell partitioning of the domain. In this study, the duality between Voronoi and Delaunay tessellations is used to couple distinct regions represented by VCLM and the finite element method (FEM). By introducing an edge-based smoothing scheme in the FEM, the element frame is transformed from the conventional triangular body to the edge entity. Therefore, along each of the Delaunay edges, both the lattice and finite elements can be defined, which provides several advantages: (a) The regions modeled by each respective approach are clearly distinguished without the need for interface elements, (b) algorithmic efficiency is enhanced during element-wise computations during explicit time integration, and (c) the element performance of the three-node triangular element is improved by introducing the edge-based strain smoothing technique. Selected examples are used to validate the VCLM-FEM coupling approach. Simulations of elastic behavior, geometric nonlinearity, and fracture are conducted. The simulation results agree well with the corresponding theoretical, numerical, and experimental results, which demonstrates the capabilities of the proposed compatible coupling scheme.

## 2

: In order to prepare for floods and droughts that occur as a result of climate change, various studies in water-related fields are being carried out in various countries around the world. Among them, special attention is being paid to the low-impact development (LID) technique. This study measured the annual maximum daily rainfall data from 2019 to 2100 by climate change scenario, which is the annual maximum daily rainfall series observed for rainfall stations, and tested the hydrological data using statistical analysis. After determining whether the data could be analyzed, the probability distribution was selected, and the parameters of the selected probability distribution were calculated using the L-moment method for each rainfall station. The probabilities of rainfall data were derived using GEV distribution, and the United States Environmental Protection Agency Storm Water Management Model (SWMM), a runoff simulation program, was used to compare and analyze the runoff reduction rate before and after the installation of a permeable pavement as an LID facility. The results of the analysis showed that representative concentration pathway (RCP) 4.5 and RCP 8.5 had the effect of reducing the runoff for more than 100 years at a 30% reduction rate compared with before installation.

## 3

The main aim of this study is to derive a simple ductility-based seismic retrofit procedure for buildings using hysteretic devices. In the procedure it is tried to minimize the assumptions while taking more details into consideration compared to previous approaches. With an available seismic retrofit device, the proposed procedure distributes the devices along the height of the structure

e to satisfy a given performance limit state through a numerical iterative method. The pushover curve and the modal characteristics of the structure are the only required input parameters, and the estimation of the pushover curve after retrofit, its modal characteristics, and maximum interstory drift ratios are byproducts of this procedure. The derived method was applied to seismic retrofit of a case study structure using buckling-restrained braces. The structure was designed to meet the enhanced performance objective of the 1% maximum drift ratio under the Maximum Considered Earthquake level. The nonlinear time history analysis of the structure subjected to 22 pairs of earthquake records showed that the median of the drift ratio demand is 0.8%, satisfying the given limit state. The results show that the proposed procedure can be implemented efficiently in estimating the required amount of seismic retrofit devices and the structural behavior after retrofit.

## 4 We explore the spatial layouts of mega-shelters and suggest better spatial planning strategies. A mega-shelter for refugees contains multiple functions, such as dormitory, dining, medical, kitchen, storage, and community areas. Post-disaster refugees often suffer from PTSD that affects their mental health and spatial cognitive ability. The spatial configuration of a mega-shelter can accelerate their recovery by providing an environment that not only satisfies the basic needs, but one that can improve their spatial cognitive ability and promote a sense of community in this new, albeit temporary, small society. Four mega-shelters in the U.S., Australia, and Japan were analyzed using space syntax methods, specifically axial line analysis and visibility graph analysis (VGA), as well as justified graph analysis. The comparative analysis shows that while specific spatial layouts are different, all shelters were designed from a manager's perspective. The movements of the refugees were sometimes unnecessarily exposed to supervision and control, and community areas were often found in locations with low accessibility. By incorporating strategies such as siting community space in areas with high global integration values and adopting transition areas, mega-shelters can create an environment that can enhance the refugees' will to recover and rebuild by promoting communications with neighbors and various community activities.

## 5

Image registration technology is widely applied in various matching methods. In this study, we aim to evaluate the feature matching performance and to find an optimal technique for detecting three types of behaviors: facing displacement, settlement, and combined displacement in reinforced soil retaining walls (RSWs). For a single block with an artificial target and a multiblock structure with artificial and natural targets, five popular detectors and descriptors: KAZE, SURF, MinEigen, ORB, and BRISK were used to evaluate the resolution performance. For comparison, the repeatability, matching score, and inlier matching features were analyzed based on the number of extracted and matched features. The axial registration error (ARE) was used to verify the accuracy of the methods by comparing the position between the estimated and real features. The results showed that the KAZE method was the best detector and descriptor for RSWs (block shape target), with the highest probability of successfully matching features. In the multiblock experiment, the block used as a natural target showed similar matching performance to that of the block with an artificial target attached. Therefore, the behaviors of RSW blocks can be analyzed using the KAZE method without installing an artificial target.