# Management of duplicates in ris files

Marius Bottin

2023-04-09

# Contents

```r
knitr::opts_chunk$set(tidy.opts = list(width.cutoff = 70), tidy = TRUE, cache=F)
def.chunk.hook  <- knitr::knit_hooks$get("chunk")
knitr::knit_hooks$set(chunk = function(x, options) {
  x <- def.chunk.hook(x, options)
  paste0("\n \\", "footnotesize","\n\n", x, "\n\n \\normalsize\n\n")
})
```

# 1 Extracting and analysing ris Files

## 1.1 Creating the reading function

```r
# This function allows to read a ris files with the following
# obligatory characteristics: - no empty lines except the ones
# between registers - always at least an empty line between 2
# registers - all registers have one and only one title field, except
# for abstract which might be on more than one line - each line
# beginning is the name of a field

read_ris <- function(risFile, keepCompleteRaw = T, extractFields = T) {
    fileLines <- readLines(risFile)
    emptyLines <- grep("^[::blank::]*$", fileLines)
    sep_registers <- numeric()
    for (i in 1:length(emptyLines)) {
        if (i != length(emptyLines) && emptyLines[i] != (emptyLines[i +
            1] - 1))
            sep_registers <- c(sep_registers, emptyLines[i])
    }
    titleLines <- grep("^TI +-.*", fileLines)
    abstractLines <- grep("^AB +-.*", fileLines)
    stopifnot(length(sep_registers) + 1 == length(titleLines))
    registers <- data.frame(id = 1:(length(sep_registers) + 1), begin = c(1,
        sep_registers + 1), end = c(sep_registers, length(fileLines)))
```

```r
    idLines <- rep(registers$id, (registers$end - registers$begin) + 1)
    lineFields <- grepl("^?([A-Z][A-Za-z0-9]) +-.*$", fileLines)
    lineOtherFields <- grepl("^Total Times Cited: +[0-9]+$", fileLines) |
        grepl("^Cited Reference Count: +[0-9]+$", fileLines)
    emptyLinesBool <- grepl("^[::blank::]*$", fileLines)
    followingMuliLineAB <- NULL
    for (i in 1:length(abstractLines)) {
        ct <- (abstractLines[i] + 1)
        while (!(lineFields[ct] | lineOtherFields[ct] | emptyLinesBool[ct])) {
            followingMuliLineAB <- c(followingMuliLineAB, ct)
            ct <- ct + 1
        }
    }
    fileLines[!(lineFields | lineOtherFields | emptyLinesBool)]
    fieldName <- gsub("^?([A-Z][A-Za-z0-9]) +-.*$", "\\1", fileLines)
    fieldName[followingMuliLineAB] <- "AB_f"
    fieldName[lineOtherFields] <- gsub("^([ A-Za-z]+): +([0-9]+)$", "\\1",
        fileLines[lineOtherFields])
    fieldName[emptyLinesBool] <- NA
    if (extractFields) {
        fieldContent <- gsub("^?[A-Z][A-Za-z0-9] +- +(.*)$", "\\1",
            fileLines)
        fieldContent[grepl("^?[A-Z][A-Za-z0-9] +-$", fileLines)] <- ""
        fieldContent[followingMuliLineAB] <- gsub("^ *([A-Za-z0-9].*)$",
            "\\1", fileLines[followingMuliLineAB])
        fieldContent[lineOtherFields] <- gsub("^([ A-Za-z]+): +([0-9]+)$",
            "\\2", fileLines[lineOtherFields])
        fieldContent[emptyLinesBool] <- NA
    }
    res <- list(nbRecords = max(registers$id), registers = registers, fieldName = fieldName,
        lineRegId = idLines)
    if (keepCompleteRaw) {
        res$raw <- fileLines
    }
    if (extractFields) {
        res$content <- fieldContent
    }
    return(res)
}
```

## 1.2 Reading the files

```r
risFiles <- c("../../Search/Results/scopus20230403/scopus.ris", "../../Search/Results/wosCore20230403/wosCore.ris",
    "../../Search/Results/wosSciELO20230403/savedrecs.ris")
names(risFiles) <- c("scopus", "WoS_core", "WoS_SciELO")
risExtracted <- lapply(risFiles, read_ris)
```

# 2 Extracting fields

```r
extractFields <- function(extractedRis, fields = c("DO", "TI", "AU"), sep = "|") {
    # Are fields unique by record
    tab <- data.frame(fieldName = extractedRis$fieldName[!is.na(extractedRis$fieldName) &
        extractedRis$fieldName %in% fields], record = extractedRis$lineRegId[!is.na(extractedRis$fieldName) &
        extractedRis$fieldName %in% fields], content = extractedRis$content[!is.na(extractedRis$fieldName) &
        extractedRis$fieldName %in% fields])
    f_multi <- sapply(by(tab, INDICES = tab$fieldName, FUN = duplicated),
        any)
    #
    reg <- extractedRis$registers$id
    no_multi_field <- names(f_multi)[!f_multi]
    tab2 <- tab[tab$fieldName %in% no_multi_field, ]
    if (any(f_multi)) {
        multi_field <- names(f_multi)[f_multi]
        names(multi_field) <- multi_field
```

```
        dataMF <- lapply(multi_field, function(f, t, s) tapply(t$content[t$fieldName ==
            f], t$record[t$fieldName == f], paste, collapse = s), t = tab,
            s = sep)
        dataMF_tab <- data.frame(fieldName = rep(names(dataMF), sapply(dataMF,
            length)), record = as.numeric(unlist(lapply(dataMF, names))),
            content = unlist(dataMF))
        tab2 <- rbind(tab2, dataMF_tab)
    }
    COLS <- unique(as.character(tab2$fieldName))
    ROWS <- unique(as.character(extractedRis$registers$id))
    arr.which <- cbind(row = match(as.character(tab2$record), ROWS), col = match(tab2$fieldName,
        COLS))
    res <- matrix(nrow = length(ROWS), ncol = length(COLS), dimnames = list(as.character(ROWS),
        COLS))
    res[arr.which] <- tab2$content
    return(as.data.frame(cbind(id = ROWS, res)))
}
```

## 2.1 Application

```
require(kableExtra)
```

```
## Loading required package: kableExtra
```

```
extractFields(risExtracted$scopus, c("DO", "TI")) %>%
    head() %>%
    kable(booktabs = T) %>%
    column_spec(2, width = "8cm")
```

| id | TI | DO |
|----|----|----|
| 1 | What influences college students to continue using business simulation games? The Taiwan experience | 10.1016/j.compedu.2009.05.009 |
| 2 | Predicting academic performance of students from VLE big data using deep learning models | 10.1016/j.chb.2019.106189 |
| 3 | Classroom social climate, self-determined motivation, willingness to communicate, and achievement: A study of structural relationships in instructed second language settings | 10.1016/j.lindif.2016.11.005 |
| 4 | Internet-based prevention for alcohol and cannabis use: Final results of the Climate Schools course | 10.1111/j.1360-0443.2009.02853.x |
| 5 | Direct protective and buffering protective factors in the development of youth violence | 10.1016/j.amepre.2012.04.029 |
| 6 | Living labs and co-production: University campuses as platforms for sustainability science | 10.1016/j.cosust.2015.06.005 |

# 3 Finding duplicates

In order to find duplicate when comparing ris files, we will use the following steps:

1. documents having the same doi will be tested for resemblance in their titles considered duplicates, if they are not chapters of a book, because those may be referenced with the DOI of the book...
2. documents not having a DOI will be compared (between them and with the ones having a doi) in terms of resemblance in their titles and years. If they have a very similar title and a maximum of one year difference, they will be considered as duplicates.

In order to compare titles: 1. we suppress all the characters which are not classical letters (space, punctuation, etc.) 1. we put every character in lower case 1. we compare strings

At first we used the following distance-based string comparation: String Alignment distance (restricted Damerau-Levenshtein distance), divide it by the length of the smallest string. Values inferior to 0.05 are considered identical, i.e. if less than 5% of the characters of the smallest string must be added or modified to obtain the larger one, they are considered identical. However, some obviously different documents were

considered duplicates, so we now just consider exact matches in modified titles. The result is that some of the duplicates are not found, but no false dupe is deleted.

Also note that if 2 documents have the same title and year but have different type, the articles and book chapters will have the priority over the conferences (for the duplicate which will be kept in the ris file).

## 3.1 function definitions

### 3.1.1 finding duplicates

```r
stopifnot(require(stringdist))
```

```
## Loading required package: stringdist
```

```r
risInternalDuplicate <- function(extractedRis, priority = c("JOUR", "CHAP")) {
    AcceptedDupli <- DupliToCheck <- data.frame()
    tabFields <- extractFields(extractedRis = extractedRis, c("DO", "TI",
        "PY", "AU", "TY"))
    tabFields$title_simp <- tolower(gsub("[^A-Za-z0-9]", "", tabFields$TI,
        perl = T))
    # Searching duplicated DOI
    tabFields_do <- tabFields[!is.na(tabFields$DO) & tabFields$TY != "CHAP",
        ]
    do_d <- duplicated(tabFields_do$DO)
    doDupli <- unique(tabFields_do[do_d, "DO"])
    toTest <- lapply(doDupli, function(x, tab) tab[tab$DO == x, ], tab = tabFields_do)
    tested <- sapply(toTest, function(x) {
        length(unique(x$title_simp)) == 1
    })
    if (any(!tested)) {
        sapply(toTest[!tested], function(x) warning("\nRecords: ", paste(x$id,
            collapse = ","), "\nhave the same DOI (", unique(x$DO), ") but differences in their titles:\n",
            paste(x$TI, collapse = "\n"), "\n\nthey will NOT be considered as duplicates\n"))
    }
    AcceptedDupli <- rbind(AcceptedDupli, Reduce(rbind, lapply(toTest[tested],
        function(x, p) {
            tab <- x[order(match(x$TY, p)), ]
            data.frame(step = "doi", ref = tab$id[1], toSupp = tab$id[2:nrow(x)])
        }, p = priority)))
    DupliToCheck <- rbind(DupliToCheck, Reduce(rbind, lapply(toTest[!tested],
        function(x) data.frame(step = "doi", ref = x$id[1], toSupp = x$id[2:nrow(x)]))))
    # Searching duplicated titles and years (max difference:1) (both
    # either do not have DOI or are chapters)
    tabFields_tiye <- tabFields[is.na(tabFields$DO) | tabFields$TY == "CHAP",
        ]
    title_simp_dup <- unique(tabFields_tiye$title_simp[duplicated(tabFields_tiye$title_simp)])
    toTest <- by(tabFields_tiye[which(tabFields_tiye$title_simp %in% title_simp_dup),
        ], tabFields_tiye$title_simp[tabFields_tiye$title_simp %in% title_simp_dup],
        function(x) x)
    diff_Ymax1 <- sapply(toTest, function(x) diff(range(as.numeric(x$PY))) <=
        1)
    sameType <- sapply(toTest, function(x) length(unique(x$TY)) == 1)
    sameAuth <- sapply(toTest, function(x) length(unique(x$AU)) == 1)
    if (any(diff_Ymax1 & (!sameAuth | !sameType))) {
        {
            sapply(toTest[diff_Ymax1 & (!sameAuth | !sameType)], function(x) warning("\nRecords: ",
                paste(x$id, collapse = ","), "\nhave comparable titles and pulication years \n(",
                paste(unique(x$TI), collapse = "\n"), ")\n but have differences in their authors and/or types:\n\nNote that they
        }
    }
    if (any(!diff_Ymax1)) {
        {
            sapply(toTest[!diff_Ymax1], function(x) warning("\nRecords: ",
                paste(x$id, collapse = ","), "\nhave comparable titles (",
                paste(unique(x$TI), collapse = "\n"), ") but have a difference of more than one publication year!\n\nNote that t
        }
    }
```

4

```r
AcceptedDupli <- rbind(AcceptedDupli, Reduce(rbind, lapply(toTest[diff_Ymax1],
    function(x, p, a) {
        tab <- x[order(match(x$TY, p), !x$id %in% a), ]
        return(data.frame(step = "ti_ye_noDoi", ref = tab$id[1], toSupp = tab$id[2:nrow(x)]))
    }, p = priority, a = AcceptedDupli$ref)))
DupliToCheck <- rbind(DupliToCheck, Reduce(rbind, lapply(toTest[!diff_Ymax1],
    function(x, p) {
        tab <- x[order(match(x$TY, p)), ]
        return(data.frame(step = "ti_ye_noDoi", ref = tab$id[1], toSupp = tab$id[2:nrow(x)]))
    }, p = priority)))
# Searching duplicated titles and years (max difference:1) (one
# of them do not have DOI or is a chapter)
m_tiye_do <- match(tabFields_tiye$title_simp, tabFields_do$title_simp)
if (sum(!is.na(m_tiye_do))) {
    title_simp_dup <- tabFields_tiye$title_simp[!is.na(m_tiye_do)]
    toTest <- by(tabFields[tabFields$title_simp %in% title_simp_dup,
        ], tabFields$title_simp[tabFields$title_simp %in% title_simp_dup],
        function(x) x)
    diff_Ymax1 <- sapply(toTest, function(x) diff(range(as.numeric(x$PY))) <=
        1)
    sameType <- sapply(toTest, function(x) length(unique(x$TY)) ==
        1)
    sameAuth <- sapply(toTest, function(x) length(unique(x$AU)) ==
        1)
    if (any(diff_Ymax1 & (!sameAuth | !sameType))) {
        {
            sapply(toTest[diff_Ymax1 & (!sameAuth | !sameType)], function(x) warning("\nRecords: ",
                paste(x$id, collapse = ","), "\nhave comparable titles and pulication years \n(",
                paste(unique(x$TI), collapse = "\n"), ")\n but have differences in their authors and/or types:\n\nNote that th
        }
    }
    if (any(!diff_Ymax1)) {
        {
            sapply(toTest[!diff_Ymax1], function(x) warning("\nRecords: ",
                paste(x$id, collapse = ","), "\nhave comparable titles (",
                paste(unique(x$TI), collapse = "\n"), ") but have a difference of more than one publication year!\n\nNote that
        }
    }
    AcceptedDupli <- rbind(AcceptedDupli, Reduce(rbind, lapply(toTest[diff_Ymax1],
        function(x, p) {
            tab <- x[order(match(x$TY, p)), ]
            return(data.frame(step = "ti_ye_1Doi", ref = tab$id[1],
                toSupp = tab$id[2:nrow(x)]))
        }, p = priority)))
    DupliToCheck <- rbind(DupliToCheck, Reduce(rbind, lapply(toTest[!diff_Ymax1],
        function(x, p) {
            tab <- x[order(match(x$TY, p)), ]
            return(data.frame(step = "ti_ye_1Doi", ref = tab$id[1],
                toSupp = tab$id[2:nrow(x)]))
        }, p = priority)))
}
# Checking that we dont suppress a document which is the
# reference a couple of duplicates If documents are cited more
# than once in AcceptedDupli, we need to make groups
nbSeen <- table(c(unique(AcceptedDupli$ref), AcceptedDupli$toSupp))
if (any(nbSeen > 1)) {
    gps = data.frame(gp = 1, id = c(AcceptedDupli[1, "ref"], AcceptedDupli[1,
        "toSupp"]))
    for (i in 2:nrow(AcceptedDupli)) {
        if (!AcceptedDupli$ref[i] %in% gps$id & !AcceptedDupli$toSupp[i] %in%
            gps$id)
            gps <- rbind(gps, data.frame(gp = max(gps$gp) + 1, id = c(AcceptedDupli[i,
                "ref"], AcceptedDupli[i, "toSupp"])))
        if (AcceptedDupli$ref[i] %in% gps$id & AcceptedDupli$toSupp[i] %in%
            gps$id) {
            gpAlready <- gps$gp[gps$id == AcceptedDupli$ref[i] | gps$id ==
                AcceptedDupli$toSupp[i]]
```

```r
                    gp <- min(gpAlready)
                    gps$gp[gps$gp %in% gpAlready] <- gp
                }
                if (AcceptedDupli$ref[i] %in% gps$id & !AcceptedDupli$toSupp[i] %in%
                    gps$id) {
                    gp <- gps$gp[gps$id == AcceptedDupli$ref[i]]
                    gps <- rbind(gps, data.frame(gp = gp, id = c(AcceptedDupli[i,
                        "toSupp"])))
                }
                if (!AcceptedDupli$ref[i] %in% gps$id & AcceptedDupli$toSupp[i] %in%
                    gps$id) {
                    gp <- gps$gp[gps$id == AcceptedDupli$toSupp[i]]
                    gps <- rbind(gps, data.frame(gp = gp, id = c(AcceptedDupli[i,
                        "ref"])))
                }

            }
        gps <- gps[order(gps$gp), ]
        gpList <- tapply(gps$id, gps$gp, function(x, t) t[x, ], t = tabFields)
        nbRefGp <- sapply(gpList, nrow)
        AD_gp <- gps$gp[match(AcceptedDupli$ref, gps$id)]
        AcceptedDupli <- AcceptedDupli[nbRefGp[as.character(AD_gp)] ==
            2, ]
        AcceptedDupli <- rbind(AcceptedDupli, Reduce(rbind, lapply(gpList[nbRefGp >
            2], function(x, p) {
            tab <- x[order(match(x, p)), ]
            return(data.frame(step = "reorganization", ref = tab$id[1],
                toSupp = tab$id[2:nrow(x)]))
        }, p = priority)))
    }
    return(list(tabInfo = tabFields, accepted_dupes = AcceptedDupli, to_check_dupes = DupliToCheck,
        toSupp = AcceptedDupli$toSupp))
}


compareRisDuplicate <- function(risToFilter, risReference) {
    # DOI
    AcceptedDupli <- DupliToCheck <- data.frame()
    rf_fieldTab <- extractFields(extractedRis = risToFilter, c("DO", "TI",
        "PY", "AU", "TY"))
    rr_fieldTab <- extractFields(extractedRis = risReference, c("DO", "TI",
        "PY", "AU", "TY"))
    rf_fieldTab$title_simp <- tolower(gsub("[^A-Za-z0-9]", "", rf_fieldTab$TI,
        perl = T))
    rr_fieldTab$title_simp <- tolower(gsub("[^A-Za-z0-9]", "", rr_fieldTab$TI,
        perl = T))
    # Searching duplicated DOI
    rr_fieldTab_do <- rr_fieldTab[!is.na(rr_fieldTab$DO) & rr_fieldTab$TY !=
        "CHAP", ]
    rf_fieldTab_do <- rf_fieldTab[!is.na(rf_fieldTab$DO) & rf_fieldTab$TY !=
        "CHAP", ]
    do_dupes <- unique(rf_fieldTab_do$DO[rf_fieldTab_do$DO %in% rr_fieldTab_do$DO])
    toTest <- lapply(do_dupes, function(x, a, b) rbind(data.frame(ref = F,
        a[a$DO == x, ]), data.frame(ref = T, b[b$DO == x, ])), a = rf_fieldTab_do,
        b = rr_fieldTab_do)
    tab <- Reduce(rbind, lapply(toTest, function(x) {
        m <- match(x[!x$ref, "title_simp"], x[x$ref, "title_simp"])
        data.frame(toSupp = x[!x$ref, "id"], ref = x[x$ref, "id"][m])
    }))
    if (nrow(tab)) {
        AcceptedDupli <- rbind(AcceptedDupli, data.frame(step = "doi",
            tab[!is.na(tab$ref), c("ref", "toSupp")]))
    }
    pbs <- Reduce(rbind, lapply(toTest, function(x) {
        pb <- x[!x$ref, "id"][!x[!x$ref, "title_simp"] %in% x[x$ref, "title_simp"]]
        if (length(pb) > 0) {
            data.frame(toSupp = pb, ref = x[x$ref, "id"])
        }
```

```r
    }))
    if (nrow(pbs)) {
        DupliToCheck <- rbind(AcceptedDupli, data.frame(step = "doi", tab[,
            c("ref", "toSupp")]))
    }
    # No Doi
    rr_fieldTab_tiye <- rr_fieldTab[is.na(rr_fieldTab$DO) | rr_fieldTab$TY ==
        "CHAP", ]
    rf_fieldTab_tiye <- rf_fieldTab[is.na(rf_fieldTab$DO) | rf_fieldTab$TY ==
        "CHAP", ]
    tisimp_dupes <- unique(rf_fieldTab_tiye$title_simp[rf_fieldTab_tiye$title_simp %in%
        rr_fieldTab_tiye$title_simp])
    toTest <- lapply(tisimp_dupes, function(x, a, b) rbind(data.frame(ref = F,
        a[a$title_simp == x, ]), data.frame(ref = T, b[b$title_simp ==
        x, ])), a = rf_fieldTab_tiye, b = rr_fieldTab_tiye)
    ok <- sapply(toTest, function(x) (diff(range(as.integer(x$PY))) <=
        1))
    if (any(ok)) {
        AcceptedDupli <- rbind(AcceptedDupli, Reduce(rbind, lapply(toTest[ok],
            function(x) data.frame(step = "no_doi", ref = x$id[x$ref],
                toSupp = x$id[!x$ref]))))
    }
    if (any(!ok)) {
        DupliToCheck <- rbind(DupliToCheck, Reduce(rbind, lapply(toTest[!ok],
            function(x) data.frame(step = "no_doi", ref = x$id[x$ref],
                toSupp = x$id[!x$ref]))))
    }
    # ref no doi
    tisimp_dupes <- unique(rf_fieldTab_do$title_simp[rf_fieldTab_do$title_simp %in%
        rr_fieldTab_tiye$title_simp])
    toTest <- lapply(tisimp_dupes, function(x, a, b) rbind(data.frame(ref = F,
        a[a$title_simp == x, ]), data.frame(ref = T, b[b$title_simp ==
        x, ])), a = rf_fieldTab_do, b = rr_fieldTab_tiye)
    ok <- sapply(toTest, function(x) (diff(range(as.integer(x$PY))) <=
        1))
    if (any(ok)) {
        AcceptedDupli <- rbind(AcceptedDupli, Reduce(rbind, lapply(toTest[ok],
            function(x) data.frame(step = "ref_no_doi", ref = x$id[x$ref],
                toSupp = x$id[!x$ref]))))
    }
    if (any(!ok)) {
        DupliToCheck <- rbind(DupliToCheck, Reduce(rbind, lapply(toTest[!ok],
            function(x) data.frame(step = "ref_no_doi", ref = x$id[x$ref],
                toSupp = x$id[!x$ref]))))
    }
    # fil no doi
    tisimp_dupes <- unique(rf_fieldTab_tiye$title_simp[rf_fieldTab_tiye$title_simp %in%
        rr_fieldTab_do$title_simp])
    toTest <- lapply(tisimp_dupes, function(x, a, b) rbind(data.frame(ref = F,
        a[a$title_simp == x, ]), data.frame(ref = T, b[b$title_simp ==
        x, ])), a = rf_fieldTab_tiye, b = rr_fieldTab_do)
    ok <- sapply(toTest, function(x) (diff(range(as.integer(x$PY))) <=
        1))
    if (any(ok)) {
        AcceptedDupli <- rbind(AcceptedDupli, Reduce(rbind, lapply(toTest[ok],
            function(x) data.frame(step = "fil_no_tiye", ref = x$id[x$ref],
                toSupp = x$id[!x$ref]))))
    }
    if (any(!ok)) {
        DupliToCheck <- rbind(DupliToCheck, Reduce(rbind, lapply(toTest[!ok],
            function(x) data.frame(step = "fil_no_tiye", ref = x$id[x$ref],
                toSupp = x$id[!x$ref]))))
    }
    message(paste(length(unique(AcceptedDupli$toSupp)), "/", risToFilter$nbRecords,
        "records are already in the reference file"))
    return(list(tabInfoRef = rr_fieldTab, tabInfoFil = rf_fieldTab, accepted_dupes = AcceptedDupli,
        to_check_dupes = DupliToCheck, toSupp = unique(AcceptedDupli$toSupp)))
```

```
}
```

### 3.1.2 Modifying and writing ris objects

```
writeRis <- function(extractedRis, filename = tempfile()) {
    writeLines(extractedRis$raw, con = filename)
    return(filename)
}

filterRis <- function(extractedRis, idToSupp, writeFile = NA) {
    filteredRis <- list()
    linesToSupp <- extractedRis$lineRegId %in% idToSupp
    filteredRis$fieldName <- extractedRis$fieldName[!linesToSupp]
    if ("raw" %in% names(extractedRis)) {
        filteredRis$raw <- extractedRis$raw[!linesToSupp]
    }
    if ("content" %in% names(extractedRis)) {
        filteredRis$content <- extractedRis$content[!linesToSupp]
    }
    filteredRis$lineRegId <- extractedRis$lineRegId[!linesToSupp]
    filteredRis$registers <- data.frame(id = unique(filteredRis$lineRegId),
        begin = tapply(1:length(filteredRis$lineRegId), filteredRis$lineRegId,
            min), end = tapply(1:length(filteredRis$lineRegId), filteredRis$lineRegId,
            max))
    filteredRis$nbRecords <- nrow(filteredRis$registers)
    if (!"nbDeleted" %in% names(extractedRis)) {
        filteredRis$nbDeleted <- extractedRis$nbRecords - filteredRis$nbRecords
    } else {
        filteredRis$nbDeleted <- extractedRis$nbDeleted + (extractedRis$nbRecords -
            filteredRis$nbRecords)
    }
    if (!is.na(writeFile)) {
        print(paste(writeRis(filteredRis, filename = writeFile), "written!"))
    }
    return(filteredRis)
}
```

## 3.2 Finding duplicates in one Ris

```
findDupes <- lapply(risExtracted, risInternalDuplicate)
```

Example of the scopus Ris File.
```
A <- rbind(data.frame(supp = F, findDupes$scopus$tabInfo[unique(findDupes$scopus$accepted_dupes$ref),
    ]), data.frame(supp = T, findDupes$scopus$tabInfo[findDupes$scopus$accepted_dupes$toSupp,
    ]))
A[order(A$TI), c("id", "TY", "TI", "PY", "DO", "AU")] %>%
    kable(booktabs = T, longtable = T) %>%
    row_spec(which(A$supp[order(A$TI)]), background = "red") %>%
    row_spec(which(!A$supp[order(A$TI)]), background = "green") %>%
    column_spec(4, width = "4cm") %>%
    column_spec(7, width = "4cm") %>%
    landscape()
```

| | id | TY | TI | PY | DO | AU |
|---|---|---|---|---|---|---|
| 3410 | 3410 | JOUR | A bayesian approach to more stable estimates of group-level effects in contextual studies | 2015 | 10.1080/00273171.2015.1090899 | Zitzmann, S.|Lüdtke, O.|Robitzsch, A. |
| 5158 | 5158 | JOUR | A bayesian approach to more stable estimates of group-level effects in contextual studies | 2015 | 10.1080/00273171.2015.1090899 | Zitzmann, S.|Lüdtke, O.|Robitzsch, A. |
| 1617 | 1617 | JOUR | A committee on well-being of medical students and house staff | 1983 | NA | Weinstein, H.M. |
| 6509 | 6509 | JOUR | A committee on well-being of medical students and house staff | 1983 | NA | Weinstein, H.M. |
| 14600 | 14600 | CONF | A new program in sustainable engineering: A platform for integrating research and service into the classroom through global engagement | 2015 | NA | Brennan, R.A. |
| 8897 | 8897 | CHAP | A New Program in Sustainable Engineering: A Platform for Integrating Research and Service into the Classroom Through Global Engagement | 2016 | 10.1007/978-3-319-32933-8_2 | Brennan, R.A.|Riley, D.R. |
| 10657 | 10657 | JOUR | Achievements in environmental education and urban sustainability in Mexico | 2016 | NA | Muñoz-Cadena, C.E.|Estrada-Izquierdo, I.E.|Morales-Pérez, R.E. |
| 7691 | 7691 | JOUR | Achievements in Environmental Education and Urban Sustainability in Mexico | 2016 | NA | Muñoz-Cadena, C.E.|Estrada-Izquierdo, I.E.|Morales-Pérez, R.E. |
| 14957 | 14957 | JOUR | An engineering curriculum track for IT for sustainability | 2012 | NA | Morell, L.|Bash, C.|Trucco, M.|Patel, C. |
| 9198 | 9198 | CONF | An engineering curriculum track for IT for sustainability | 2012 | 10.1109/FIE.2012.6462324 | Morell, L.|Trucco, M.|Bash, C.|Patel, C. |
| 400 | 400 | JOUR | Anatomy education in a changing medical curriculum | 1998 | 10.1002/(SICI)1097-0185(199802)253:1<28::AID-AR11>3.0.CO;2-E | Drake, R.L. |
| 2697 | 2697 | JOUR | Anatomy education in a changing medical curriculum | 1999 | NA | Drake, R.L. |
| 6286 | 6286 | CHAP | Beyond innovation in mobile learning: Towards sustainability in schools | 2016 | 10.4324/9780203076095 | Cerratto-Pargman, T.|Milrad, M. |
| 9065 | 9065 | CHAP | Beyond innovation in mobile learning: Towards sustainability in schools | 2015 | 10.4324/9780203076095-10 | Cerratto-Pargman, T.|Milrad, M. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 9154 | 9154 | CHAP | Conferences as learning spaces on climate change and sustainability: Insights from university students' experiences | 2015 | 10.4018/978-1-4666-8764-6.ch003 | | El Zoghbi, M.B. |
| 10668 | 10668 | CHAP | Conferences as learning spaces on climate change and sustainability: Insights from university students' experiences | 2016 | 10.4018/978-1-5225-0803-8.ch042 | | El Zoghbi, M.B. |
| 14386 | 14386 | CHAP | Crafting pedagogical pathways that disrupt and transform anthropocentric mindsets of higher education students | 2017 | 10.1007/978-3-319-47877-7_9 | | Tillmanns, T.|Holland, C. |
| 8925 | 8925 | CHAP | Crafting Pedagogical Pathways that Disrupt and Transform Anthropocentric Mindsets of Higher Education Students | 2017 | 10.1007/978-3-319-47868-5_19 | | Tillmanns, T.|Holland, C. |
| 14536 | 14536 | CHAP | Creating a sustainable program for ethics in student affairs: A look at climate, assessment, and curriculum | 2014 | 10.4018/978-1-4666-6433-3.ch017 | | Van Wingerden, C.|Ellis, M.|Pratt, T.W. |
| 14581 | 14581 | CHAP | Creating a sustainable program for ethics in student affairs: A look at climate, assessment, and curriculum | 2014 | 10.4018/978-1-4666-6046-5.ch077 | | Van Wingerden, C.|Ellis, M.|Pratt, T.W. |
| 7622 | 7622 | CHAP | Design principles for an intelligent-augmented-reality-based m-learning application to improve engineering students' english language skills | 2017 | 10.4018/978-1-5225-2110-5.ch011 | | Bozdogan, D.|Kasap, B.|Kose, U. |
| 13841 | 13841 | CHAP | Design principles for an intelligent-Augmented-Reality-based M-learning application to improve engineering students' English language skills | 2018 | 10.4018/978-1-5225-5469-1.ch018 | | Bozdogan, D.|Kasap, B.|Köse, U. |
| 13990 | 13990 | CHAP | Design Principles for an intelligent-augmented-reality-based m-learning application to improve engineering students' English language skills | 2018 | 10.4018/978-1-5225-5643-5.ch036 | | Bozdogan, D.|Kasap, B.|Kose, U. |
| 10874 | 10874 | CHAP | Designing sustainability curricula: A case following chemical engineering curriculum redesign | 2014 | 10.4018/978-1-4666-5856-1.ch024 | | Sheehan, M. |

| | | | | | | |
|---|---|---|---|---|---|---|
| 10959 | 10959 | CHAP | Designing sustainability curricula: A case following chemical engineering curriculum redesign | 2015 | 10.4018/978-1-4666-8246-7.ch040 | Sheehan, M. |
| 14439 | 14439 | CHAP | Designing sustainability curricula: A case following chemical engineering curriculum redesign | 2015 | 10.4018/978-1-4666-8246-7.ch040 | Sheehan, M. |
| 11058 | 11058 | CONF | Developing sustainable architecture education approaches in Malaysia: A case study of critiques session in 2 nd Year design studio of Architecture Department, National University of Malaysia (UKM) | 2011 | 10.4028/www.scientific.net/AMM.71-78.5003 | Utaberta, N.|Hassanpour, B.|Abdullah, N.A.G.|Tahir, M.|Che Ani, A. |
| 14980 | 14980 | CONF | Developing sustainable architecture education approaches in Malaysia: A case study of critiques session in 2nd year design studio of architecture department, National University of Malaysia (UKM) | 2011 | NA | Utaberta, N.|Hassanpour, B.|Che-Ani, A.I.|Zaharim, A.|Abdullah, N.A.G. |
| 4681 | 4681 | JOUR | EarthLabs: Supporting teacher professional development to facilitate effective teaching of climate science | 2014 | 10.5408/13-059.1 | Ellins, K.K.|Ledley, T.S.|Haddad, N.|McNeal, K.|Gold, A.|Lynds, S.|Libarkin, J. |
| 10942 | 10942 | JOUR | EarthLabs: Supporting teacher professional development to facilitate effective teaching of climate science | 2014 | 10.5408/13-059.1 | Ellins, K.K.|Ledley, T.S.|Haddad, N.|McNeal, K.|Gold, A.|Lynds, S.|Libarkin, J. |
| 9150 | 9150 | CHAP | Education collaboration development: A blended shore education approach to sustainable development | 2013 | 10.4018/978-1-4666-4458-8.ch026 | Strohschen, G. |
| 14675 | 14675 | CHAP | Education collaboration development: A blended shore education approach to sustainable development | 2013 | 10.4018/978-1-4666-4852-4.ch082 | Strohschen, G. |
| 5568 | 5568 | CHAP | Educational solutions | 2005 | 10.1016/B978-012088441-4/50004-6 | Mihelcic, J.R.|Hokanson, D.R. |
| 15343 | 15343 | CHAP | Educational solutions | 2005 | 10.1016/B978-012088441-4/50004-6 | Mihelcic, J.R.|Hokanson, D.R. |
| 8115 | 8115 | JOUR | Effect of Nutrition Education on Diet Quality, Sustainable Nutrition and Eating Behaviors among University Students | 2022 | 10.1080/07315724.2021.1955420 | Yolcuoğlu, İ.Z.|Kızıltan, G. |

| | | | | | | |
|---|---|---|---|---|---|---|
| 13451 | 13451 | JOUR | Effect of Nutrition Education on Diet Quality, Sustainable Nutrition and Eating Behaviors among University Students | 2021 | 10.1080/07315724.2021.1955420 | Yolcuoğlu, İ.Z.|Kızıltan, G. |
| 9252 | 9252 | CONF | Engineering education for a sustainable, just and peaceful society | 2010 | NA | Zandvoort, H. |
| 15119 | 15119 | CONF | Engineering education for a sustainable, just and peaceful society | 2010 | NA | Zandvoort, H. |
| 15244 | 15244 | CONF | Enhancing bioengineering opportunities for engineering majors | 2003 | 10.1109/FIE.2003.1265972 | Tyson, V.K.|Baumann, M.J.|Baldwin, R.G.|Fisher, P.D. |
| 15381 | 15381 | CONF | Enhancing bioengineering opportunities for engineering majors | 2003 | NA | Tyson, V.K.|Baumann, M.J.|Baldwin, R.G.|Fisher, P.D. |
| 10876 | 10876 | JOUR | Establishing an applied training session in metrology at an agricultural engineering school (CIRAD, Montpellier, France) | 2014 | 10.1051/ijmqe/2014018 | Calchera, G.|Linck, O.|Bouniol, A.|Ricci, J.|Servent, A.|Grevesse, C.|Martin, L. |
| 14752 | 14752 | CONF | Establishing an applied training session in metrology at an agricultural engineering school (CIRAD, Montpellier, France) | 2014 | NA | Calchera, G.|Linck, O.|Bouniol, A.|Ricci, J.|Servent, A.|Grevesse, C.|Martin, L. |
| 14447 | 14447 | JOUR | Experiment in teaching sustainable concept: Integrating sustainable design in interior design curriculum | 2015 | NA | Rashdan, W. |
| 14456 | 14456 | JOUR | Experiment in teaching sustainable concept: Integrating sustainable design in interior design curriculum | 2015 | 10.18848/2325-128x/cgp/v09i03/57911 | Rashdan, W. |
| 2593 | 2593 | JOUR | From concept to commercialisation: student learning in a sustainable engineering innovation project | 2007 | 10.1080/03043790601118689 | SchäFer, A.I.|Richards, B.S. |
| 5057 | 5057 | JOUR | From concept to commercialisation: student learning in a sustainable engineering innovation project | 2007 | 10.1080/03043790601118689 | Schäfer, A.I.|Richards, B.S. |
| 15357 | 15357 | CONF | Higher education in hydrology for sustainable development in the 21st century | 1999 | NA | Qingqiu, C.|Zhongji, Z.|Yunhao, C. |

| 15393 | 15393 | CONF | Higher education in hydrology for sustainable development in the 21st century | 1999 | NA | Qingqiu, C.|Zhongji, Z.|Yunhao, C. |
| 93 | 93 | JOUR | How much do engineering students know about sustainable development? The findings of an international survey and possible implications for the engineering curriculum | 2005 | 10.1080/03043790512331313804 | Azapagic, A.|Perdan, S.|Shallcross, D. |
| 647 | 647 | JOUR | How much do engineering students know about sustainable development? The findings of an international survey and possible implications for the engineering curriculum | 2005 | 10.1080/03043790512331313804 | Azapagic, A.|Perdan, S.|Shallcross, D. |
| 6587 | 6587 | CHAP | IEA study in civic education | 2010 | 10.1016/B978-0-08-044894-7.01478-0 | Torney-Purta, J.|Amadeo, J.-A.|Schwille, J. |
| 15023 | 15023 | CHAP | IEA Study in Civic Education | 2009 | 10.1016/B978-0-08-044894-7.01478-0 | Torney-Purta, J.|Amadeo, J.-A.|Schwille, J. |
| 4716 | 4716 | JOUR | Improving climate change communication starting with environmental educators | 2014 | 10.5408/13-062.1 | Lambert, J.L.|Bleicher, R.E. |
| 14443 | 14443 | JOUR | Improving climate change communication starting with environmental educators | 2014 | 10.5408/13-062.1 | Lambert, J.L.|Bleicher, R.E. |
| 5072 | 5072 | JOUR | Improving climate literacy with project-based modules rich in educational rigor and relevance | 2014 | 10.5408/13-056.1 | DeWaters, J.E.|Andersen, C.|Calderwood, A.|Powers, S.E. |
| 10989 | 10989 | JOUR | Improving climate literacy with project-based modules rich in educational rigor and relevance | 2014 | 10.5408/13-056.1 | DeWaters, J.E.|Andersen, C.|Calderwood, A.|Powers, S.E. |
| 9186 | 9186 | CHAP | Including children with disabilities in regular classes: Current climate and future direction in Malaysia | 2012 | 10.4324/9780203113585 | Ibrahim, H. |
| 14844 | 14844 | CHAP | Including children with disabilities in regular classes: Current climate and future direction in Malaysia | 2012 | 10.4324/9780203113585-12 | Ibrahim, H. |
| 8882 | 8882 | CONF | Inclusion of sustainability analysis in a national airport design competition | 2018 | NA | Johnson, M.E.|Gu, Y.|Holtaway, L.E. |

| 13651 | 13651 | CONF | Inclusion of sustainability analysis in a national airport design competition | 2019 | NA | Johnson, M.E.|Gu, Y.|Holtaway, L.E. |
|---|---|---|---|---|---|---|
| 14345 | 14345 | CHAP | Integrating sustainable engineering principles in material science engineering education | 2016 | 10.4018/978-1-4666-9619-8.ch059 | Mainali, B.|Petrolito, J.|Russell, J.|Ionescu, D.|Al Abadi, H. |
| 15611 | 15611 | CHAP | Integrating sustainable engineering principles in material science engineering education | 2015 | 10.4018/978-1-4666-8183-5.ch014 | Mainali, B.|Russell, J.|Petrolito, J.|Ionescu, D.|Al Abadi, H. |
| 14432 | 14432 | CHAP | Intercultural communication and sustainable leadership: The case of a joint master course | 2015 | 10.4018/978-1-4666-8246-7.ch074 | Kostoulas-Makrakis, N. |
| 14524 | 14524 | CHAP | Intercultural communication and sustainable leadership: The case of a joint master course | 2015 | 10.4018/978-1-4666-8246-7.ch074 | Kostoulas-Makrakis, N. |
| 14692 | 14692 | CHAP | Intercultural communication and sustainable leadership: The case of a joint master course | 2014 | 10.4018/978-1-4666-5856-1.ch036 | Kostoulas-Makrakis, N. |
| 7898 | 7898 | JOUR | Internationalization of the curriculum through student-led climate change teaching activity | 2014 | 10.5408/13-033.1 | McGregor, H.V.|O'Shea, B.|Brewer, C.|Abuodha, P.|Pharo, E.J. |
| 14577 | 14577 | JOUR | Internationalization of the curriculum through student-led climate change teaching activity | 2014 | 10.5408/13-033.1 | McGregor, H.V.|O'Shea, B.|Brewer, C.|Abuodha, P.|Pharo, E.J. |
| 2535 | 2535 | JOUR | Investigating climate change understandings of urban middle-level students | 2014 | 10.5408/13-042.1 | Bodzin, A.M.|Anastasio, D.|Sahagian, D.|Peffer, T.|Dempsey, C.|Steelman, R. |
| 14511 | 14511 | JOUR | Investigating climate change understandings of urban middle-level students | 2014 | 10.5408/13-042.1 | Bodzin, A.M.|Anastasio, D.|Sahagian, D.|Peffer, T.|Dempsey, C.|Steelman, R. |
| 9465 | 9465 | JOUR | Issues in engineering education | 1985 | 10.1061/(ASCE)1052-3928(1985)111:3(108) | Muspratt, M.A. |
| 11381 | 11381 | JOUR | ISSUES IN ENGINEERING EDUCATION. | 1985 | 10.1061/(ASCE)1052-3928(1985)111:3(108) | Muspratt, Murray A. |
| 5154 | 5154 | JOUR | JiFUNzeni: A blended learning approach for sustainable teachers' professional development | 2014 | NA | Onguko, B.B. |
| 14653 | 14653 | CONF | JiFUNzeni: A blended learning approach for sustainable teachers' professional development | 2013 | NA | Onguko, B. |

| | | | | | | |
|---|---|---|---|---|---|---|
| 14571 | 14571 | CHAP | Learning about sustainability in a non-formal laboratory context for secondary level students: A module on climate change, the ozone hole, and summer smog | 2014 | 10.4018/978-1-4666-7363-2.ch046 | Garner, N.\|Lischke, M.L.\|Siol, A.\|Eilks, I. |
| 15660 | 15660 | CHAP | Learning about sustainability in a non-formal laboratory context for secondary level students: A module on climate change, the ozone hole, and summer smog | 2014 | 10.4018/978-1-4666-5856-1.ch012 | Garner, N.\|De Lourdes Lischke, M.\|Siol, A.\|Eilks, I. |
| 14262 | 14262 | CHAP | Learning and climate change adaptation: Moving towards resilience in an era of escalating instability | 2016 | 10.4018/978-1-5225-0803-8.ch074 | Wilson, L.A. |
| 14535 | 14535 | CHAP | Learning and climate change adaptation: Moving towards resilience in an era of escalating instability | 2015 | 10.4018/978-1-4666-8764-6.ch013 | Wilson, L.A. |
| 14318 | 14318 | CHAP | Life quality and health of women workers from family agriculture | 2016 | NA | Pereira, A.L.S.\|Sá, F.D.\|Almeida, J.J.\|Nascimento, L.S.\|Carvalho, C.J.\|Dantas, D.S. |
| 14427 | 14427 | CONF | Life quality and health of women workers from family agriculture | 2016 | NA | Pereira, A.L.S.\|Sá, F.D.\|Almeida, J.J.\|Nascimento, L.S.\|Carvalho, C.J.\|Dantas, D.S. |
| 14700 | 14700 | CONF | Live energy: An initiative for teaching energy and sustainability topics with the most up-to-date and relevant content | 2013 | NA | Ehlig-Economides, C.\|Yalvac, B.\|Binks-Cantrell, E.\|Gilman, D.R.\|Ayar, M.\|Robinson, T.A.\|Kovscek, A.R.\|Toossi, R.\|Pisupati, S.V.\|Aghara, S.K. |
| 14824 | 14824 | CONF | Live energy: An initiative for teaching energy and sustainability topics with the most up-to-date and relevant content | 2012 | NA | Ehlig-Economides, C.\|Aghara, S.K.\|Pisupati, S.V.\|Toossi, R.\|Kovscek, A.R.\|Ayar, M.\|Binks-Cantrell, E.\|Gilman, D.R.\|Smith, D.L.\|Robinson, T.A.\|Yalvac, B. |
| 10871 | 10871 | JOUR | Online and on-campus historical geology students' prior ideas about global climate change | 2014 | 10.5408/12-385.1 | Baldwin, K.A.\|Cooper, C.M. |
| 14557 | 14557 | JOUR | Online and on-campus historical geology students' prior ideas about global climate change | 2014 | 10.5408/12-385.1 | Baldwin, K.A.\|Cooper, C.M. |

| 10338 | 10338 | CONF | Participatory digital educational game production with fifth graders on biodiversity (SDG15) | 2020 | 10.34190/GBL.20.018 | Küchler, M.G. |
|---|---|---|---|---|---|---|
| 16020 | 16020 | CONF | Participatory digital educational game production with fifth graders on biodiversity (SDG15) | 2020 | 10.34190/GBL.20.018 | Küchler, M.G. |
| 14542 | 14542 | CHAP | Pass, fail, or incomplete? Analyzing environmental education in nova scotia's sixth grade curriculum | 2014 | 10.4018/978-1-4666-7363-2.ch068 | Spence, E.|Wright, T.|Castleden, H. |
| 14618 | 14618 | CHAP | Pass, fail, or incomplete? Analyzing environmental education in Nova Scotia's sixth grade curriculum | 2014 | 10.4018/978-1-4666-5856-1.ch011 | Spence, E.|Wright, T.|Castleden, H. |
| 14368 | 14368 | CHAP | Performance evaluation in teaching environments on climate change | 2016 | NA | Rodrigues, R.|Silva, L.|Souza, E.|Souza, A.|Brito, F. |
| 10906 | 10906 | CONF | Performance evaluation in teaching environments on climate change | 2016 | NA | Rodrigues, R.|Silva, L.|Souza, E.|Souza, A.|Brito, F. |
| 15448 | 15448 | JOUR | School discipline in the United States: prevention, correction and long term social development | 1998 | 10.53841/bpsecp.1998.15.1.15 | Bear, G.G. |
| 821 | 821 | JOUR | School discipline in the United States: Prevention, correction, and long-term social development | 1998 | NA | Bear, G.G. |
| 14470 | 14470 | CHAP | Sharing pds teacher expertise with pre-service teachers through videos | 2014 | 10.4018/978-1-4666-6367-1.ch010 | Siebert, C.J.|Wyss, V.L.|Jackson, T. |
| 14410 | 14410 | CHAP | Sharing PDS teacher expertise with pre-service teachers through videos | 2015 | 10.4018/978-1-4666-8632-8.ch084 | Siebert, C.J.|Wyss, V.L.|Jackson, T. |
| 14323 | 14323 | CONF | Stand-Alone educational center | 2017 | NA | Shali, R.A.|Qadir, G.|Naffar, A.E.|Hittini, B.|Obaid, A.|Rida, H. |
| 14195 | 14195 | CONF | STAND-ALONE EDUCATIONAL CENTER | 2017 | 10.14455/ISEC.res.2017.198 | Al Shali, R.|Qadir, G.|El Naffar, A.|Hittini, B.|Obaid, A.|Rida, H. |
| 2610 | 2610 | JOUR | Supporting and understanding students' epistemological discourse about climate change | 2014 | 10.5408/13-036.1 | Holthuis, N.|Lotan, R.|Saltzman, J.|Mastrandrea, M.|Wild, A. |
| 10954 | 10954 | JOUR | Supporting and understanding students' epistemological discourse about climate change | 2014 | 10.5408/13-036.1 | Holthuis, N.|Lotan, R.|Saltzman, J.|Mastrandrea, M.|Wild, A. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 14634 | 14634 | CHAP | Supporting sustainability education and leadership: Strategies for students, faculty, and the planet | 2013 | 10.4018/978-1-4666-4852-4.ch081 | | Cassidy, A.|Sipos, Y.|Nyrose, S. |
| 15503 | 15503 | CHAP | Supporting sustainability education and leadership: Strategies for students, faculty, and the planet | 2013 | 10.4018/978-1-4666-4458-8.ch012 | | Cassidy, A.|Sipos, Y.|Nyrose, S. |
| 14407 | 14407 | CHAP | Sustainability in higher education through basic science research: Strategies for corporate bodies in pharmaceuticals | 2014 | 10.4018/978-1-4666-7363-2.ch035 | | Yogeeswari, P.|Sriram, D. |
| 14606 | 14606 | CHAP | Sustainability in higher education through basic science research: Strategies for corporate bodies in pharmaceuticals | 2013 | 10.4018/978-1-4666-2845-8.ch004 | | Yogeeswari, P.|Sriram, D. |
| 11147 | 11147 | CONF | Sustainability initiatives at several U.S. institutions of higher education | 2011 | NA | | Sharma, M.P.|Winslett, M.|Peters, R.W. |
| 14820 | 14820 | CONF | Sustainability initiatives at several U.S. Institutions of higher education | 2011 | NA | | Sharma, M.P.|Winslett, M.|Peters, R.W. |
| 14832 | 14832 | CONF | Sustainability initiatives at several U.S. Institutions of higher education | 2011 | NA | | Sharma, M.P.|Winslett, M.|Peters, R.W. |
| 14981 | 14981 | CONF | Sustainability initiatives at several U.S. Institutions of higher education | 2011 | NA | | Sharma, M.P.|Winslett, M.|Peters, R.W. |
| 14854 | 14854 | CONF | Sustainability initiatives at Several U.S. Institutions of Higher Education | 2011 | NA | | Sharma, M.P.|Winslett, M.|Peters, R.W. |
| 14413 | 14413 | CHAP | Sustainability, ethics and education: What's accounting got to do with it? | 2015 | 10.4018/978-1-4666-8195-8.ch049 | | Khan, T. |
| 14472 | 14472 | CHAP | Sustainability, ethics and education: What's accounting got to do with it? | 2014 | 10.4018/978-1-4666-5856-1.ch021 | | Khan, T. |
| 9052 | 9052 | CHAP | Sustainable development in business education: The role of entrepreneurship as pedagogy | 2014 | 10.4018/978-1-4666-5856-1.ch006 | | Lourenço, F.|Sappleton, N.|Lei, W.S.|Cheng, R. |
| 14494 | 14494 | CHAP | Sustainable development in business education: The role of entrepreneurship as pedagogy | 2015 | 10.4018/978-1-4666-8468-3.ch072 | | Lourenço, F.|Sappleton, N.|Lei, W.S.|Cheng, R. |
| 7886 | 7886 | CHAP | Sustainable educational reform | 2010 | 10.1016/B978-0-08-044894-7.01096-4 | | Hargreaves, A. |
| 15129 | 15129 | CHAP | Sustainable Educational Reform | 2009 | 10.1016/B978-0-08-044894-7.01096-4 | | Hargreaves, A. |

| 5558 | 5558 | JOUR | Sustainable happiness, living campus, and wellbeing for all | 2016 | NA | O'Brien, C.|Adam, C. |
| 3993 | 3993 | JOUR | Sustainable happiness, Living Campus, and wellbeing for all | 2016 | NA | O'Brien, C.|Adam, C. |
| 9105 | 9105 | CONF | Sustainable minerals education - We care, but do you? | 2011 | NA | Laurence, D.|Hebblewhite, B. |
| 11159 | 11159 | CONF | Sustainable minerals education - We care, but do you? | 2011 | NA | Laurence, D.|Hebblewhite, B. |
| 11036 | 11036 | CHAP | Sustainable sport education in primary education: An English case study | 2011 | 10.4324/9780203807156-7 | O'Donovan, T.M.|Macphail, A.|Kirk, D. |
| 14941 | 14941 | CHAP | SUSTAINABLE SPORT EDUCATION IN PRIMARY EDUCATION: An English case study | 2011 | 10.4324/9780203807156-3 | O'donovan, T.M.|Macphail, A.|Kirk, D. |
| 10692 | 10692 | CHAP | Systematic support for STEM pre-service teachers: An authentic and sustainable four-pillar professional development model | 2016 | 10.4018/978-1-4666-9624-2.ch058 | Rogers, R.R.H.|Winship, J.|Sun, Y. |
| 7617 | 7617 | CHAP | Systematic support for STEM pre-service teachers: An authentic and sustainable four-pillar professional development model | 2015 | 10.4018/978-1-4666-9471-2.ch005 | Rogers, R.R.H.|Winship, J.|Sun, Y. |
| 14241 | 14241 | CHAP | Systematic support for STEM pre-service teachers: An authentic and sustainable four-pillar professional development model | 2016 | 10.4018/978-1-5225-0978-3.ch029 | Rogers, R.R.H.|Winship, J.|Sun, Y. |
| 14349 | 14349 | CHAP | Systematic support for STEM pre-service teachers: An authentic and sustainable four-pillar professional development model | 2016 | 10.4018/978-1-5225-1624-8.ch027 | Rogers, R.R.H.|Winship, J.|Sun, Y. |
| 14417 | 14417 | CHAP | Tackling climate change through educational awareness: A case study on Georgia house resolution 689 | 2015 | 10.4018/978-1-4666-8764-6.ch008 | Drenner, K. |
| 14296 | 14296 | CHAP | Tackling climate change through educational awareness: A case study on Georgia House Resolution 689 | 2016 | 10.4018/978-1-5225-1671-2.ch057 | Drenner, K. |
| 8887 | 8887 | JOUR | Teachers Perceived Disruption at School and Related Variables from Teachers and School Functioning | 2017 | 10.1017/sjp.2017.67 | Fernández, M.B.M.|Gómez, J.C.C.|Babarro, J.M.|Jalón, M.J.D.-A.|Arias, R.M. |

| 10628 | 10628 | JOUR | Teachers' Perceived Disruption at School and Related Variables from Teachers and School Functioning | 2017 | 10.1017/sjp.2017.67 | Martínez Fernández, M.B.|Chacón Gómez, J.C.|Martín Babarro, J.|Díaz-Aguado Jalón, M.J.|Martínez Arias, R. |
|---|---|---|---|---|---|---|
| 5995 | 5995 | CONF | Teaching multinational, multidisciplinary sustainable product development | 2008 | 10.1115/DETC2008-49388 | Agogino, A.M.|Beckman, S.L.|Borja, V.|López, M.|Shedroff, N.|Ramírez, A.C. |
| 11228 | 11228 | CONF | Teaching multinational, multidisciplinary sustainable product development | 2009 | NA | Agogino, A.M.|Beckman, S.L.|Borja, V.|López, M.|Shedroff, N.|Ramírez, A.C. |
| 6242 | 6242 | CONF | Teaching sustainable engineering and industrial ecology using a hybrid problem-project based learning approach | 2017 | NA | Sriraman, V.|Torres, A.|Ortiz, A.M. |
| 10652 | 10652 | CONF | Teaching sustainable engineering and industrial ecology using a hybrid problem-project based learning approach | 2017 | NA | Sriraman, V.|Torres, A.|Ortiz, A.M. |
| 2366 | 2366 | JOUR | Teaching sustainable entrepreneurship to engineering students: the case of Delft University of Technology | 2006 | 10.1080/03043790600566979 | Bonnet, H.|Quist, J.|Hoogwater, D.|Spaans, J.|Wehrmann, C. |
| 3895 | 3895 | JOUR | Teaching sustainable entrepreneurship to engineering students: the case of Delft University of Technology | 2006 | 10.1080/03043790600566979 | Bonnet, H.|Quist, J.|Hoogwater, D.|Spaans, J.|Wehrmann, C. |
| 11329 | 11329 | JOUR | The Baltic Sea project: an international challenge for environmental education | 1995 | NA | Jaaskelainen, L. |
| 15270 | 15270 | JOUR | The Baltic Sea project: an international challenge for environmental education | 1995 | NA | Jaaskelainen, L. |
| 13808 | 13808 | JOUR | The educator's perspective: Knowledge, attitude and practices on occupational safety and health at school among primary and secondary school teachers | 2019 | NA | Yusoff, N.B.M.|How, V.|Azmi, E.S.B.|Bin Othman, K. |

| | | | | | | |
|---|---|---|---|---|---|---|
| 8782 | 8782 | JOUR | THE EDUCATOR'S PERSPECTIVE: KNOWLEDGE, ATTITUDE AND PRACTICES ON OCCUPATIONAL SAFETY AND HEALTH AT SCHOOL AMONG PRIMARY AND SECONDARY SCHOOL TEACHERS | 2019 | 10.37268/MJPHM/VOL.19/NO.1/ART.52 | Mohamad Yusoff, N.B.|How, V.|Binti Azmi, E.S.|Othman, K.B. |
| 7687 | 7687 | CHAP | The Sustainability Journey of USM: Solution Oriented Campus Ecosphere for Vitalising Higher Education Action on GAP | 2017 | 10.1007/978-3-319-47895-1_23 | Osman, O.|Ibrahim, K.|Koshy, K.|Akib, N.A.M.|Shabudin, A.F.A. |
| 10714 | 10714 | CHAP | The Sustainability Journey of USM: Solution Oriented Campus Ecosphere for Vitalising Higher Education Action on GAP | 2017 | 10.1007/978-3-319-47868-5_17 | Osman, O.|Ibrahim, K.|Koshy, K.|Akib, N.A.M.|Shabudin, A.F.A. |
| 7818 | 7818 | CONF | The teaching of sustainable development at Aston University | 2008 | NA | Fletcher, J.P.|Drahun, J.A.|Davies, P.A.|Knowles, P. |
| 15052 | 15052 | CONF | The teaching of sustainable development at Aston University | 2009 | NA | Fletcher, J.P.|Drahun, J.A. |
| 11171 | 11171 | CONF | Towards a pedagogical model of technology integration: Using web 2.0 technologies to learn critical citizenship at a South African University | 2012 | NA | Rambe, P. |
| 14785 | 14785 | CONF | Towards a pedagogical model of technology integration: Using web 2.0 technologies to learn critical citizenship at a South African University | 2012 | NA | Rambe, P. |
| 9214 | 9214 | CHAP | Transforming mindsets through education for sustainable development | 2010 | 10.1016/B978-0-08-044894-7.00224-4 | Segovia, V.M. |
| 15056 | 15056 | CHAP | Transforming Mindsets Through Education for Sustainable Development | 2009 | 10.1016/B978-0-08-044894-7.00224-4 | Segovia, V.M. |
| 14587 | 14587 | CHAP | Using project-based learning to teach sustainability issues to elementary students | 2014 | 10.4018/978-1-4666-7363-2.ch042 | Weiland, I.|Pokral, E.|Cook, K. |
| 14642 | 14642 | CHAP | Using project-based learning to teach sustainability issues to elementary students | 2014 | 10.4018/978-1-4666-5856-1.ch010 | Weiland, I.|Pokral, E.|Cook, K. |

| | | | | | | |
|---|---|---|---|---|---|---|
| 6364 | 6364 | JOUR | Using student-generated blogs to create a global perspective on climate change | 2014 | 10.5408/13-065.1 | Schuenemann, K.|Wagner, R. |
| 10816 | 10816 | JOUR | Using student-generated blogs to create a global perspective on climate change | 2014 | 10.5408/13-065.1 | Schuenemann, K.|Wagner, R. |
| 11178 | 11178 | CHAP | What do university students learn? | 2010 | 10.1016/B978-0-08-044894-7.00864-2 | Kuh, G.D.|Bridges, B.K. |
| 15048 | 15048 | CHAP | What Do University Students Learn? | 2009 | 10.1016/B978-0-08-044894-7.00864-2 | Kuh, G.D.|Bridges, B.K. |
| 11113 | 11113 | CHAP | Young children and sustainable consumption: An early childhood education agenda | 2013 | 10.4324/9780203146057 | Nichols, S. |
| 14630 | 14630 | CHAP | Young children and sustainable consumption: An early childhood education agenda | 2013 | 10.4324/9780203146057-14 | Nichols, S. |
| 2569 | 2569 | JOUR | Zoo and aquarium professionals' concerns and confidence about climate change education | 2014 | 10.5408/13-048.1 | Swim, J.|Fraser, J. |
| 14482 | 14482 | JOUR | Zoo and aquarium professionals' concerns and confidence about climate change education | 2014 | 10.5408/13-048.1 | Swim, J.|Fraser, J. |

```
int_filt_ris <- mapply(function(ris, dupe) {
    filterRis(ris, dupe$toSupp)
}, risExtracted, findDupes, SIMPLIFY = F)
```

## 3.3 Filtering a Ris file from the records of another

```
finalRis <- list(scopus = int_filt_ris$scopus)
```

### 3.3.1 Wos: Core Collection from scopus

```
filWosCore <- compareRisDuplicate(int_filt_ris$WoS_core, finalRis$scopus)
```

```
## 7480 / 9146 records are already in the reference file
finalRis$WoS_core <- filterRis(extractedRis = int_filt_ris$WoS_core, idToSupp = filWosCore$toSupp)
```

### 3.3.2 Wos: SciELO from scopus and WoS Core Collection

```
filWosSciELO <- compareRisDuplicate(int_filt_ris$WoS_SciELO, finalRis$scopus)
```

```
## 59 / 428 records are already in the reference file
int_filt_ris$WoS_SciELO <- filterRis(extractedRis = int_filt_ris$WoS_SciELO,
    idToSupp = filWosSciELO$toSupp)
filWosSciELO <- compareRisDuplicate(int_filt_ris$WoS_SciELO, finalRis$WoS_core)
```

```
## 16 / 369 records are already in the reference file
finalRis$WosScielo <- filterRis(int_filt_ris$WoS_SciELO, idToSupp = filWosSciELO$toSupp)
```

# 4 Writing final Ris files

```
DIR <- "../../Search/finalRis"
if (dir.exists(DIR)) {
    unlink(DIR, recursive = T)
}
dir.create(DIR)
writeRis(finalRis$scopus, filename = paste(DIR, "scopus.ris", sep = "/"))
```

```
## [1] "../../Search/finalRis/scopus.ris"
writeRis(finalRis$WoS_core, filename = paste(DIR, "WoS_Core.ris", sep = "/"))
```

```
## [1] "../../Search/finalRis/WoS_Core.ris"
writeRis(finalRis$WosScielo, filename = paste(DIR, "WoS_SciELO.ris", sep = "/"))
```

```
## [1] "../../Search/finalRis/WoS_SciELO.ris"
```