**AMSTERDAM EFFICIENT DEDUPLICATION (AED) METHOD**

**HANDLEIDING**

|  |  |
| --- | --- |
| René Otten, Linda Schoonmade, Ralph de Vries  Universiteitsbibliotheek, afdeling Research Support  Vrije Universiteit Amsterdam |  |

Inhoud

[INTRODUCTION 3](#_Toc34729742)

[1. PubMed 4](#_Toc34729743)

[1.1 Verzamelen ID’s uit PubMed (tijdens search) 4](#_Toc34729744)

[1.2 Verzamelen ID’s uit EndNote (achteraf) 4](#_Toc34729745)

[1.3 Verwijderen oude PubMed records in toekomstige update 4](#_Toc34729746)

[2. Embase.com 4](#_Toc34729747)

[2.1 Verzamelen ID’s uit Embase.com (tijdens search) 5](#_Toc34729748)

[2.2 Verzamelen ID’s uit EndNote (achteraf) 5](#_Toc34729749)

[2.3 Verwijderen PubMed records 5](#_Toc34729750)

[2.4 Verwijderen oude Embase records in toekomstige update 6](#_Toc34729751)

[3. CINAHL (EBSCO) 6](#_Toc34729752)

[3.1 Verzamelen ID’s uit CINAHL (tijdens search) 6](#_Toc34729753)

[3.2 Verzamelen ID’s uit EndNote (achteraf) 6](#_Toc34729754)

[3.3 Verwijderen PubMed records 6](#_Toc34729755)

[3.4 Verwijderen oude CINAHL records in toekomstige update 7](#_Toc34729756)

[3.5 andere EBSCO databases 7](#_Toc34729757)

[4. APA PsycInfo (EBSCO) 7](#_Toc34729758)

[4.1 Verzamelen ID’s uit PsycInfo (tijdens search) 7](#_Toc34729759)

[4.2 Verzamelen ID’s uit EndNote (achteraf) 7](#_Toc34729760)

[4.3 Verwijderen PubMed records 8](#_Toc34729761)

[4.4 Verwijderen oude PsycInfo records in toekomstige update 8](#_Toc34729762)

[4.5 andere EBSCO databases 8](#_Toc34729763)

[5. Cochrane Library (Wiley) 8](#_Toc34729764)

[5.1 Verzamelen ID’s uit Cochrane (tijdens search) 8](#_Toc34729765)

[5.2 Verzamelen ID’s uit EndNote (achteraf) 9](#_Toc34729766)

[5.3 Verwijderen PubMed / Embase records 9](#_Toc34729767)

[5.4 Verwijderen oude Cochrane records in toekomstige update 10](#_Toc34729768)

[5.5 andere EBSCO databases 10](#_Toc34729769)

[6. Scopus 10](#_Toc34729770)

[6.1 Verzamelen ID’s uit Scopus (tijdens search) 10](#_Toc34729771)

[6.2 Verzamelen ID’s uit EndNote (achteraf) 11](#_Toc34729772)

[6.3 Verwijderen PubMed records 11](#_Toc34729773)

[6.4 Verwijderen oude Scopus records in toekomstige update 11](#_Toc34729774)

[7. Web of Science 12](#_Toc34729775)

[7.1 Verzamelen ID’s uit Web of Science (tijdens search) 12](#_Toc34729776)

[7.2 Verzamelen ID’s uit EndNote (achteraf) 12](#_Toc34729777)

[7.3 Verwijderen PubMed records 12](#_Toc34729778)

[7.4 Verwijderen oude Web of Science records in toekomstige update 13](#_Toc34729779)

[DOI van alle versies 13](#_Toc34729780)

[Citeer dit document als 13](#_Toc34729781)

[Appendix: handout 13](#_Toc34729782)

# INTRODUCTION

Bij het uitgebreid zoeken van literatuur, met name voor systematische reviews, is het gebruikelijk en meestal vereist om in meerdere databases te zoeken. Afhankelijk van het specifieke onderwerp kan dit variëren van 2 tot meer dan 10 databases. De resultaten zullen daarom veel duplicaten bevatten. Ontdubbeling is noodzakelijk om te voorkomen dat hetzelfde artikel twee keer of meer wordt gescreend, gelezen en geciteerd. Maar ontdubbeling kan een tijdrovend proces zijn en is gevoelig voor fouten.

In dit document beschrijven we onze methode om te ontdubbelen:

*AMSTERDAM EFFICIENT DEDUPLICATION (AED) METHOD*

Deze methode is 100% betrouwbaar.  
Deze methode kan ook worden gebruikt bij updates van oudere searches. Dit laatste vraagt een kleine investering tijdens de initiële search maar levert bij een update een substantiële tijdsbesparing.

In deze methode wordt per database / host uitgelegd welke stappen ondernomen moeten worden.

1. Verzamelen van de accession numbers (indien mogelijk) tijdens de search. Doel: bewaren voor een latere update en in het geval van PubMed voor het ontdubbelen tussen PubMed en andere databases
2. Verzamelen van de accession numbers uit EndNote omdat de oude niet werden verzameld bij de initiële search of omdat ze niet makkelijk uit de database gehaald kunnen worden
3. Ontdubbelen van PubMed records in een andere database
4. Ontdubbelen van oude records in een database bij update

# 1. PubMed

## 1.1 Verzamelen ID’s uit PubMed (tijdens search)

* Download de lijst PMIDs in PubMed > Save > Format : **PMID**
* Bewerk deze lijst in bijvoorbeeld Notepad++ of Word

1. Voor ontdubbeling in PubMed in een toekomstige update

* Zet regeleinde om in <spatie>1
* Sla bv. op als: *PMIDs voor PubMed.txt*

1. Voor ontdubbeling in andere databases

* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie> door <spatie>OR<spatie>   
  Haal (indien nodig) de laatste OR weg
* Sla bv. op als: *PMIDs voor andere databases.txt*

## 1.2 Verzamelen ID’s uit EndNote (achteraf)

Dit gaat er vanuit dat je alle oude records uit PubMed geïmporteerd hebt in EndNote.

* Gebruik dan de Output Style “[Accession Numbers](https://zenodo.org/record/3582928/files/Accession%20Numbers.ens)”
* Selecteer alle titels [Ctrl-A] > rechter muisklik > Copy Formatted
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie>OR<spatie> door <spatie>  
  Haal (indien nodig) de laatste OR weg
* Deze versie kun je gebruiken voor het ontdubbelen in PubMed

## 1.3 Verwijderen oude PubMed records in toekomstige update

* Na het uitvoeren van een PubMed zoekactie
* Knip en plak de PMID **1.1**-lijst (zonder OR) in PubMed en voer deze zoekactie uit
* Trek die af van de nieuwe search
* #nieuwesearch NOT #oudesearch

# 2. Embase.com

## 2.1 Verzamelen ID’s uit Embase.com (tijdens search)

* Export de Embase indentification numbers (PUI) via format: “CSV-Fields by Column” en “Specify fields to be exported” > Selecteer : **PUI**
* Klik op Openen > het bestand wordt in Excel geopend
* Verwijder de eerste rij (met “PUI”)
* Selecteer alles [Ctrl-A] en kopieer [Ctrl-C]
* Ga naar Notepad++ of Word en plak als platte tekst [Ctrl-V]
* Zet regeleinde om in <spatie>1
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie> door <spatie>OR<spatie>   
  **Haal (indien nodig) de laatste OR weg**
* Deze versie kun je gebruiken voor het **ontdubbelen in Embase** toekomstige update

## 2.2 Verzamelen ID’s uit EndNote (achteraf)

Dit gaat er vanuit dat je alle oude records uit Embase geïmporteerd hebt in EndNote.

* Kopieer het veld “Notes” met Embase indentification number (PUI) naar het veld “Accession Number” (>Tools>Change/Move/Copy fields>Move Fields)
* Gebruik dan de Output Style “[Accession Numbers](https://zenodo.org/record/3582928/files/Accession%20Numbers.ens)”
* Selecteer alle titels [Ctrl-A] > rechter muisklik > Copy Formatted
* Gebruik bv. Excel of Word om deze output te scheiden of te sorteren zodat je de Embase indentification numbers (PUIs ; beginnen met L) apart in één blok / kolom hebt
* Selecteer dit blok / deze kolom
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Vervang regeleinde door <spatie>OR<spatie>
* Haal (indien nodig) de laatste OR weg
* Deze versie kun je gebruiken voor het ontdubbelen in Embase toekomstige update

## 2.3 Verwijderen PubMed records

* Na het uitvoeren van een Embase zoekactie
* Knip en plak de PMID **1.2**-lijst (met OR) in Embase tussen () met uitgang **:ui**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* #nieuwesearch NOT #oudesearch

* Of rechtstreeks:
* #x NOT (30198739 OR 30183539 OR …):ui

## 2.4 Verwijderen oude Embase records in toekomstige update

* Na het uitvoeren van een Embase zoekactie
* Knip en plak de Embase **2.1 of 2.2**-lijst (met OR) in Embase tussen () met uitgang **:id**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* #x NOT (L30198739 OR L30183539):id

# 3. CINAHL (EBSCO)

## 3.1 Verzamelen ID’s uit CINAHL (tijdens search)

* NIET AAN DE ORDE! Te omslachtig door het maximum van 50 item per keer “Add to folder”

## 3.2 Verzamelen ID’s uit EndNote (achteraf)

Dit gaat er vanuit dat je alle oude records uit CINAHL geïmporteerd hebt in EndNote.

* Gebruik de Output Style “[Accession Numbers](https://zenodo.org/record/3582928/files/Accession%20Numbers.ens)”
* Selecteer alle titels > rechter muisklik > Copy Formatted
* Gebruik bv. Excel om deze output te scheiden of te sorteren zodat je de CINAHL Accession Numbers apart in één blok in één kolom hebt
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Deze versie kun je gebruiken voor het **ontdubbelen in CINAHL** toekomstige update

## 3.3 Verwijderen PubMed records

* Na het uitvoeren van een CINAHL zoekactie
* Ga naar Notepad++ of Word en plak de PMID **1.2**-lijst (met OR) (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie>OR<spatie> door <spatie>OR<spatie>NLM. Voeg nog NLM toe aan het begin
* Knip en plak de bewerkte lijst in CINAHL tussen () met voorloper **PM**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* #nieuwesearch NOT #oudesearch
* Of rechtstreeks:
* Sx NOT PM (NLM30198739 OR NLM30183539 OR …)

## 3.4 Verwijderen oude CINAHL records in toekomstige update

* Na het uitvoeren van een CINAHL zoekactie
* Knip en plak de CINAHL **3.2**-lijst (met OR) in CINAHL tussen () met voorloper **AN**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* #nieuwesearch NOT #oudesearch
* Of rechtstreeks:
* Sx NOT AN (103870552 OR 123658192 OR …)

## 3.5 andere EBSCO databases

* Voor andere EBSCO databases (niet CINAHL/PsycInfo), zoals ERIC en SPORTDiscus zijn bovengenoemde stappen 2 en 4 bruikbaar. De ontdubbeling met PubMed is niet mogelijk omdat Eruic en SPORTDisccus geen PMID’s bevatten.

# 4. APA PsycInfo (EBSCO)

## 4.1 Verzamelen ID’s uit PsycInfo (tijdens search)

* NIET AAN DE ORDE! Te omslachtig door het maximum van 50 item per keer “Add to folder”.

## 4.2 Verzamelen ID’s uit EndNote (achteraf)

Dit gaat er vanuit dat je alle oude records uit PsycInfo geïmporteerd hebt in EndNote.

* Gebruik de Output Style “[Accession Numbers](https://zenodo.org/record/3582928/files/Accession%20Numbers.ens)”
* Selecteer alle titels > rechter muisklik > Copy Formatted
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Deze versie kun je gebruiken voor het **ontdubbelen in PsycInfo** toekomstige update

## 4.3 Verwijderen PubMed records

* Na het uitvoeren van een PsycInfo zoekactie
* Knip en plak de PMID 1.2-lijst (met OR) in PsycInfo tussen () met voorloper **PM**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* Sx NOT PM (30198739 OR 30183539 OR …)

## 4.4 Verwijderen oude PsycInfo records in toekomstige update

* Na het uitvoeren van een PsycInfo zoekactie
* Knip en plak de PsycInfo 4.2-lijst (met OR) in CI PsycInfo tussen () met voorloper **AN**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* Sx NOT AN (2006-09996-006 OR 2013-37963-006 OR …)

## 4.5 andere EBSCO databases

* Voor andere EBSCO databases (niet CINAHL/PsycInfo), zoals ERIC en SPORTDiscus zijn bovengenoemde stappen 2 en 4 bruikbaar. De ontdubbeling met PubMed is niet mogelijk omdat Eric en SPORTDiscus geen PMID’s bevatten.

# 5. Cochrane Library (Wiley)

## 5.1 Verzamelen ID’s uit Cochrane (tijdens search)

* HET IS NIET MOGELIJK UITSLUITEND DE ACCESSION NUMBERS TE DOWNLOADEN

## 5.2 Verzamelen ID’s uit EndNote (achteraf)

Dit gaat er vanuit dat je alle oude records uit Cochrane geïmporteerd hebt in EndNote.

* Gebruik de Output Style “[Accession Numbers](https://zenodo.org/record/3582928/files/Accession%20Numbers.ens)”
* Selecteer alle titels > rechter muisklik > Copy Formatted
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Deze versie kun je gebruiken voor het ontdubbelen in Cochrane toekomstige update

## 5.3 Verwijderen PubMed / Embase records

DIT WERKT NIET BIJ GROTE AANTALLEN IN PUBMED / EMBASE

**Stappen PMID-lijst:**

* Knip en plak de PMID **1.2**-lijst (met OR) in bijvoorbeeld Notepad++ of Word:
* Zet in de PMID **1.2**-lijst voor het eerste nummer OR<spatie> en na het laatste nummer <spatie>OR
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H] vervang OR<spatie> door OR (“**PubMed** <spatie>
* Vervang <spatie>OR door “)**:an** <spatie>OR
* Haal de eerste en de laatste OR in het blok weg
* Voer hiermee een zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* #x NOT (("PubMed 30403654"):an OR ("PubMed 30322793"):an OR ("PubMed 30247967"):an)

**Stappen Embase-lijst:**

* Knip en plak de Embase **2.1 of 2.2**-lijst (met OR) in bijvoorbeeld Notepad++ of Word:
* Zet in de Embase-lijst voor het eerste nummer OR<spatie> en na het laatste nummer <spatie>OR
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H] OR<spatie> door OR (“**Embase** <spatie>
* Vervang <spatie>OR door “)**:an** <spatie>OR
* Haal de eerste en de laatste OR in het blok weg
* Voer hiermee een zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* #x NOT (("Embase 624095871"):an OR ("Embase 615604185"):an OR ("Embase 615669641"):an)

## 5.4 Verwijderen oude Cochrane records in toekomstige update

* Na het uitvoeren van een Cochrane zoekactie
* Knip en plak de Cochrane **5.2**-lijst met OR in Cochrane tussen ()   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* #x NOT (CN-01577879 OR …)

## 5.5 andere EBSCO databases

* Voor andere EBSCO databases (niet CINAHL/PsycInfo), zoals ERIC en SPORTDiscus zijn bovengenoemde stappen 2 en 4 bruikbaar. De ontdubbeling met PubMed is niet mogelijk omdat Eric en SPORTDiscus geen PMID’s bevatten.

# 6. Scopus

## 6.1 Verzamelen ID’s uit Scopus (tijdens search)

* Selecteer: All >>> Export > method of export: CSV >Selecteer EID (andere vakjes leeg laten!)
* Open de CSV-file in Excel en verwijder de bovenste regel
* Gebruik Excel om deze output te scheiden zodat je de Scopus EIDs (in het formaat 2-s2.0-77954009273) apart in één kolom hebt
* Selecteer deze kolom
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie> door <spatie>OR<spatie>
* Deze versie kun je gebruiken voor het ontdubbelen in Scopus toekomstige update

NB. Bij meer dan 2000 items opsplitsen in delen!

## 6.2 Verzamelen ID’s uit EndNote (achteraf)

Dit gaat er vanuit dat je alle oude records uit Scopus geïmporteerd hebt in EndNote.

* Kopieer het veld “URL” naar het veld “Accession Number” (>Tools>Change/Move/Copy fields>Move Fields)
* Gebruik de Output Style “[Accession Numbers](https://zenodo.org/record/3582928/files/Accession%20Numbers.ens)”
* Selecteer alle titels > rechter muisklik > Copy Formatted
* Plak de file in Excel om deze output te scheiden zodat je de Scopus EIDs (in het formaat 2-s2.0-77954009273) apart in één kolom hebt
* Selecteer deze kolom
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie> door <spatie>OR<spatie>
* Deze versie kun je gebruiken voor het ontdubbelen in Scopus toekomstige update

## 6.3 Verwijderen PubMed records

* Na het uitvoeren van een Scopus zoekactie
* Knip en plak de PMID **1.2-**lijst (met OR) in Scopus tussen () met voorloper **PMID**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* PMID(30198739 OR 30183539 OR …)

NB. Werkt alleen met combinatie van set nummers (#1 AND NOT #2)!

## 6.4 Verwijderen oude Scopus records in toekomstige update

* Na het uitvoeren van een Scopus zoekactie
* Knip en plak de Scopus **6.1**-lijst (met OR) in Scopus tussen () met voorloper **EID**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* EID( 2-s2.0-85064602041 OR 2-s2.0-85064120601 OR … )

NB. Werkt alleen met combinatie van set nummers (#1 AND NOT #2)!

# 7. Web of Science

## 7.1 Verzamelen ID’s uit Web of Science (tijdens search)

* Export to Other File Formats > Full Record > Tab-delimited (Win)
* Open de file in Excel
* Selecteer de kolom UT met de accession numbers
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie> door <spatie>OR<spatie>
* Deze versie kun je gebruiken voor het **ontdubbelen in Web of Science** toekomstige update

## 7.2 Verzamelen ID’s uit EndNote (achteraf)

Dit gaat er vanuit dat je alle oude records uit Web of Science geïmporteerd hebt in EndNote.

* Gebruik de Output Style “[Accession Numbers](https://zenodo.org/record/3582928/files/Accession%20Numbers.ens)”
* Selecteer alle titels > rechter muisklik > Copy Formatted
* Ga naar Notepad++ of Word en plak (als platte tekst [Ctrl-V])
* Zet regeleinde om in <spatie>1
* Selecteer alles [Ctrl-A] > Zoeken/Vervangen [Ctrl-H]: vervang <spatie> door <spatie>OR<spatie>
* Deze versie kun je gebruiken voor het ontdubbelen in Web of Science toekomstige update

## 7.3 Verwijderen PubMed records

* Na het uitvoeren van een Web of Science zoekactie
* Knip en plak de PMID **1.2**-lijst (met OR) Web of Science tussen () met voorloper **PMID=**   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* #x NOT PMID=(30198739 OR 30183539 OR …)

## 7.4 Verwijderen oude Web of Science records in toekomstige update

* Na het uitvoeren van een Web of Science zoekactie
* Knip en plak de Web of Science **7.1 of 7.2**-lijst (met OR) in Web of Science tussen () met voorloper UT=   
  en voer deze zoekactie uit
* Trek die af van de nieuwe search
* Of rechtstreeks:
* #x NOT UT=(WOS:000436466400018 OR WOS:00043646654321 OR …)

# DOI van alle versies

10.5281/zenodo.3582927

# Citeer dit document als

René Otten, Ralph de Vries, & Linda Schoonmade. Amsterdam Efficient Deduplication (AED) method - handleiding. Zenodo.org . http://doi.org/10.5281/zenodo.3582927

# Appendix: handout

Een handout (in het Engels) is te vinden op

<http://doi.org/10.5281/zenodo.3582927>

Maart 2020

Zet regeleinde om in <spatie>

In Notepad++: Ga naar Bewerken > Uitlijnen > Regeluiteinden omzetten in spaties  
In Word: Ctrl-H > Zoeken naar: ^p > Vervangen door <spatie>