Contents

[Step by step: 1](#_Toc513728611)

[Important notes and decisions: 2](#_Toc513728612)

[Import what seems to be duplicated publications, but it is not. 2](#_Toc513728613)

[Search on ePrint database 2](#_Toc513728614)

# Step by step:

1. **Create, select or delete a project**

* Add a project name, description, and the scores number correspondent to A\*, A, B and C.
* Project names must be unique
* You must always have at least one project created. If they are all deleted the first one must be created from the DB.

1. **Upload XLSX template**

* Is it worth having three options?
  + Current employee
  + Ex MDX employee
  + External author
* Import XLSX file based on the template
* Click *upload* to get XLSX file data

1. **Checks if current employee**:

* If: the value in the "Current Employee" column from the XLSX is Y then is current employee.
* Else if: email ends with @mdx.ac.uk it sets as current employee
* Else: not a current employee

Show the following columns (not reaching ePrint yet, just the local database for previously imported information):

* Imported name
* Imported email
* Current calculated employee status
* If previously imported and processed:
  + Number of publications as main author for the REF year (>= 01/01/2014)
  + Number of publications as co-author (>= 01/01/2014)
* Tick box to “ignore user”. All his papers will be ignored. Author will not be imported to the DB.
* Click to process current list

1. **Processing each author at time.**

* Only *counts* the authors publications at this point. It is not getting their details yet.
* Maybe slower. To be tested.
* Able to show progress feedback to the user. Shows a tick next to each processed name.
* If author (first or not) are **not** a current employee:
  + It will be moved to the end of the processing list
  + It will have the “Ignore user” tick box ticked by default
  + If not ignored, it will be processed just like a current employee.

Show the following columns:

* Author name
* Processed (tick after imported)
* Current calculated employee status (drop down so it can be changed if wrongly calculated).
* Number of publications as main author for the REF year (>= 01/01/2014)
* Number of publications as co-author (>= 01/01/2014)
* *Even if there are* ***no*** *valid publications, author details will still be saved to the BD showing 0 processed publications.*
* Tick box to “ignore user”. All his papers will be ignored. Author will not be imported to the DB.
* ~~Drop down with REF options~~. It will assign all valid publications where author is the first author to a chosen REF.
* ~~ERA voting~~. Same situation as REF assignment above.
* Click to process current list (this will import all non-ignored authors + their publications).

1. **Processing list:**

* If combination of author + publication is already in the DB:
  + Check if is assigning any REF to it
    - If REF is currently null, updates it
  + Check if is assigning any ERA to it
    - If ERA is currently null, updates it
* Else:
* Check

# Important notes and decisions:

## Import what seems to be duplicated publications, but it is not.

It can happen because of automated process called SWORD (example: [1](http://eprints.mdx.ac.uk/23121/) and [2](http://eprints.mdx.ac.uk/23224/))

It can be because author added a new version with publisher formatting (example: [1](http://eprints.mdx.ac.uk/15311/) and [2](http://eprints.mdx.ac.uk/19480/))

Conclusion by Balbir: don’t touch them (it will be defined on case by case) – Face to face meeting on 17/04/2018

## Search on ePrint database

Is only done with First Name + Last Name from the spreadsheet or DB (if previously added)

Emails are only used to check if staff is current employee, ex-mdx, or external