

OpenHarvester

A java prototype that processes the result set of pre-downloaded data (from a database) and allows one to claim his/her publications from a ranked list.

Muhammad Javed, Ph.D.

Ontology Engineer

Tech Lead (Scholars@Cornell)

Cornell University Library

It's a preliminary work



WORK ZONE AHEAD
Next 15 Miles Mins

Work is inspired from Symplectic Element's processes.
and has a high overlap with ReCiter

Name-based search terms

Name variants: +
(last name first)

Javed, M. ✖

Use the format "Smith, A. G.". This format clearly distinguishes your last name and each of your initials.

Add all combinations of your last name and initials under which you publish. The system will use the exact formats best required when searching each data source. Do not use the format "Smith, A. G." write your last name first, followed by a comma.

If a name variant is only used by one online database, you can customise this setting in the database settings.

Addresses:

Dublin City University
Cornell University
Dublin City
Cornell

[Switch to advanced view](#)

Add keywords associated with previous institutions, or those with which you collaborate.

Start date:

This date should be in one of the following formats:
14 Mar 2001 (dd MMM yyyy)
Mar 2001 (MMM yyyy)
1991 (yyyy)

Keywords: +

You have not supplied any keywords.

[Switch to advanced view](#)

Journals: +

You have not supplied any journals.

[Switch to advanced view](#)

Specific article IDs

Article ID: +

Online database:

If searches are not picking up specific articles from a database, you can add their IDs here. For example, enter the Article ID "28399581" with the online database "PubMed" selected to enqueue retrieval of the PubMed article with PubMedID 28399581.

You have not supplied any article IDs.

Home > Menu > Research Admin > Module Admin > System Admin

Elements

Publication search settings

Three reasons why I did this preliminary work:

Reason 1: Symplectic Elements do not “search” publications, neither in CrossRef nor in Epubmed Central. (Supplementary Sources)

Reason 2: Adrenaline in my veins to explore what data can be harvested using open citation APIs.

Reason 2.1: Now I can access the data, can I harvest publications for a researcher ?

Reason 3: To understand, what data is required to successfully find publications for a researcher ?

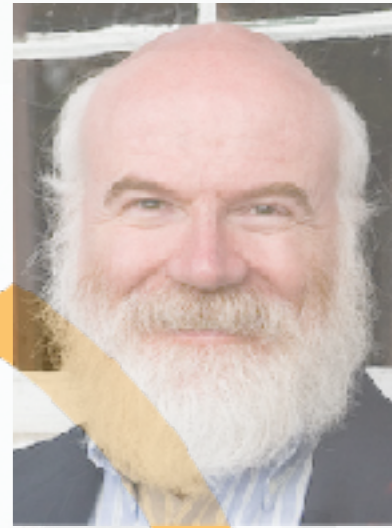


Limitations:

1. Works on pre-downloaded data.
 1. *Step I: Search database and download result set.*
 2. *Step II: Process result set and harvest researcher's publications.*
2. No name diacritic handling.
3. Name handling requires some more tweaks.
4. and more..



...



• Dean Krafft

Search name in the author list (Dean OR Krafft)

2

3

Result Set

5

List Review from top

Ranking

4

Claim a Publication

6

(String)

Set of Publications

Ranked list of Publications

List of my Publications

1

Re-Ranking (based on update profile)

8

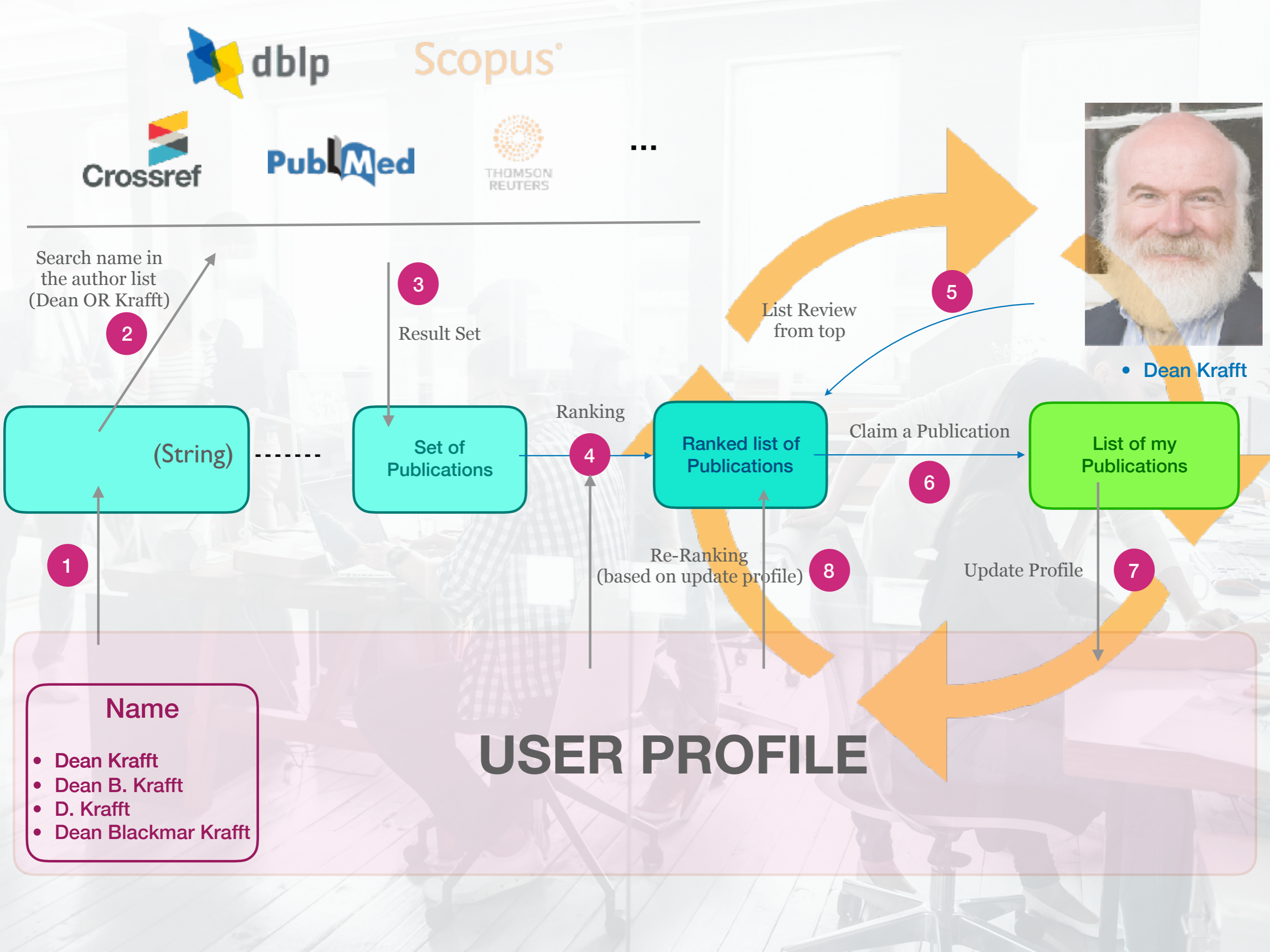
Update Profile

7

Name

- Dean Krafft
- Dean B. Krafft
- D. Krafft
- Dean Blackmar Krafft

USER PROFILE



User Profile:

1. Name Variants
2. Start/End Year
3. Affiliations
4. Co-authors
5. Subject Areas
6. Identifiers (personal)
7. Identifiers (publications)
8. and more...

Two Step Process

Step 1: Download Data

search string *output base-folder*
↓ ↓

```
java -jar CrossrefDataDownloader.jar Dean+Krafft CROSSREF
```

Downloads result set and save files in a folder named as “Dean+Krafft”

Step 2: Process Result Set

Processes result set files from folder “Dean+Krafft”



LIVE DEMO



Roberts, Patricia
| Professor and Chair

Search Publications

Positions

► Professor, [English](#), [College of Arts and Humanities](#)
1997 -



My research is focused on Derrida and the nature of political discourse in the era of electracy.

Research Areas

[Derrida](#) | [Electracy](#) | [Political discourse](#) | [Rhetoric](#)

Publications in VIVO



2 in the last 10 full years

[Co-author Network](#)

[Map of Science](#)

Affiliation | Publications | Teaching | Service | Background | Contact | View All

head of

[English Chair](#) 2013 -

member of

[University Undergraduate Curriculum Committee Chair](#) 2015 -

[College of Arts and Humanities Tenure Committee Member](#) 2001 - 2003

OpenHarvester

User Profile

Search Publications

Load Claimed Publications to VIVO