

## Dataset Metadata

### 1. General Information

- **Title:** AoI2WoS: Mapping Area of Interest in Google Scholar Profile to Web of Science (WoS) categories
- **Description:** Google Scholar Profiles (GSP) provide a very efficient way for researchers to showcase their academic achievements, publications, and research interests. It is possible to check who is citing articles, graph citations over time, and report several citation metrics that can be used in Scientometrics. One important section in each GSP is called "Areas of interest". It can include maximum five scientific fields for everyone. Because, these fields are filled manually by the authors found on their own perception and knowledge about the broad scientific disciplines, they are not necessarily matched to Web of Science (WoS) categories of sciences including 7 Domains and 134 Areas. As a result, we have gathered and created the dataset that have classified nonstandard and divers scientific fields available in 85285 records of GSP in to 7 Domains and 134 Areas according to WoS categories.

### 2. Content

- **Number of Records:** 85285 records
- **Format:** CSV
- **Size:** 9.49 MB
- **Variables/Columns:**
  - Fields (Google scholar profile information: Areas of interest including maximum 5 scientific fields in the Google scholar profile of each researcher)
  - Domain (different fields of science contain 7 domains according to WoS categories)
  - Area (subdomains for domains contain 134 according to WoS categories)

### 3. Data Structure

- **Input Variables:**
  - *Area of Interest* of each author available in their Google Scholar Profile
- **Target Variable:**
  - Output:
    - *Domain*
    - *Area*

### 4. Data Quality

- **Data Cleaning Steps:**
  - 1-remove noise
  - 2-remove punctuations
  - 3-remove stop words
  - 4-stemming

### 5. Acknowledgments

- The dataset compilation was automated using AI tools.

## **6. Contact Information**

- **Contact Person:**

- Mahdi Bahaghighat, Email: [Bahaghighat@eng.ikiu.ac.ir](mailto:Bahaghighat@eng.ikiu.ac.ir)
- Pouyan Jahani rad, Email: [pouyan.jahani@gmail.com](mailto:pouyan.jahani@gmail.com)

- **Organization**

- Computer Engineering Department, Imam Khomeini International University, Qazvin, Iran