**Data Bibliography Science Articles – DBSA**

Mục lục

[1. Các module chính của chương trình 2](#_Toc273301309)

[1.1. Thu thập các bài báo khoa học: 2](#_Toc273301310)

[1.2. Import dữ liệu từ cơ sở dữ liệu DBLP 2](#_Toc273301311)

[1.3. Rút trích thông tin chỉ mục của các bài báo. 2](#_Toc273301312)

[1.4. Phân loại các bài báo theo chủ đề 2](#_Toc273301313)

[2. Kiến trúc của chương trình 3](#_Toc273301314)

[2.1. Core 3](#_Toc273301315)

[2.2. Kiến trúc phân lớp của chương trình: 3](#_Toc273301316)

# Các module chính của chương trình

## Thu thập các bài báo khoa học:

Thu thập các bài báo khoa học từ thư viện số ACM[[1]](#footnote-2), IEEExplore[[2]](#footnote-3) và CiteSeer[[3]](#footnote-4), ScienceDriect [[4]](#footnote-5)

## Import dữ liệu từ cơ sở dữ liệu DBLP[[5]](#footnote-6)

Module sẽ lấy thông tin từ dữ liệu có sẵn của DBLP từ đó xác định thêm chủ đề của bài báo bổ xung vào cơ sở dữ liệu.

## Rút trích thông tin chỉ mục của các bài báo.

Rút trích các thông tin chỉ mục của các bài báo từ kết quả trả về của module thu thập.

+Tên bài báo (title).

+ Tác giả bài báo.

+ Tóm tắt của bài báo (abstract).

+ Năm công bố của bài báo.

+ Nhà xuất bản (Publisher).

## Phân loại các bài báo theo chủ đề

Dựa vào tên hoặc tóm tắt của bài báo thu thập được, module sẽ phân loại các bài báo theo các chủ đề trong khung phân loại ngành khoa học máy tính.

# Kiến trúc của chương trình



## Core

* + 1. Digitals : Các thư viện số nơi lấy thông tin các bài báo.
    2. Query: Xử lý các câu truy vấn xuống database.
    3. Classificaiton: phân lớp bài báo khoa học.
    4. Harvest: Lấy thông tin chỉ mục các bài báo từ kết quả trả của thư viện số.

## Kiến trúc phân lớp của chương trình:

* GUI :

uit.tkorg.dbsa.gui.\*

.fetcher

.classification

* Action :

uit.tkorg.dbsa.actions.\*

.fetcher

.classification

* Core :

uit.tkorg.dbsa.core.\*

.fetcher

.classification

* Utilities :

uit.tkorg.dbsa.utilies.\*

.fetcher

.classification

* Documentation :

uit.tkorg.dbsa.documentation.\*

.doc

.references

.presentations

…

* Resources :

uit.tkorg.dbsa.resources.\*

.images

.icon

…

* Properties :

Uuit.tkorg.dbsa.properties.\*

.file

…

1. <http://portal.acm.org/portal.cfm> [↑](#footnote-ref-2)
2. <http://ieeexplore.ieee.org> [↑](#footnote-ref-3)
3. <http://citeseerx.ist.psu.edu/> [↑](#footnote-ref-4)
4. <http://www.sciencedirect.com>/ [↑](#footnote-ref-5)
5. <http://dblp.mpi-inf.mpg.de/dblp-mirror/index.php> [↑](#footnote-ref-6)