Design Document

Download — Write → *facultyInfo* (.html files)

Extract → Call → Parser

Read ← *facultyInfo*

Extract — Write → *Tables*

Parser — Write → *Tables* (.csv files)

Tables — Read → Database — Write → **Database** *facultyInfo*

Legend:
- Modules written in python and C
- Directories for storing files
- Database
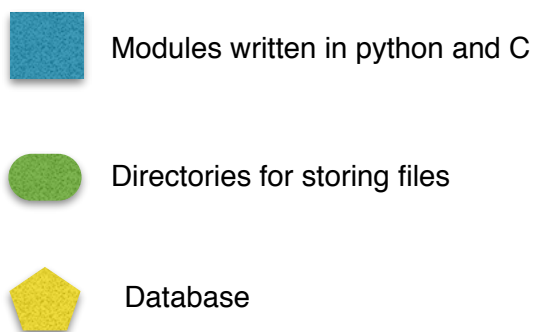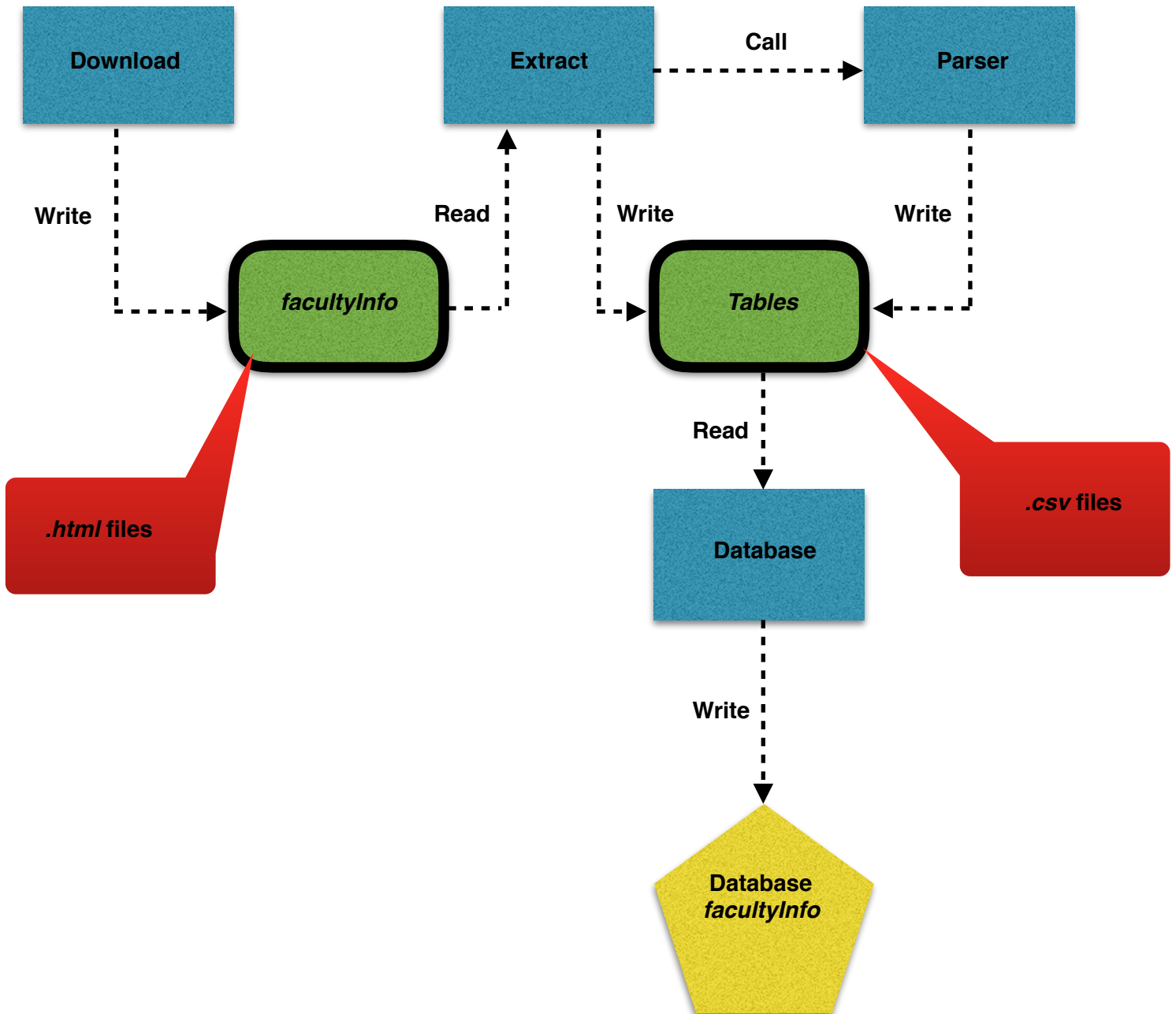
Module Description

1. **Download**

    a. This module is written in python.

    b. It Downloads the faculty HTML pages from the web and stores the .html files in the 'facultyInfo' directory.

2. **Parser**

    a. This module is written in C.

    b. It extracts the relevant information from .html files and writes it into .csv files that are stored in 'Tables' directory.

3. **Extract**

    a. This module is written in C.

    b. It reads the .html files from 'facultyInfo' directory and calls parser program on each of these files.

4. **Database**

    a. This module is written in python.

    b. It reads the .csv files from the 'Tables' directory and stores it in the database named as 'facultyInfo'.