

# studyExtraction

## Installation

```
pip install -r requirements.txt
python setup.py install
```

## Introduction

To use with the version 2016.2 of Vitic, you'll need to have the database loaded into the Oracle DB server and run the script from the server where Oracle is installed. Also, you'll need to set up the ORACLE\_HOME and LD\_LIBRARY\_PATH environment variables.

## Manual

Extract studies' findings based on the given filtering and the organs' and morphological changes' ontologies-based expansions of these findings.

- Required arguments:
  - -a / --organ ORGAN Anatomical entity that the finding refers to. You can filter for more than one organ by passing a blank space-separated list.
- Optional arguments:
  - Version-related arguments:
    - -v / --version {2016.1,2016.2} Vitic database version (default: 2016.2).
    - -d / --sid SID If working with Vitic database version 2016.2, provide the Oracle SID's. -u / --user USER If working with Vitic database version 2016.2, provide the Oracle database user name.
    - -p / --passw PASSW If working with Vitic database version 2016.2, provide the Oracle database password.
  - Study design-related arguments:
    - -i / --min\_exposure MIN\_EXPOSURE Minimum exposure period (days).
    - -e / --max\_exposure MAX\_EXPOSURE Maximum exposure period (days).
    - -r / --route {Cutaneous, Diertary, Oral, Oral gavage, Intragastric, Nasogastric, Oropharyngeal, Endotracheal, Intra-articular, Intradermal, Intraesophageal, Intraileal, Intramuscular, Subcutaneous, Intraocular, Intraperitoneal, Intrathecal, Intrauterine, Intravenous, Intravenous bolus, Intravenous drip, Parenteral, Nasal, Respiratory (inhalation), Percutaneous, Rectal, Vaginal, Subarachnoid} Administration route. You can filter for more than one administration route by passing a blank space-separated list.
    - -s / --species {Mouse, Rat, Hamster, Guinea pig, Rabbit, Dog, Pig, Marmoset, Monkey, Baboon} Species. You can filter for more than one species by passing a blank space-separated list.
    - -x / --sex {F,M,Both} Study design sex.
  - Finding-related arguments:
    - -m / --observation OBSERVATION Morphological change type that the finding refers to. You

can filter for more than one morphological change by passing a blank space-separated list.

- `-t / --treatment_related` Keep only treatment-related findings.
- Output-related arguments:
- `-o / --output_basename OUTPUT_BASENAME` Output file base name. Two output files will be generated: `basename_quant.tsv` and `basename_qual.tsv`, with quantitative and qualitative results respectively. (default: output).

## Use examples

### 1. Extract all studies with liver-related findings

- vitic 2016.1:

```
python extract.py -v 2016.1 -a liver
```

- vitic 2016.2:

```
python extract.py -v 2016.2 -d ORACLE_SID -u ORACLE_USER -p ORACLE_PASSWORD -a liver
```

### 2. Extract all studies with liver- and kidney-related findings

Note that you can filter for more than one organ by passing a blank space-separated list.

- \* vitic 2016.1:

```
python extract.py -v 2016.1 -a liver kidney
```

- \* vitic 2016.2:

```
python extract.py -v 2016.2 -d ORACLE_SID -u ORACLE_USER -p ORACLE_PASSWORD -a liver  
kidney
```

### 3. Extract only studies of interest

Filter the studies of interest based on exposure time (days), administration route, and species. Note that for route and species you can filter for more than one value by passing a blank space-separated list.

- \* Using long arguments:

```
python extract.py -v 2016.1 --organ liver --min_exposure 1 --max_exposure 10 --route  
ORAL --species MOUSE RAT
```

- \* Using short arguments:

```
python extract.py -v 2016.1 -a liver -i 1 -e 10 -r ORAL -s MOUSE RAT
```

### 4. Extract treatment-related findings only

```
python extract.py -v 2016.1 -a liver -i 1 -e 10 -r ORAL -s MOUSE RAT -t
```