

Writing scripts via Pywikibot

Run as part of the “Small wiki toolkits” project



WIKIMEDIA
FOUNDATION

Upcoming SWT workshops on bots development

- How to run a popular bot on your wiki (as part of the Wikimedia Hackathon)
- How to host a bot on Toolforge
- How to maintain bots
- Intro to Wikimedia Tools (PAWS, Quarry, WDQS) for running queries and scripts
- How to interact with Wikidata via Pywikibot

Agenda

- Writing custom scripts
- Pywikibot project structure
- Key classes, methods and properties
- Implementing a search and replace script



Writing custom scripts



WIKIMEDIA
FOUNDATION

Pywikibot framework

- Writing custom scripts
- Pywikibot project structure (userscript directory)
- Installing development packages
- Pywikibot Python classes, methods & properties
- Exploring the structure of a script
- Implementing Search & Replace script
- Exploring Archivebot script

Poll question
How many SWT workshops
have you attended
previously?



WIKIMEDIA
FOUNDATION

Pywikibot project structure



WIKIMEDIA
FOUNDATION

Poll question
Do you have Pywikibot
installation on your local
machine?



WIKIMEDIA
FOUNDATION

Pywikibot framework

Project structure

```
~~~~ project tree structure ~~~~
```

```
|— CODE_OF_CONDUCT.rst
|— dev-requirements.txt
|— docs
|   |— Makefile
|— pwb.py
|— pywikibot
|— scripts
|   |— userscripts
|       |— __init__.py
|— tests
|   |— README.rst
~~~~~ truncated ~~~~~
```

Pywikibot framework

Installing development packages

→ Installing 'dev_requirements.txt' packages

```
(venv)core:$ pip install -r dev_requirements.txt
```

```
~~~~~ output ~~~~~
```

```
Ignoring pytest-cov: markers 'python_version < "3.6"' don't match your environment
```

```
Ignoring pytest-cov: markers 'python_version >= "3.10"' don't match your environment
```

```
Ignoring cffi: markers 'python_version == "3.5"' don't match your environment
```

```
Collecting pytest>=5.3
```

```
~~~~~ truncated ~~~~~
```

Key classes, methods and properties



WIKIMEDIA
FOUNDATION

Poll question

How familiar are you with the concepts of classes, methods and properties in Python?



WIKIMEDIA
FOUNDATION

Poll question

How familiar are you with
Pywikibot classes, methods
and properties?



WIKIMEDIA
FOUNDATION

Pywikibot framework

Quick overview of OOP in Python

→ Classes, methods and properties

```
class MyClass():  
    # ...  
  
class MyClass(BaseClass): # inheritance  
    # ...
```

Pywikibot framework

Quick overview of OOP in Python

→ Methods

```
class MyClass():
    def __init__(self): # constructor, a special method
        # ...
    def operate(self):
        #....
```

Pywikibot framework

Quick overview of OOP in Python

→ Properties

```
class MyClass():
    attr1 = 1                # class attribute
    def __init__(self, arg1):
        self.arg1 = arg1    # instance attribute/property
        # ...
    # other methods here
```


Pywikibot framework

Some Key Pywikibot classes, methods & properties

→ Classes

`CurrentPageBot`: Sets the current page to be treated.

`ExistingPageBot`: Treats only existing pages.*

`CreatingPageBot`: A class that only treats nonexistent pages.

`Page`: Represents a Mediawiki page.

Pywikibot framework

Some Key Pywikibot classes, methods & properties

→ Methods

`treat_page`: process a page. This is inherited and overwritten in custom scripts.

`put_current`: Writes to current page. Also inherited

Pywikibot framework

Some Key Pywikibot classes, methods & properties

→ Properties

`current_page`: The current active wiki page. Give you access to the text, title etc of a page

`opt`: An attribute that holds command line arguments passed to script.

`update_options`: Stores the *valid* flags for the script

Exploring the basic structure of a script



WIKIMEDIA
FOUNDATION

Implementing a search and replace script



WIKIMEDIA
FOUNDATION

Pywikibot framework

Writing a search and replace script

→ Classes, methods, and properties

ExistingPageBot class

treat_page() method

put_current() method

opt property

current_page property

update_options property

Pywikibot framework

Writing a search and replace script

→ Script invocation

```
(venv)core:$ python pwb.py search_replace -page:"User:Almoksey" -tosearch:"Token"  
-toreplace:"Komla" -summary:"my custom summary info"
```

```
1 pages read
```

```
1 pages written
```

```
0 pages skipped
```

```
Execution time: 12 seconds
```

```
.....
```

Pywikibot framework

Writing a search and replace script

→ Class implementation

```
class SearchReplace(ExistingPageBot):  
    update_options = {  
        'summary': 'Bot: search and replace operation',  
        'tosearch': "",  
        'toreplace': ""  
    }  
  
    def treat_page(self):  
        page_text = self.current_page.text # retrieve page content at text  
        text = page_text.replace(self.opt.tosearch, self.opt.toreplace)  
        self.put_current(text, summary=self.opt.summary) # save back to modified text/content
```


Try-out time

- Set up a custom script in your project, give it your name and import 'pywikibot' and 'ExistingPageBot'.
- Add a class that inherits from 'ExistingPageBot'



WIKIMEDIA
FOUNDATION

Pywikibot framework

Writing a search and replace script

→ main function invocation pattern

```
def main():  
    options = {}  
  
    gen_factory = pagegenerators.GeneratorFactory()  
  
    args = pywikibot.handle_args() # global options  
  
    local_args = gen_factory.handle_args(args) # generators options  
  
    for arg in local_args:  
        opt, sep, value = arg.partition(':')  
  
        if opt in ('-tosearch', '-toreplace'):  
            options[opt[1:]] = value  
  
    bot = SearchReplace(generator=gen_factory.getCombinedGenerator(), **options).run()  
  
if __name__ == '__main__':  
    main()
```

Pywikibot framework

Testing the script

→ Searching and replacing a work on *test* wiki page

```
(venv)core:$ python pwb.py search_replace -page:"User:Almoksey"  
-tosearch:"Python" -toreplace:"Javascript" -summary:"Bot: my custom summary  
info"
```

```
(venv)core:$ python pwb.py -simulate search_replace -page:"User:Almoksey"  
-tosearch:"Python" -toreplace:"Javascript" -summary:"Bot: my custom summary  
info"
```

Exploring archivebot script



WIKIMEDIA
FOUNDATION

Try-out time

- Extend the previous example to receive command line arguments and print them on standard output(console)



WIKIMEDIA
FOUNDATION

THANK YOU

Seyram Komla Sapaty

✉ ssapaty@wikimedia.org

🐦 [@komlasapaty](https://twitter.com/komlasapaty)



WIKIMEDIA
FOUNDATION



Copyright © 2022, [Seyram Komla Sapaty](#), [Srishti Sethi](#) and the [Wikimedia Foundation](#).
This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](#).