

bots

installing, configuring and running pywikibot

Merlijn van Deen, Maarten Dammers
user:valhallasw / user:multichill

1. Download packages (might take a while)
2. Bot background
3. Creating a bot user account
4. Install, configure
5. Make some edits on wikidata!

http://www.mediawiki.org/wiki/Pywikipediabot/Hackathon_2013

Pywikipediabot/Hackathon 2013

< [Pywikipediabot](#)

- For windows users: download python - <http://www.python.org/ftp/python/2.7.5/python-2.7.5.msi> 
- Download pywikipedia: <http://toolserver.org/~pywikipedia/pywikipedia-rewrite.zip> 
- Create a bot user account: <https://www.mediawiki.org/w/index.php?title=Special:UserLogin&type=signup> 





bot workflow

1. get a list of pages
2. do something with each page on your list

1. Download packages (might take a while)
2. Bot background
- 3. Creating a bot user account**
4. Install, configure
5. Make some edits on wikidata!



Special page

Log in / create account

Create account

Already have an account? [Log in.](#)

Language: [Deutsch](#) | [English](#) | [Esperanto](#) | [français](#) | [español](#) | [italiano](#) | [Nederlands](#)

Username:

Use a temporary random password and send it to the email address specified below

Password:

Retype password:

Email:

- E-mail (optional): Enables others to contact you through your user or user_talk page without needing to reveal your identity.

You can also choose to let others contact you by email through a link on your user or talk page. Your email address is not revealed when other users contact you.

Reason:

Remember my login on this browser (for a maximum of 30 days)

Please read [what this site is and is not about](#) before going on.

- [Main page](#)
- [Get MediaWiki](#)
- [Get extensions](#)
- [Communication](#)
- [Tech blog](#)

- Support
 - [User help](#)
 - [Technical manual](#)
 - [FAQ](#)
 - [Support desk](#)

[MediaWiki.org](#)

[Development](#)

[Toolbox](#)

1. Download packages (might take a while)
2. Bot background
3. Creating a bot user account
- 4. Install, configure**
5. Make some edits on wikidata!

Organiseren Openen Delen met Nieuwe map

- .svn Bestandsmap
- pywikibot Bestandsmap
- distribute_setup.py Python File 14,9 kB
- INSTALL Bestand 819 bytes
- pwb.py Python File 1,86 kB
- README-conversion.txt Tekstdocument 5,66 kB
- user-config.py Python File 105 bytes
- apicache Bestandsmap
- scripts Bestandsmap
- generate_user_files.py Python File 10,1 kB
- LICENSE Bestand 1,04 kB
- pywikibot.lwp LWP-bestand 641 bytes
- setup.py Python File 1,02 kB
- version Bestand 62 bytes
- externals Bestandsmap
- tests Bestandsmap
- generate_user_files.pyc Compiled Python File 8,97 kB
- oude-user-config.py Python File 105 bytes
- README Bestand 323 bytes
- throttle.ctrl CTRL-bestand 0 bytes

Programma's (1)

cmd.exe

Documenten (21)

- pywikipedia-rewrite.zip
- contents.html
- rino_tools.js
- mediawiki118_fresh_installation.sql
- SimpleSeleniumTestDB.sql
- tables-1.18.sql
- tables-1.17.sql
- tables-1.16.sql
- tables-1.15.sql

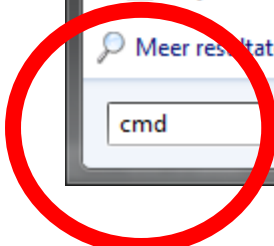
Bestanden (25)

- CmdLine.py
- CmdLine.py
- pywikipedia-rewrite.zip
- EasyInstall.txt
- setuptools.txt
- .spyder.ini
- pywikipedia-rewrite.externals-nightly (1).zip
- EasyInstall.txt

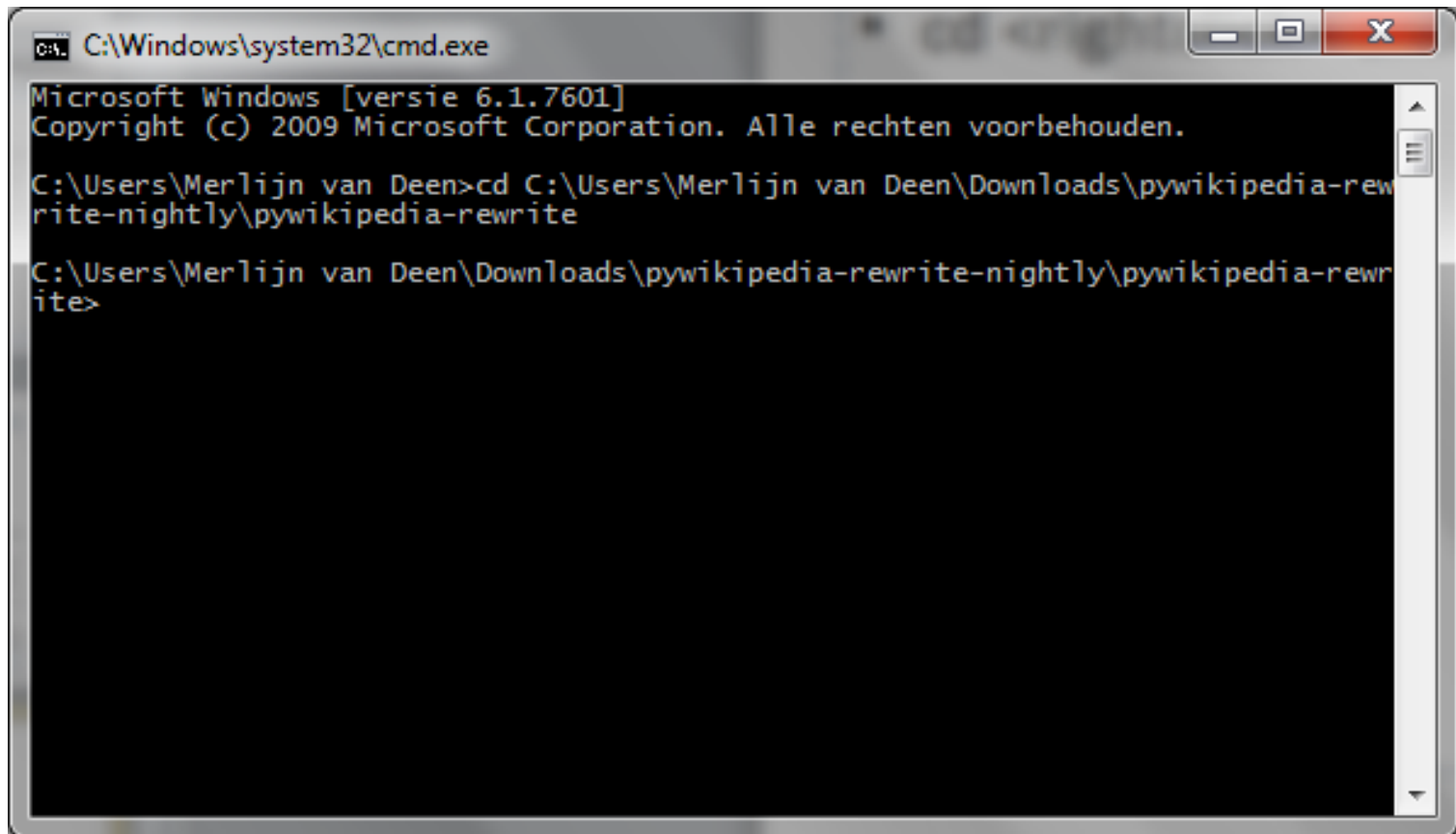
Meer resultaten weergeven

cmd

Afsluiten



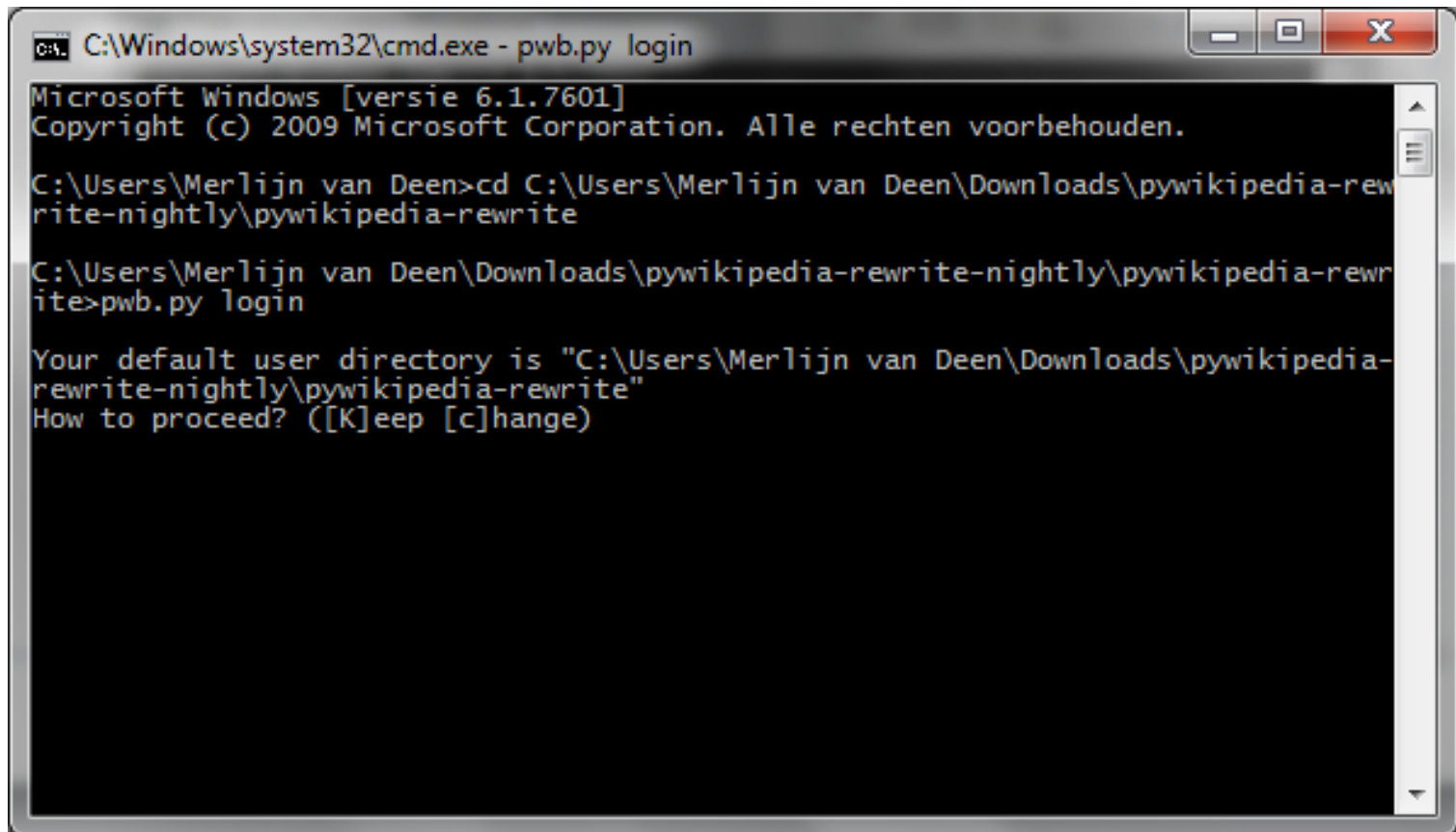
- cd <rightclick> enter



```
C:\Windows\system32\cmd.exe
Microsoft Windows [versie 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Alle rechten voorbehouden.

C:\Users\Merlijn van Deen>cd C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly\pywikipedia-rewrite
C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly\pywikipedia-rewrite>
```

- `pwb.py login -all`
- **on linux:** `python pwb.py login -all`



```
C:\Windows\system32\cmd.exe - pwb.py login
Microsoft Windows [versie 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. Alle rechten voorbehouden.

C:\Users\Merlijn van Deen>cd C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly\pywikipedia-rewrite
C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly\pywikipedia-rewrite>pwb.py login

Your default user directory is "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly\pywikipedia-rewrite"
How to proceed? ([K]eep [c]hange)
```

magic key combination

- K <enter>
- n <enter>
- y <enter>
- <enter>
- (your language code)
<enter>
- (your bot user name)
<enter>
- S <enter>
- Y <enter>
- 22 <enter>
- wikidata <enter>
- (your bot user name)
<enter>
- N <enter>
- N <enter>
- (enter password)
<enter>

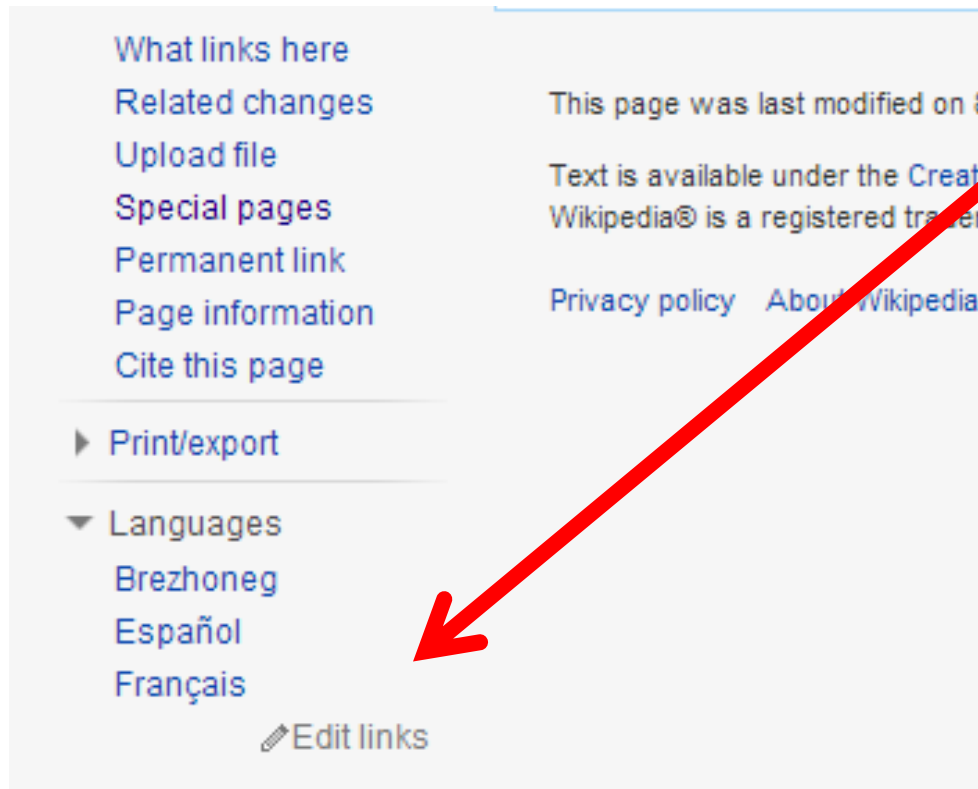
add claims on wikidata: country

- choose a category
 - e.g. en:Category:Bridges_in_France



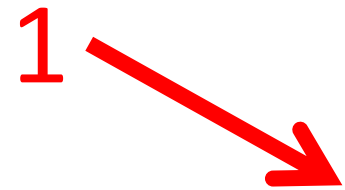
Determining Q & P

- Browse the category, click an article
- Click on 'Edit links'



Statements

[add]



Statements

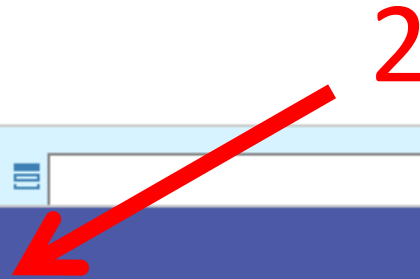
I accept the
this messa

[save|cancel]

country

country

sovereign state of this item



[add]

country calling code

dialed on phone for a country after the international dialing prefix (precede value by +)

Also known as: country dialing code

I accept these terms for
this message again.

Statements

country



France

[save|cancel] ?

3

France

European Parliament constituency

[add]

France

Wikipedia disambiguation page

List of pages linked to thi

Language



Code

brezhoneg

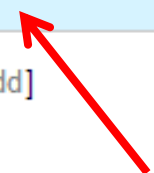
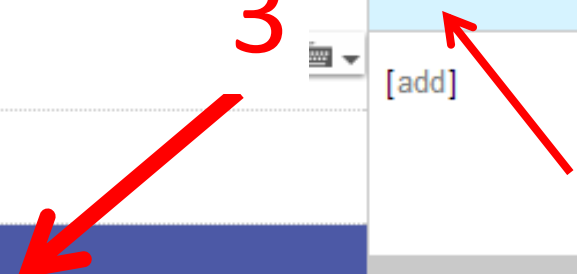
br

France

country mostly in Western Europe

[edit]

4



Create a new item
Item by title
Recent changes
Random item
Help
Donate

Language select
English

Statements

[country](#)

France

▼ 0 sources

www.wikidata.org/wiki/Property:P17



P17 = country

Item by title
Recent changes
Random item
Help
Donate

Language select
English
Select

Statements

country

[France](#)

▼ 0 sources

www.wikidata.org/wiki/Q142



Q142 = France

```
pwb.py claimit  
-cat:Bridges_in_France  
P17 Q142
```

adds the claim **P17=Q142** (country=France) to all pages in **Category:Bridges_in_France**

**Run for a few edits,
then abort with ctrl-c**

```
C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite>pwb.py claimit -cat:Category:Bridges_in_France P17 Q142
Processing [[en:List of medieval bridges in France]]
A claim for p17 already exists. Skipping
Processing [[en:Barentin Viaduct]]
A claim for p17 already exists. Skipping
Processing [[en:Pont de Bonpas]]
Adding p17 --> q142
Sleeping for 9.1 seconds, 2013-05-25 11:55:21
Processing [[en:Bras de la Plaine Bridge]]
Adding p17 --> q142
Sleeping for 8.6 seconds, 2013-05-25 11:55:32
```

```
File "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite\scripts\clai.py", line 114, in <module>
    main()
File "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite\scripts\clai.py", line 111, in main
    bot.run()
File "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite\scripts\clai.py", line 81, in run
    item.addClaim(claim)
File "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite\pywikibot\page.py", line 2533, in addClaim
    self.repo.addClaim(self, claim, bot=bot)
File "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite\pywikibot\site.py", line 3393, in addClaim
    data = req.submit()
File "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite\pywikibot\data\api.py", line 234, in submit
    self.site.throttle(write=self.write)
File "C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite\pywikibot\throttle.py", line 266, in __call__
    time.sleep(wait)
```

KeyboardInterrupt

```
C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly (2)\pywikipedia-rewrite>_
```



wikidata: convert infoboxes

- import biological classification from pages
 - e.g. from en:Template:Taxobox



OpenCage, CC-BY-SA-2.5,
[commons.wikipedia.org/wiki/
File:Catopuma_temminckii.jpg](https://commons.wikipedia.org/wiki/File:Catopuma_temminckii.jpg)

template-claim mapping

```
{{Taxobox
| name = Asian golden cat
...
| regnum = [[Animal]]ia          P75
| phylum = [[Chordate|Chordata]] P76
| classis = [[Mammal]]ia        P77
| ordo = [[Carnivora]]          P70
| familia = [[Felidae]]         P71
...
}}
```



```
pwb.py harvest_template
```

```
-cat:Felines
```

```
-template:"Taxobox"
```

```
-namespace:0
```

```
regnum P75
```

```
phylum P76
```

sets the claims **P75** and **P76** based on the **Taxobox** template for all pages in **Category:Felines**

```
C:\Users\Merlijn van Deen\Downloads\pywikipedia-rewrite-nightly  
(2)\pywikipedia-rewrite>pwb.py harvest_template -cat:Felines -  
template:"Taxobox" -namespace:0 regnum P75 phylum P76
```

```
Processing [[en:Felinae]]
```

```
Adding p75 --> q729
```

```
Sleeping for 8.9 seconds, 2013-05-25 13:49:37
```

```
Adding p76 --> q10915
```

```
Sleeping for 6.4 seconds, 2013-05-25 13:49:50
```

kingdom (taxonomy)

P75

animal

[edit]

▼ 1 source

[edit]

imported from

English
Wikipedia

[add source]

[add]

phylum

P76

chordates

[edit]

▼ 1 source

[edit]

imported from

English
Wikipedia

[add source]

[add]

KONINKRIJK BELGIE



RIJBEWIJS

Permis de conduire
Permiso de Conducción
Řidičský průkaz
Kørekort
Juhiluba
Άδεια οδήγησης
Driving Licence
Ceadúnas Tiomána
Patente di guida
Vadītāja apliecība

Führerschein
Vairuotojo pažymėjimas
Vezetői engedély
Licenzja tas-Sewqan
Prawo Jazdy
Carta de Condução
Vodičský preukaz
Vozniško dovoljenje
Ajokortti
Körkort

Model van de
EUROPESE GEMEENSCHAPPEN

Other widely-used scripts:

- `replace.py` (search-and-replace on pages)
- `category.py` (add/change/remove categories)
- `redirect.py` (fix double redirects)

`data_ingestion.py` (import images to commons)