

KONSUMERA WIKIMEDIAS DATA



WIKIMEDIA
SVERIGE

Hack for Heritage
2018-10-06, Visby

André Costa

Operativ chef, andre.costa@wikimedia.se

@Lokal_Profil



WIKIMEDIA
SVERIGE

Action API

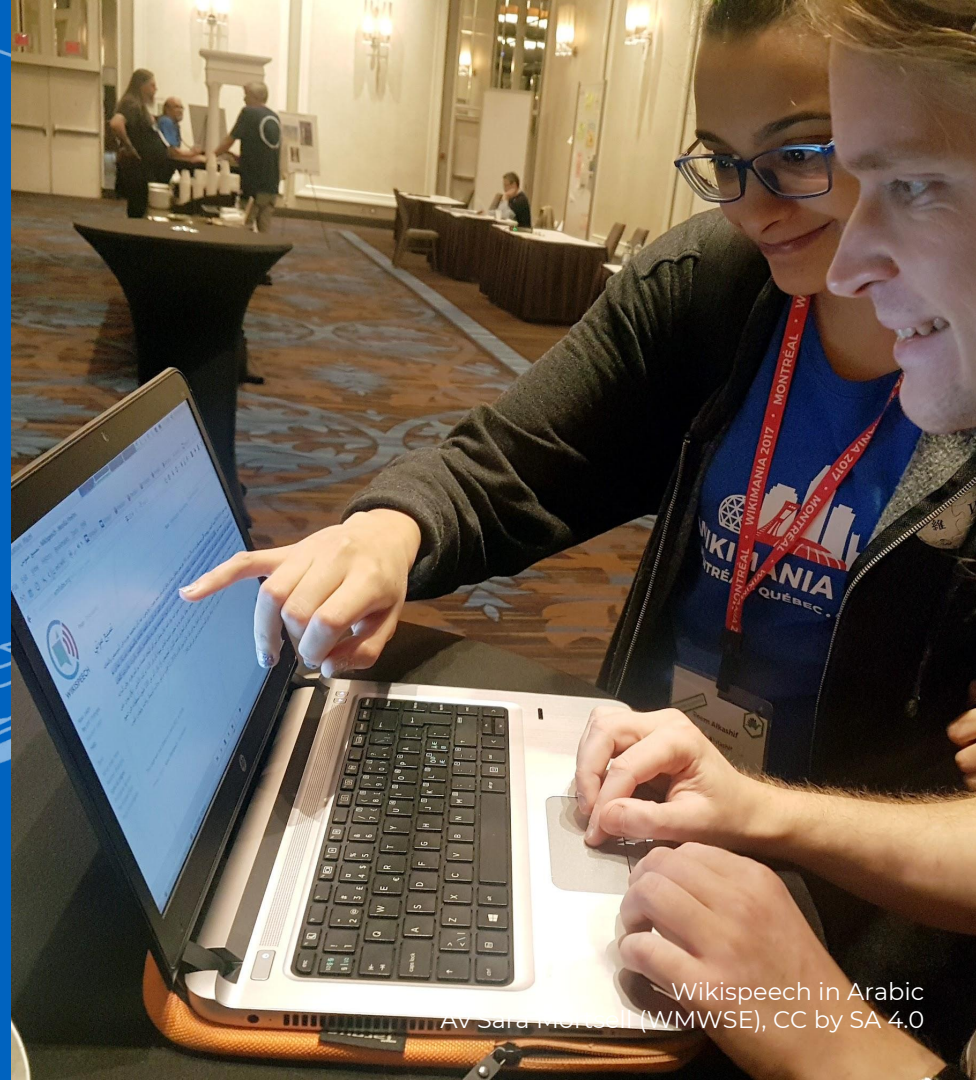
- Introduktion med några exempel
- Sandlådan
- Restriktioner

Sparql

Rest API

Dumpar

SQL



Wikimedia Action API

Endpoint:

en.wikipedia.org/w/api.php

commons.wikimedia.org/w/api.php

etc.

- Även automatisk dokumentation
- Fördjupande: mediawiki.org/wiki/API

Konstruera en fråga

- **Action:** `action=query`
 - **query** / edit / login / upload etc.
- **Action-specifika parametrar**
 - Ytterligare under-moduler för query
 - Egenskaper: Sidinnehåll, ingress, första bilden, koordinat, sidvisningar...
 - Listor: alla artiklar nära en viss koordinat, all media i en kategori...

Ett exempel

(Ingress för artiklarna API och Wikimedia)

`action=query`

`prop=extracts`

`titles=API|Wikimedia`

`redirects=`

`exintro=`



Ge mig...

- ...alla artiklar inom 500 m av vår position

- `/w/api.php?action=query&list=geosearch&gscoord=57.6307857|18.3056511&gsradius=500&gslimit=10&gsnamespace=0`

- ...alla bilder med ordet HACK

- `/w/api.php?action=query&list=search&srsearch=Hack&srnamespace=6&srinfo=totalhits&srlimit=10`

- ...alla artiklar i kategorin Cultural heritage

- `/w/api.php?action=query&list=categorymembers&cmtitle=Category:Cultural_heritage&cmlimit=10`



API sandlåda

Special:ApiSandbox



WIKIMEDIA
SVERIGE

Leker strödd i sandkasse
Av Bjørn Erik Pedersen CC BY-SA 4.0

Generatorer

En sökning kan vara grunden för en andra sökning:

- ingress för alla artiklar i en kategori
- `/w/api.php?action=query&prop=extracts&exlimit=10&exintro=&generator=categoremembers&gcmtitle=Category:Cultural_heritage&gcmlimit=10`

Notera att alla parametrar får ett g-prefix

LÄNKADE DATA / SPARQL

query.wikidata.org

[exempel](#)



REST API

mediawiki.org/wiki/REST_API



WIKIMEDIA
SVERIGE

Kücheninterieur mit schlafender Frau
Av David Teniers the Younger Public Domain



Dumpar dumps.wikimedia.org



WIKIMEDIA
SVERIGE

Aarhus container terminal
Av Guillaume Baviere CC BY 2.0

SQL much?

quarry.wmflabs.org



WIKIMEDIA

Panorama of Carrara quarry
Av Wittylama CC BY-SA 4.0



TACK!

bit.ly/H4H2018-Wikimedia



WIKIMEDIA
SVERIGE

Usage restrictions

Wikimedia Terms of Use

wikimediafoundation.org/wiki/Terms_of_Use

Copyright

For reuse follow the licenses and credit the source

API etiquette

mediawiki.org/wiki/API:Etiquette



API etiquette (summary)

- Combine your requests
- Don't query too frequently
- Use the `maxlag`-parameter if possible
- Use a descriptive `User-Agent` header
 - The name of the application
 - Contact details
 - meta.wikimedia.org/wiki/User-Agent_policy



Useful links

- mediawiki.org/wiki/API
- mediawiki.org/wiki/API:FAQ
- mediawiki.org/wiki/API:Etiquette
- mediawiki.org/wiki/API:Errors_and_warnings
- mediawiki.org/wiki/API:Data_formats
- mediawiki.org/wiki/API:Cross-site_requests
- en.wikipedia.org/wiki/Special:ApiSandbox

