

Parallella texter från Migrationsverket

SND-ID: ext0329-1.

Ingår i samling hos SND: [Parallella texter från myndigheter](#)

Skapare/primärforskare

Simon Dahlberg - Institutet för språk och folkminnen, Språkrådet
Institutet för språk och folkminnen, Språkrådet

Forskningshuvudman

[Institutet för språk och folkminnen](#) - Språkrådet

Beskrivning

Parallella texter nedladdade med hjälp av "w3m -dump" från ett ubuntu-skal, ifrån Migrationsverkets webbplats.

Texterna har laddats ner med hjälp av kommando 'w3m -dump' ifrån ett ubuntu-skal, varpå resulterande textfiler har skalats av till att innehålla endast den intressanta texten (inga menyer och dylikt).

Urvalsmetod

Flerspråkigt parallellt material.

Dataformat / datastruktur

[Text](#)

Datainsamling 1

- Insamlingsmetod: Självadministrerad text och/eller dagböcker: webbaserat
- Tidsperiod(er) för datainsamling: 2019-01-01 - 2019-01-31

Resurstyp

Korpus

Tänkt användning

Språkteknologiskt datorprogram

Text corpus

- Antal språk
 - Flerspråkig
- Språk
 - Svenska (swe)
Texter: 33
 - Amhariska (amh)
Texter: 23
 - Arabiska (ara)

- Texter: 33
- Azerbajdzjanska (aze)
Texter: 27
- Central Kurdish (ckb)
Texter: 29
- Engelska (eng)
Texter: 33
- Persiska (fas)
Texter: 32
- Kroatiska (hrv)
Texter: 23
- Armeniska (hye)
Texter: 24
- Georgiska (kat)
Texter: 1
- Northern Kurdish (kmr)
Texter: 28
- Mongoliska (mon)
Texter: 25
- Dari (prs)
Texter: 28
- Pashto (pus)
Texter: 28
- Romani (rom)
Arli (dialekt)
Texter: 24
- Ryska (rus)
Texter: 33
- Somaliska (som)
Texter: 29
- Spanska (spa)
Texter: 31
- Albanska (sqi)
Texter: 27
- Thailändska (tha)
Texter: 4
- Tigrinja (tir)
Texter: 29
- Turkiska (tur)
Texter: 2

- Uzbekiska (uzb)
Texter: 25
- Kinesiska (zho)
Texter: 3
- Franska (fra)
Texter: 31

[Mer..](#)

- Modalitet

Skriftspråk

- Storlek

Ord: 29008 (swe)

Texter: 33 (swe)

Ord: 438614 (TOT)

Texter: 580 (TOT)

- Källa

migrationsverket

www.migrationsverket.se

Geografisk utbredning

Geografisk plats: [Sverige](#)

Ansvarig institution/enhet

Språkrådet

Medverkande

Institutet för språk och folkminnen, Språkrådet

Identifierare

URL:

<http://liljeholmen.sprakochfolkminnen.se/sprakresurser/version/20190124/myndighetsdata/texter>

URL:

<http://liljeholmen.sprakochfolkminnen.se/sprakresurser/version/20190124/myndighetsdata/texter/Migrationsverket>

Forskningsområde

[Samhälle och kultur](#) (CESSDA Topic Classification)

[Lagstiftning och rättssystem](#) (CESSDA Topic Classification)

[Konflikter, säkerhet och fred](#) (CESSDA Topic Classification)

[Internationell politik och organisationer](#) (CESSDA Topic Classification)

[Samhällsvetenskap](#) (Standard för svensk indelning av forskningsämnen 2011)

[Språk och litteratur](#) (Standard för svensk indelning av forskningsämnen 2011)

[Social välfärdspolitik och välfärdssystem](#) (CESSDA Topic Classification)

Nyckelord

[Invandringspolitik, Utvandring och invandring](#)

Tillgänglighetsnivå

Åtkomst till data via extern aktör
Data är fritt tillgängliga

Hemsida

[Länk till övergripande studie hos Institutet för språk och folkminnen](#)
[Direktlänk till resursen](#)

Kontakt för frågor om data

Gunnar Eriksson
gunnar.eriksson@sprakochfolkminnen.se

Denna resurs har följande relationer

Är del av [Myndighetsdata](#)

Relaterade forskningsdata i SND:s katalog

[Parallella texter från myndigheter](#)

Ingår i samling hos SND

[Parallella texter från myndigheter](#)

CLARIN Virtual Collection Registry

[Lägg till i samling](#)

En virtuell samling är kopplad till ett specifikt forskningsändamål och innehåller länkar till dataresurser i olika digitala arkiv. Samlingen är lätt att skapa, få åtkomst till och citera.

Read more about virtual collections on the [CLARIN website](#).

Ladda ner metadata

[DataCite](#)

[MetaShare](#)

[MetaShare-CMDI](#)

[DDI 2.5](#)

[DDI 3.3](#)

[DCAT-AP-SE 2.0](#)

[JSON-LD](#)

[PDF](#)

[Citering \(CLS\)](#)