

Engelsk-svensk guldstandard för ordlänkning (GES)

SND-ID: ext0283-1.

Skapare/primärforskare

Lars Ahrenberg - Linköpings universitet, Institutionen för datavetenskap

Maria Holmqvist - Linköpings universitet, Institutionen för datavetenskap

Forskningshuvudman

[Linköpings universitet](#) - Institutionen för datavetenskap

Beskrivning

En referenskorpus för ordlänkning uppdelad på tränings- och testdata. Meningarna är hämtade från Europarl engelska respektive svenska delar.

Data är skapade från den engelsk-svenska delen av den parallella korpusen Europarl och utgörs, för varje meningspar, av par av tokenidentifierare på formen n-m där n och m är heltal.

Dataformat / datastruktur

[Numeriska](#)

[Text](#)

Resurstyp

Korpus

Tänkt användning

Språkteknologiskt datorprogram

Text corpus

- Antal språk
Tvåspråkig
- Språk
 - Engelska (eng)
 - Svenska (swe)Meningar: 1164

[Mer..](#)

- Modalitet
Skriftspråk
- Storlek
Meningar: 1164
- Uppmärkning
 - Länkning
Manuell uppmärkning

Ansvarig institution/enhet

Institutionen för datavetenskap

Forskningsområde

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

[Språkteknologi \(språkvetenskaplig databehandling\)](#) (Standard för svensk indelning av forskningsämnen 2011)

Nyckelord

[Översättning](#), [Datorhantering av naturligt språk](#)

Publikationer

Maria Holmqvist and Lars Ahrenberg (2011). A Gold Standard for English-Swedish Word Alignment. In Proceedings of the 18th Nordic Conference on Computational Linguistics, Riga, Latvia, May 11-13, 2011.

Om du publicerat något baserat på det här datamaterialet, [meddela gärna SND](#) en referens till din(a) publikation(er). Är du ansvarig för katalogposten kan du själv uppdatera metadata/databeskrivningen via DORIS.

Tillgänglighetsnivå

Åtkomst till data via extern aktör

Data är fritt tillgängliga

Licens

[CC BY 4.0](#)

Hemsida

[A Gold Standard Word Alignment for English-Swedish](#)

Kontakt för frågor om data

Lars Ahrenberg

lars.ahrenberg@liu.se

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 \(CLS\)](#)