

Analyses

– from the perspective of the user

Staffan Melin 2024-05

Reasons for this work effort

Citations for analyses – precision and visibility


Documentation

Get contributions

Easier to use



Searching for an analysis



SPRÅKBANKENTEXT

A research infrastructure for language data
and a language technology research unit

Svenska


News and events Blog Research Tools **Analyses** Data FAQ About us Contact us

[Home](#) / [Analyses](#)

Analyses

On this page you can browse and search our analyses. Click on an analysis name to read more about it and get instructions on how to use it. You can work immediately with the analysis by clicking the platform icon in the Access column.

All (214) Collections (24) **Analyses (176)** Utilities (14)

Name or description	Task	Unit	Language		
Stanza	Lemmatisation	Token	Swedish		
Analysis	Task	Unit	Language	Citation	Access
Swedish lemmatisation using Stanza Sentence segments are analysed to enrich tokens with lemmas.	Lemmatisation	Token	Swedish	”	Run analysis: https://spraakbanken.gu.se/analyses/swe-lemmatization-stanza Available in:  Mink
Lemmatisation with Fossil Sentence segments are analysed to enrich tokens with lemmas.	Lemmatisation	Token	Swedish	”	Run analysis: https://spraakbanken.gu.se/analyses/swe-lemmatization-stanza

Finding an analysis



SPRÅKBANKENTEXT

A research infrastructure for language data
and a language technology research unit

Svenska

News and events Blog Research Tools **Analyses** Data FAQ About us Contact us

[Home](#) / [Analyses](#)

Swedish lemmatisation using Stanza

Sentence segments are analysed to enrich tokens with lemmas <Followed by a longer description.>

How to use it

```
from stanza.core.pipeline import Pipeline
from stanza.core.pipeline.tokenizer import Tokenizer
from stanza.core.pipeline.lemmatizer import Lemmatizer

tokenizer = Tokenizer()
lemmatizer = Lemmatizer()

pipeline = Pipeline([tokenizer, lemmatizer])

text = "Svenska texten är här."
tokens = pipeline.process(text)

print(tokens)
```

How to cite
Standard reference(s):
Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton and Christopher D. Manning. 2020. Stanza: A Python Natural Language Processing Toolkit for Many Human Languages. In Association for Computational Linguistics (ACL) System Demonstrations, 2020.
Analysis reference:
Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton and Christopher D. Manning. 2020. Stanza: A Python Natural Language Processing Toolkit for Many Human Languages. In Association for Computational Linguistics (ACL) System Demonstrations, 2020.

Related publications

Available in:  Mink

Task
Lemmatisation

Tool
[Stanza](#)

Model
<https://spraakbanken.gu.se/resurser/stanzasynt>

Language
Swedish

Unit
Token

DOI
[10.80338/60rv-4v83](https://doi.org/10.80338/60rv-4v83)

License
[MIT](#)

Analysis description

UNIVERSITY OF GOTHENBURG

SPRÅKBANKENTEXT

A research infrastructure for language data and a language technology research unit

Svenska

News and events Blog Research Tools **Analyses** Data FAQ About us Contact us

Home / Analyses

Swedish Lemmatisation using Stanza

Sentence set ... tokens with lemmas <Followed by a longer description.>

Available in: Mink

Task
Lemmatisation

Tool
[Stanza](#)

Model
<https://spraakbanken.gu.se/resurser/stanzasynt>

Language
Swedish

How to cite

Standard reference(s):
Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton and Christopher D. Manning. Language Processing Toolkit for Many Human Languages. In Association for Computational Linguistics. 2020.

Analysis reference:
Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton and Christopher D. Manning. Language Processing Toolkit for Many Human Languages. In Association for Computational Linguistics. 2020.

Related publications

Standard reference.
This citation is more general and not pointing to a specific version.

Analysis reference.
This citation is detailed and points to a specific version of the analysis.

Example of input/output

```
# sent_id 1
# ...
1 They they PRON PRP Case=Nom|Number=Plur 2 nsubj 4:nsubj -
2 buy buy VERB VBP Number=Plur|Person=3|Tense=Pres 0 root -
3 and and CONJ CC - 2 cc -
4 sell sell VERB VBP Number=Plur|Person=3|Tense=Pres 2 conj 0:root -
5 books book NOUN NNS Number=Plur 2 dobj 4:dobj SpaceAfter=No
6 . . PUNCT . - 2 punct -

# sent_id 2
# ...
```



GÖTEBORGS
UNIVERSITET

SPRÅKBANKENTEXT

