

Länkad förenkling

Evelina Rennes & Sarah Albertsson

Automatisk textförenkling

- Stort behov av enkla texter
- Problem:
 - Tidskrävande
 - Dyrt
 - Går inte att förenkla allt
- Automatisk förenkling verkar rimligt

StilLett

- Textanalys och textförenkling
- Målgrupp: webbredaktörer
- Analys: komplexitetsmått
- Förenkling: regelbaserad

The screenshot displays the StilLett web application interface. On the left, a dark sidebar contains configuration options for text simplification, each with a teal circular icon. On the right, a white panel shows the analysis results, categorized into 'Återkoppling', 'Lexikala mått', 'Enkla mått', and 'Strukturella mått'.

Configuration Options:

- ☒ Långa ord
Maximal ordlängd: 7 tecken
- ☒ Långa meningar
Maximal meningslängd: 15 ord
- Aktiva förenklingstyper:
 - ☒ Rak ordföljd
 - ☒ Passiv till aktiv form
 - ☒ Ta bort onödig information
 - ☒ Proximering
 - ☒ Citatomvändning
 - ☒ Meningsuppdelning
 - ☐ Egna regler:

Analysis Results:

Återkoppling
0 Synonymförslag
3 Meningsförenklingar
Textförenkling

Lexikala mått
SweVocTotal 0.636
SweVocC 0.386
SweVocH 0.057
SweVocD 0.011
SweVocS 0.182
SweVocK 0.011

Enkla mått
AWL 5.625
ASL 22.000
Lix 60.636
Ovix 65.163
Kix 95.509

Strukturella mått
LexDens 0.523
avgVerbalArity 2.583
numberOfContentWords 46.000
numberOfSentences 4.000
numberOfTokens 88.000
numberOfVerbs 12.000
sentenceLengthSD 3.162
numberOfVerbalRoots 3.000
avgSentenceDepth 8.250

English Wikipedia

Python (programming language)

From Wikipedia, the free encyclopedia

This article is about the programming language. For the snake genus, see [Pyt](#) ([disambiguation](#)).

Python is a widely used [high-level](#), [general-purpose](#), [interpreted](#), [dynamic programming language](#).^{[24][25]} Its design philosophy emphasizes code [readability](#), and its syntax allows programmers to express concepts in fewer [lines of code](#) than possible in languages such as [C++](#) or [Java](#).^{[26][27]} The language provides constructs intended to enable writing clear programs on both a [small and large scale](#).^[28]

Python supports multiple [programming paradigms](#), including [object-oriented](#), [imperative](#) and [functional programming](#) or [procedural](#) styles. It features a [dynamic type](#) system and automatic [memory management](#) and has a large and comprehensive [standard library](#).^[29]

Python interpreters are available for many [operating systems](#), allowing Python code to run on a wide variety of systems. [CPython](#), the [reference implementation](#) of Python, is [free and open-source software](#) and has a community-based development model, as do nearly all of its variant implementations. CPython is managed by the non-profit [Python Software Foundation](#).

Simple English Wikipedia

Python (programming language)

From Wikipedia, the free encyclopedia

This page is about the [programming language](#). For the [snake](#), see [python](#).

Python is an [open source programming language](#) made to both look good and be easy to read. A programmer named [Guido van Rossum](#) made it in 1991. Python is named after the television show [Monty Python's Flying Circus](#). Many examples and tutorials include jokes from the show.

Python is an [interpreted language](#). Interpreted languages do not need to be [compiled](#) to run. A program called an [interpreter](#) will run python code on any kind of computer it can run on itself. This means if the programmer needs to change the code they can quickly see the results. This also means Python is slower than a compiled language like [C](#), because it is not running [machine code](#) directly.

Python is a good programming language for beginners. It is a high-level language, which means a programmer can focus on what to do instead of how to do it. Writing programs in Python takes less time than in another language.

Python drew inspiration from other programming languages like C, [C++](#), [Java](#), [Perl](#), and [Lisp](#).

Linköpings kommun

 [Lyssna](#)

Mat ska vara god, näringsrik och av hög kvalitet. För att vara det måste den hanteras på rätt sätt. Miljökontorets kontroll gör att du som konsument får bra och säkra livsmedel. Miljökontoret kontrollerar även att tobakslagen följs på till exempel skolgårdar och restauranger.

Linköpings kommun LL

Mat och dryck ska vara god, näringsrik och ha hög kvalitet.

Därför ska den hanteras på rätt sätt.

Miljökontorets gör kontroller

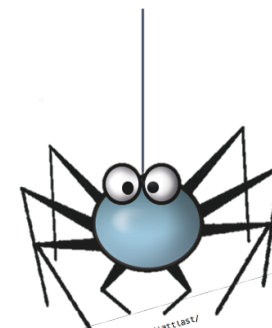
för att du ska få bra och säkra livsmedel.

När vi gör kontrollen

följer vi det som står i Livsmedelslagen och i EUs förordningar.

Insamling av korpus

- Kommuner och myndigheter
- Webbspindling



	Ordinary	Easy-To-Read	Total
Number of seed URLs	191	191	382
Number of collected URLs	182,777	2158	184,935
Number of URLs after filtering	181,983	2157	184,140
Raw crawl size	186.1 MB	2 MB	188.1 MB
Size after post-processing	2.78 GB	29.1 MB	2.81 GB
Number of pages	136,501	1629	138,130
Number of tokens	29.2M	334,491	29.6M

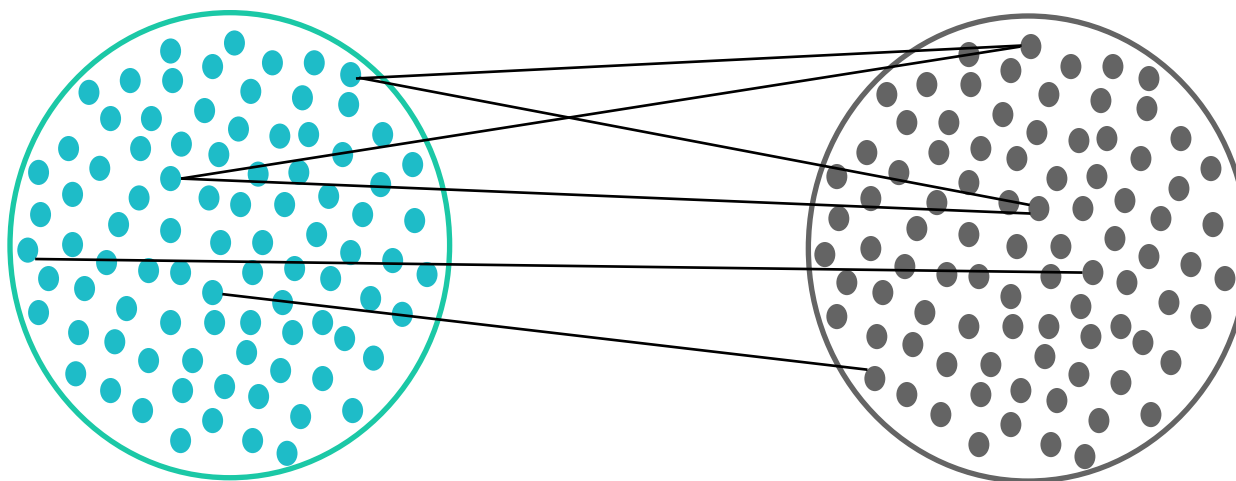
... nu då?

Länkning

Länkning

Meningar från
lättlästa dokument

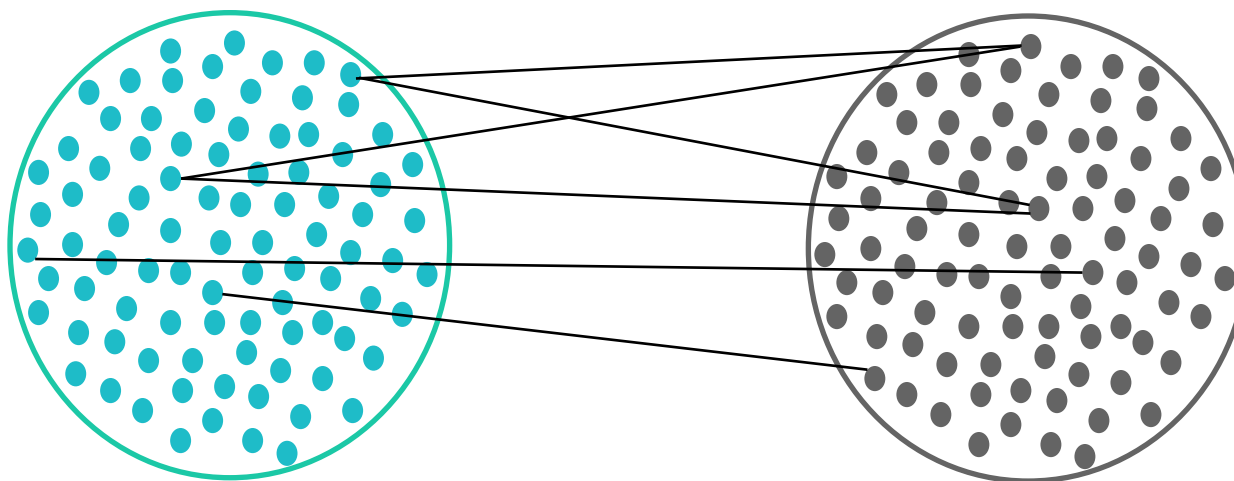
Meningar från icke-
lättlästa dokument



Länkning

Meningar från
lättlästa dokument

Meningar från icke-
lättlästa dokument



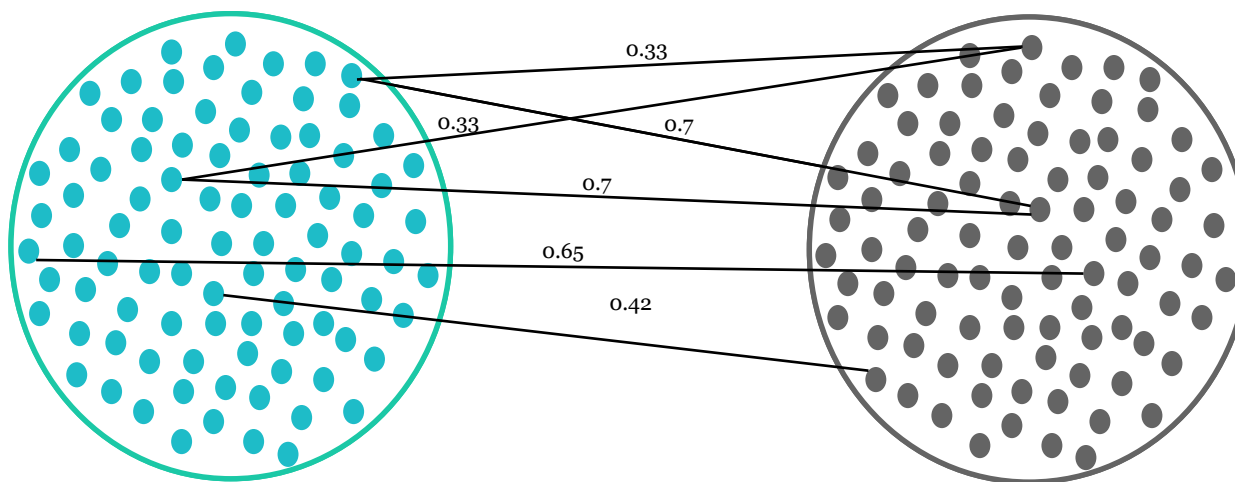
Miljökontorets gör kontroller för
att du ska få bra och säkra
livsmedel.

Miljökontorets kontroll för att
du som konsument får bra och
säkra livsmedel.

Länkning

Meningar från
lättlästa dokument

Meningar från icke-
lättlästa dokument



Mått varierar från -1 till 1 där.
Exakta kopior är 1.

Utvärdering

2%

Nedan kan du läsa en mening som systemet har slumpat fram. Meningarna som följer efter det är sådana som systemet tycker liknar den ursprungliga meningen. Din uppgift är att bedöma hur lika de är. Bedömningen gör du enligt kategoriuppdelningen, om du trycker på knappen så ser du samtliga kategorier.

Kategorier.

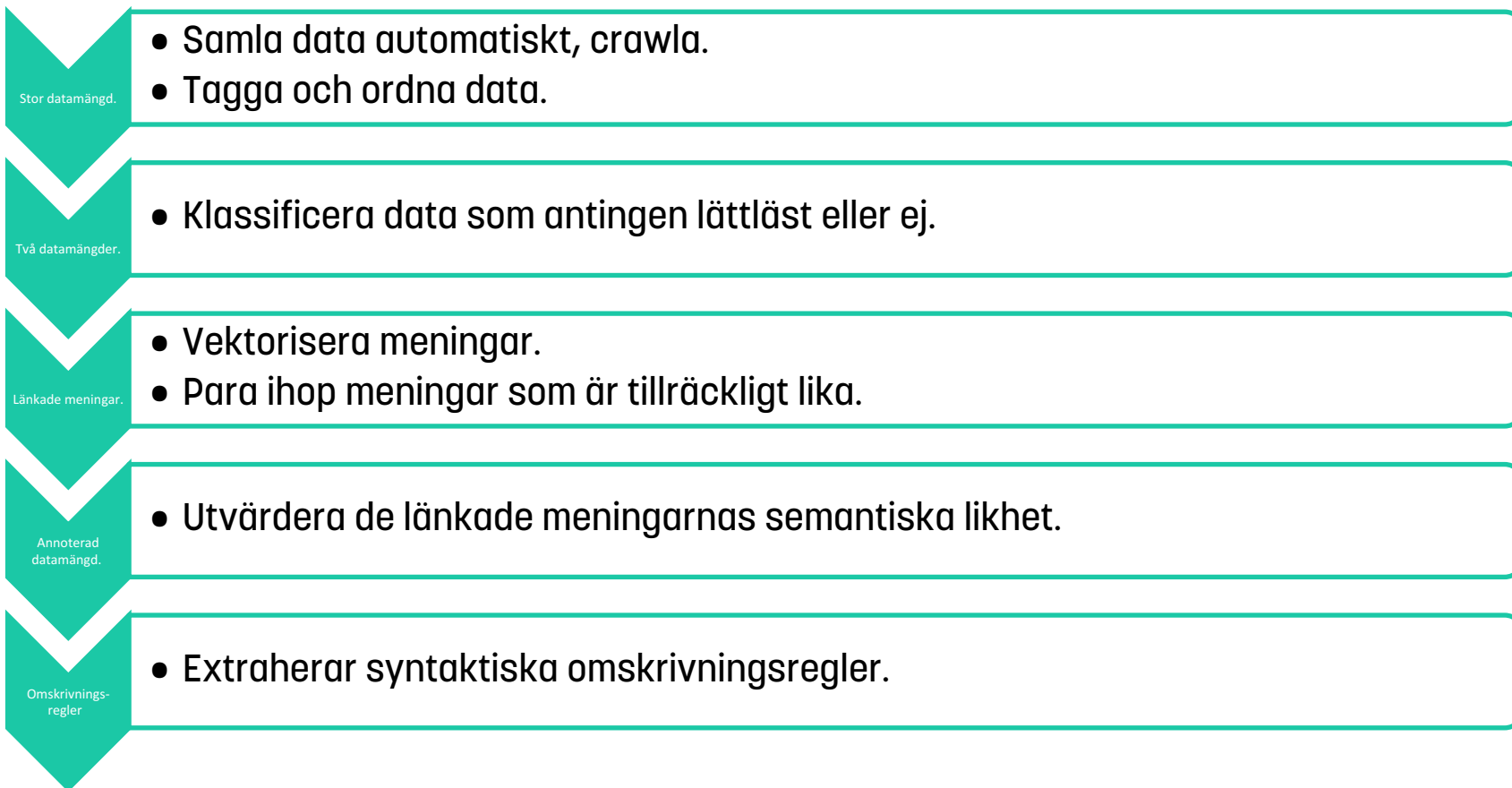
undrade han .

Hon undrade hur du har det .

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4

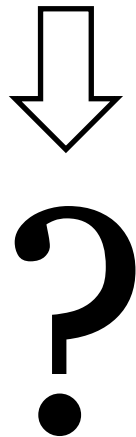
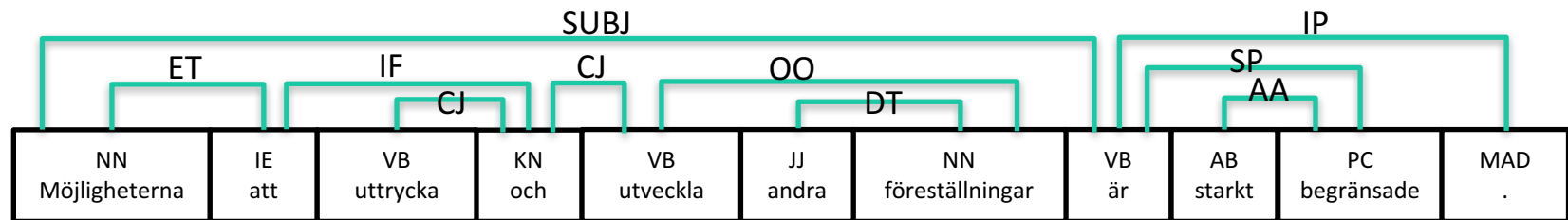
undrar Simson .

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4



FRAMTIDEN

Framtiden



Tack!

evelina.rennes@liu.se
sara1853@student.liu.se

www.liu.se

Semantisk likhet

0 De två texterna betyder inte samma sak och behandlar inte samma ämnesområde.

1 De två texterna beskriver olika koncept, idéer och handlingar, men kan hittas tillsammans i ett längre dokument på samma ämnesområde

2 De två texterna har olik mening, men de delade koncepten, idéerna och handlingarna i den mindre texten är relaterade (men inte lika) till de i den större texten.

3 De två texterna delar många viktiga idéer, koncept eller handlingar men de som uttrycks i den mindre texten är lika men inte identiska till det viktigaste i den större texten.

4 Texterna är mycket lika i mening och de viktigaste idéerna, koncepten eller handlingarna i den större texten representeras i den mindre texten.