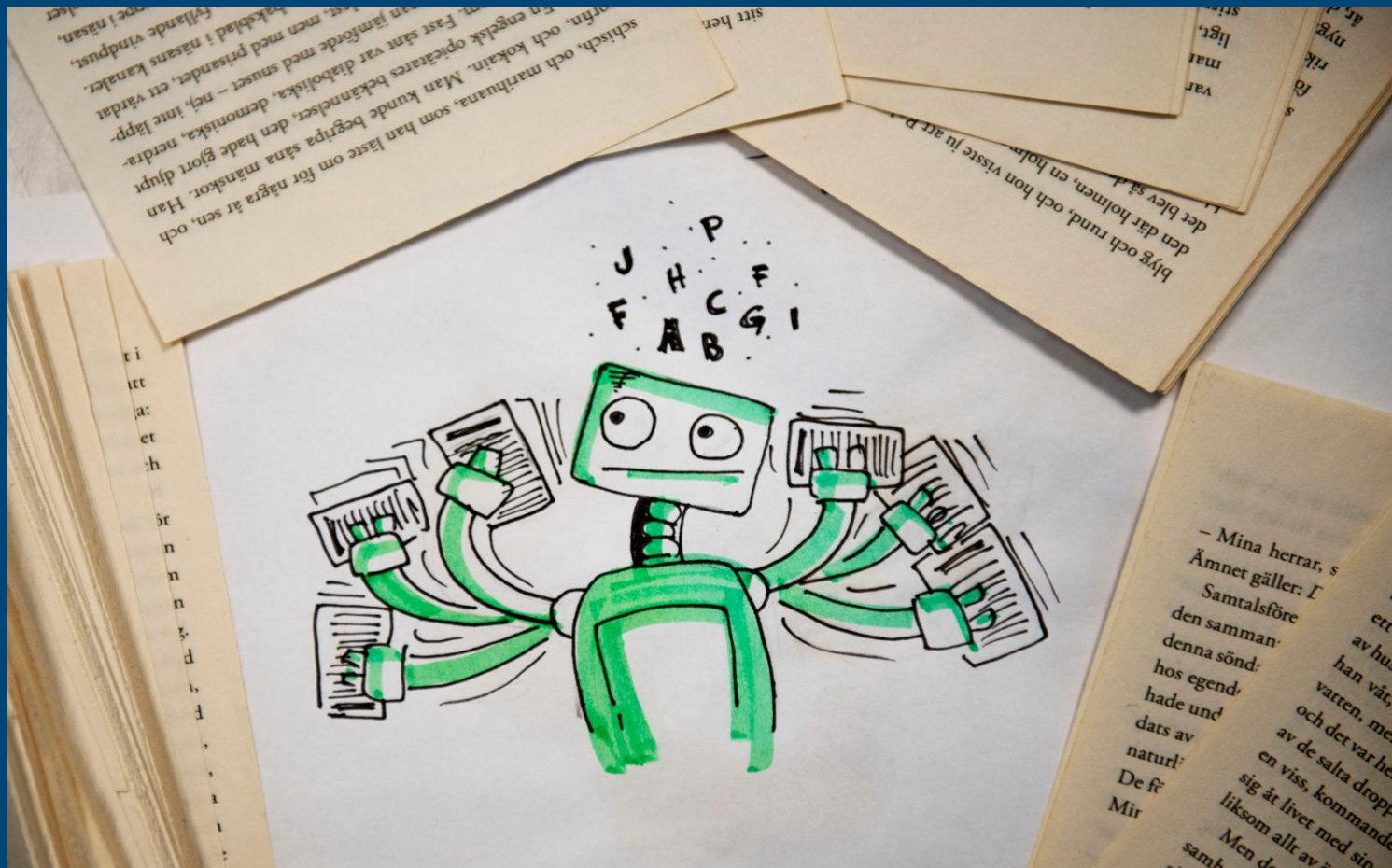
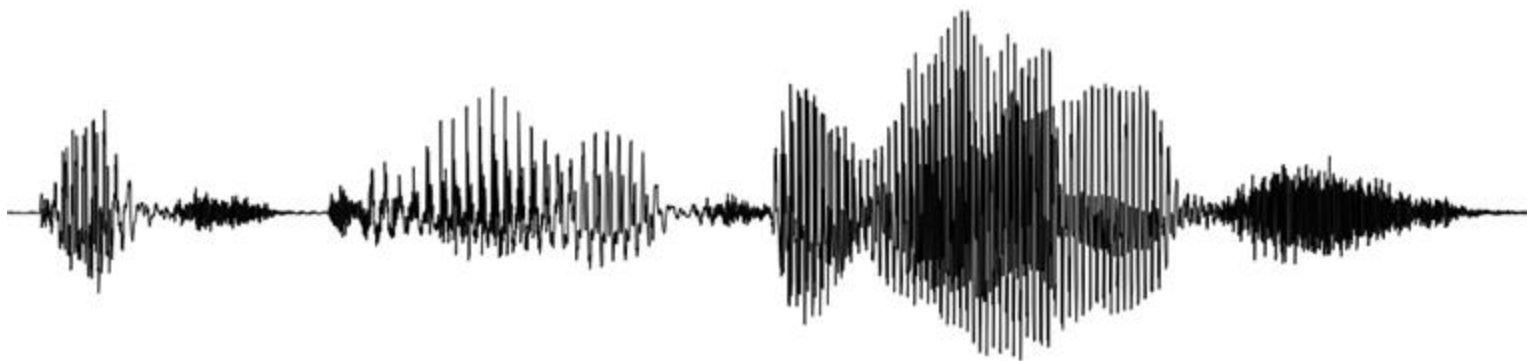


# KB-Whisper



2025-10-23



Herr talman! Det är viktigt att transkribera rätt!

# Varför är svensk taligenkänning viktigt?



The screenshot shows the Mediesök BETA search interface. At the top left is the logo and text "Mediesök BETA". At the top right is a green circle with an 'i' and the text "OM TJÄNSTEN". Below the header is a search bar containing "OpenAI" and a "Sök" button. To the right of the button is a help icon and the text "Hur fungerar det här?". Below the search bar is a hint: "Fritextsök, t.ex. 'Stockholm', 'Vädret', 'I år blir det varmt'". Below this is a summary: "Din sökning på **OpenAI** resulterade i 2 träffar." On the left side, there are filters: "KANAL" with "P4 Dalarna" (2 results) and "ÅR" with "2020" (2 results). The main content area shows two search results, both from "P4 Dalarna" and dated "2020-06-01".

**KANAL**  
P4 Dalarna 2

**ÅR**  
2020 2

**P4 Dalarna** Publiceringsdatum: 2020-06-01

*OpenAI* är ju ett ideellt forskningsbolag med sätt i USA. De jobbar med artificiell intelligens och driver på väldigt mycket för att vi ska utveckla medmänsklig AI. Alltså sådant som inte utplånar och förstör alltihopa utan vi ska dra nytta av det och skapa bra saker av det. Och i det här inkluderas att man ofta släpper nya typer av teknik och ny kod och nya algoritmer ganska fritt så att andra kan bygga vidare på den här tekniken som de skapar. Just det, så då har de kommit på någon smart teknik och sen så kan man då dela med sig den. Ja, precis.

**P4 Dalarna** Publiceringsdatum: 2020-06-01

Ja, och det här är ett annat exempel på när man har tagit den här tekniken från *OpenAI* men gått i en annan riktning. Och det här är en kille i San Francisco som heter Thomas Dimson och han har byggt ett helt uppslagsverk på internet med ord som låter som att de existerar men inte gör det. Och varje ord och varje beskrivning av det ordet har skapats av hans algoritm. Man kan söka på this word does not exist, så hittar man dit. Ett hittepåspråk i full utveckling, vad ska det vara bra för? Ja, det är ju inte bra för så mycket. Det är ju perfekt om man vill skapa ett nytt språk, kanske.

# Varför är svensk taligenkänning viktigt?

Lag (2018:1937) om tillgänglighet till digital offentlig service

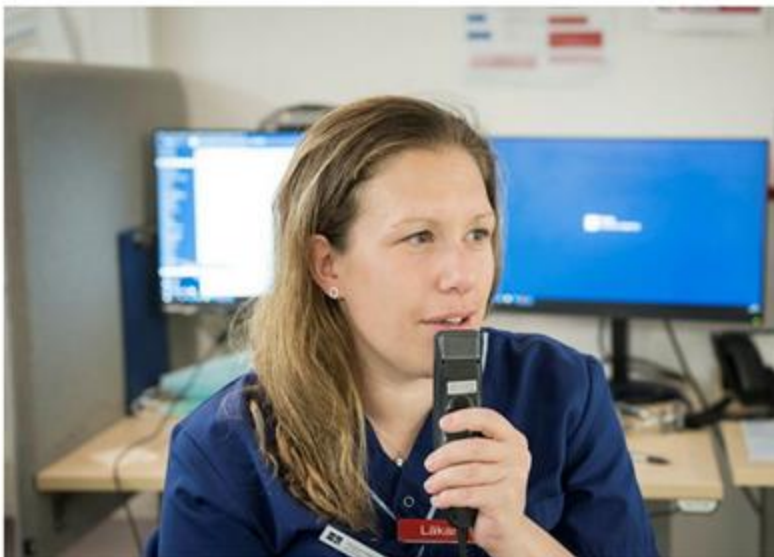
Lag (2023:254) om vissa produkters och tjänsters tillgänglighet

# Varför är svensk taligenkänning viktigt?

Dagens Medicin

## Lättare gå hem i tid när taligenkänning ersätter diktamen

Publicerad: 29 april 2022, 06:00



Pauline Hytner är ST-läkare på Capio S:t Görans akut.  
Foto: Anna Tärnhuvud/Bildbyrå

## Läkartidningen

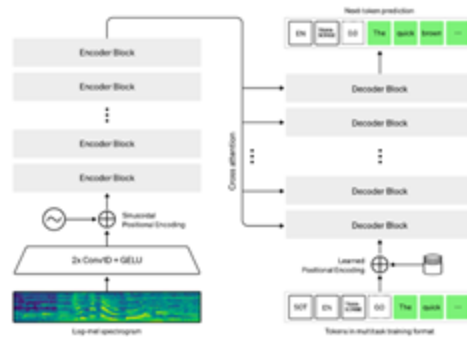
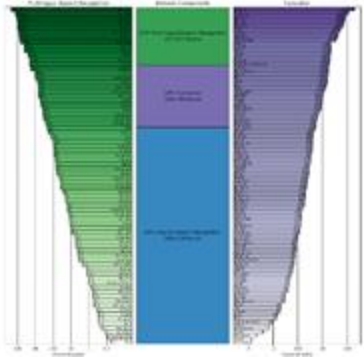
### Läkare på SU kritiska mot VGR:s taligenkänning: »Inte jättebra«

Är tidskrävande, feltolkar ord och snappar upp prat från omgivningen. VGR:s nya verktyg för taligenkänning får kritik från läkare på Sahlgrenska universitetssjukhuset i en ny enkät. »Huvudintrycket är att det inte fungerar jättebra«, säger Anders Thurin som sitter i styrelsen för läkarföreningen på SU.

### VGR får betala för kritiserat program för taligenkänning till 2032

# Hur tränar vi KB-Whisper?

680 000 h  
träningsdata



50 000 h svensk  
träningsdata

Stage 1 finträning



~15 000 h  
undertextlik  
träningsdata

~9000 h  
ordagrann  
träningsdata

Stage 2 finträning

Whisper  
tal-till-textmodell

**KB-Whisper**



# Transparens kring data

| Dataset      | Källa   | Stage 1 (h)   | Stage 2 (h)  |
|--------------|---------|---------------|--------------|
| Undertexter  | SVT     | 18,468        | 688          |
|              | SMDB    | 8,629         | 421          |
|              | YouTube | 7,164         | 2,001        |
| Riksdagen    |         | 21,949        | 5,119        |
| ISOF         |         | 54            | 54           |
| NST          |         | 250           | 250          |
| <b>Summa</b> |         | <b>56,514</b> | <b>8,533</b> |

Representerar en stor bredd och variation av svenska dialekter, sociolekter och brytningar.



Sök

Utökad sökning

Söktips

Sortering:

Relevans

Äldst först

Nyast först

Filter:

Tv (85)

1. **Canal+ SF-kanalen 2010-06-12**

TV-sändning

2. **Canal+ SF-kanalen 2010-10-03**

TV-sändning

3. **Canal+ SF-kanalen 2010-03-07**

TV-sändning

08:45-09:15 Alla vi barn i Bullerbyn, 09:15-09:45  
Alla vi barn i Bullerbyn, 09:45-10:15 Alla vi barn i  
Bullerbyn, 10:15-10:45 Alla vi barn i Bullerbyn

4. **Canal+ SF-kanalen 2010-05-16**

TV-sändning

08:45-09:15 Alla vi barn i Bullerbyn, 09:15-09:45  
Alla vi barn i Bullerbyn, 09:45-10:15 Alla vi barn i  
Bullerbyn, 10:15-10:45 Alla vi barn i Bullerbyn

5. **Canal+ SF-kanalen 2010-07-18**

TV-sändning

10:00-10:30 Alla vi barn i Bullerbyn, 10:30-11:00  
Alla vi barn i Bullerbyn, 11:00-11:30 Alla vi barn i  
Bullerbyn, 11:30-12:00 Alla vi barn i Bullerbyn

## Din sökning gav 85 träffar.

## Antal träffar fördelat på medietyper

Tv: 85

Film/Video: 1



TV

Film/Video

## Söktips datum

Det är möjligt att söka efter en specifik dag eller ett datumspann, det vill säga från och till och med ett visst datum.

1. datum:19900514

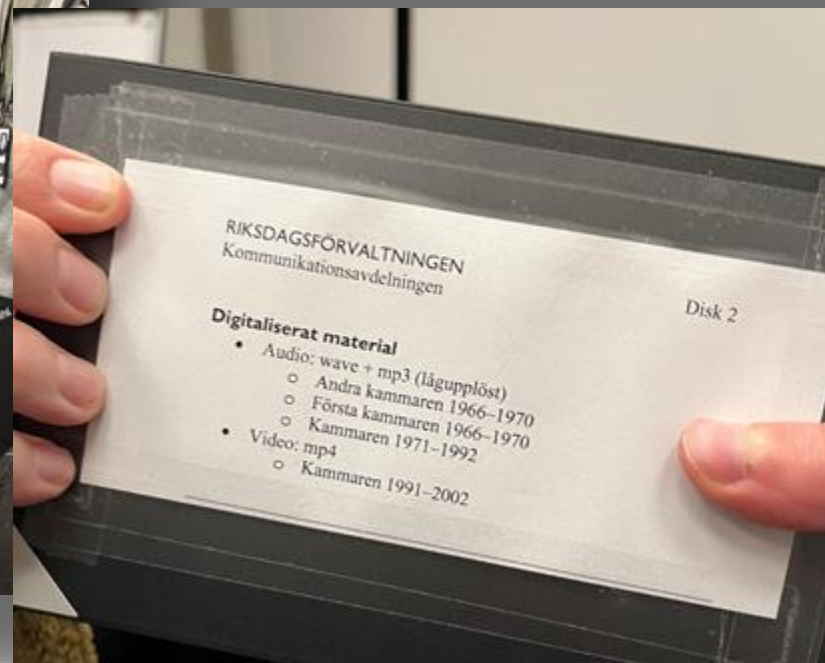
2. datum:19640816-19890101

Fler söktips finns på [vår söktipssida](#).





svt



SVERIGES  
RIKSDAG







## Aktuellt – Igår 21:00 – Nytt projekt: AI ska lära sig svenska dialekter

svtplay.se

# swedia

Project  
information

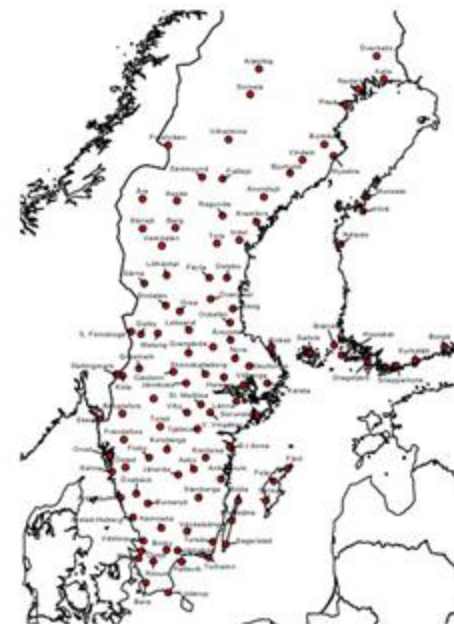
Maps

Databases

International  
cooperation

Publications

Contact



## Inspelning med Agda Carlsson

Accession

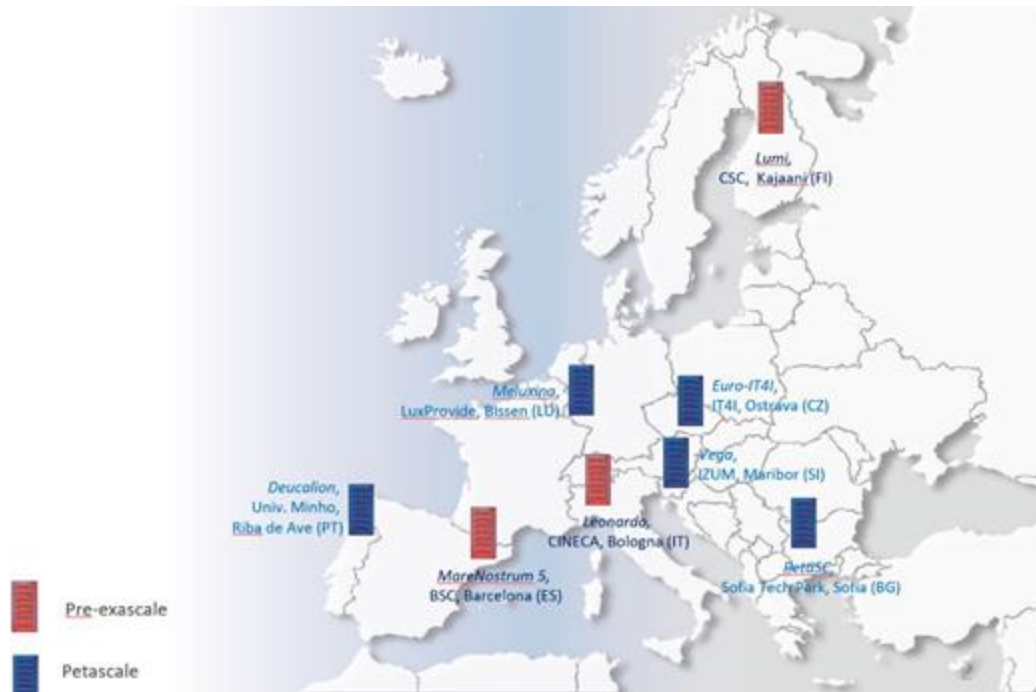
**Accessionsnummer:** ULMA BD 314 **År:** 1959 **Materialtyp:** arkiv **Arkiv:** Institutet för språk och folkminnen, Uppsala

## Beräkningskapacitet via EuroHPC JU

- 3500 nodtimmar (juni 23 - juni 24)
  - 50 000 nodtimmar (juni 24 - juni 25)
- 1 nod = 4 Nvidia A100



**EuroHPC**  
Joint Undertaking



# Modellernas prestanda mätt i felprocent på ordnivå (WER)

| Model size |        | FLEURS      | Common-Voice | NST         |
|------------|--------|-------------|--------------|-------------|
| tiny       | Our    | <b>13.2</b> | <b>12.9</b>  | <b>11.2</b> |
|            | OpenAI | 59.2        | 67.8         | 85.2        |
| base       | Our    | <b>9.1</b>  | <b>8.7</b>   | <b>7.8</b>  |
|            | OpenAI | 39.6        | 52.1         | 53.4        |
| small      | Our    | <b>7.3</b>  | <b>6.4</b>   | <b>6.6</b>  |
|            | OpenAI | 20.6        | 26.4         | 26.4        |
| medium     | Our    | <b>6.6</b>  | <b>5.4</b>   | <b>5.8</b>  |
|            | OpenAI | 12.1        | 15.8         | 17.1        |
| large-v3   | Our    | <b>5.4</b>  | <b>4.1</b>   | <b>5.2</b>  |
|            | OpenAI | 7.8         | 9.5          | 11.3        |

Vi förbättrar prestandan för små modeller avsevärt. Taligenkänning som håller hög kvalitet kan nu köras resurseffektivt på laptops.

# Whisper Web

Transkribera tal direkt i din webbläsare!



Från URL



Från fil



Spela in

**Transkriberingen utförs lokalt på din enhet. Inga uppgifter skickas någonstans.**

Skapad av Pierre Measure baserat på en demo av Xenova ❤️

<https://kb-whisper.mesu.re/>



# KBx - en samarbetsyta mellan KB-labb, utveckling och produktion

Startade för att få igång användningen av AI internt på KB

- Testa i liten skala
- Involvera många i olika delar av verksamheten
- Kompetensutveckling
- Bygga upp ambassadörsskap
- En grund i:
  - 2 x 50% data scientists
  - 1 x 50% produktledare/teamlead



# Projekt

Hur kan vi navigera och söka i samlingar där objekten saknar beskrivning och metadata?

- Bildpilot – söka bland vykort.  
Öppet tillgänglig: <https://lab.kb.se/bildsok>
- Mediesök – göra lokalradio sökbart

Hur kan vi rationalisera interna processer etc?

- Tablåkomplettering – automatisera ett flöde som idag är manuellt

På gång: Metadataberika affischer



Bildlicens: <https://id.kb.se/policy/kb/digisaml/1.0>

# Mediesök

En tjänst för att söka i radiomaterial

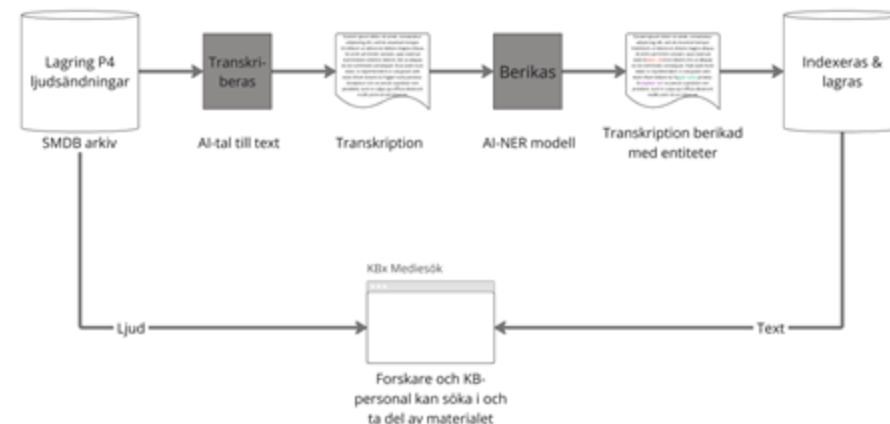
Varför ljud?

- Dold samling
- Tidsbesparande för handläggare
- Öppnar för forskning
- <https://lab.kb.se/mediesok/>

|                                         |                            |
|-----------------------------------------|----------------------------|
| 20:03-21:00 Sportextra                  |                            |
| 21:00-22:00 Ekot                        |                            |
| 22:00-22:10 Ekot                        |                            |
| 22:10-22:12 Radiosporten                |                            |
| 22:12-23:00 Karlavagnen ▾               |                            |
| 23:00-23:07 Ekot                        |                            |
| 23:07-00:00 Karlavagnen ▾               |                            |
| Utgivning ▲                             |                            |
| År/datum                                | 2008-07-21                 |
| Kanal                                   | SR P4 lokal                |
| Utgivning                               | Luleå : SR, P4, Norrbotten |
| Utgivningsland                          | Sverige                    |
| Anmärkningar [...] ▼                    |                            |
| Exemplar ▲                              |                            |
| MPEG-1 layer3 YA_sr_p4nbtn_2008-07-21 ▾ |                            |
| DVD+R L08-1499 ▾                        |                            |

# Berikning med modeller

- Transkribering av materialet
- Identifiera entiteter (personnamn, organisationsnamn, platsnamn, osv.)



Orsaken kan vara att vittnen känner en ökad rädsla för våld eller hot, skriver Dagens Nyheter. Det kan byggas en husbilsamping på Södermalm i Stockholm där Tantogården låg innan den brann ner förra våren. Det är exploateringskontoret som kommer med förslaget men det verkar inte särskilt populärt. En som ofta promonerar i området är Gertrud Höberleid och hon tycker inte att någon camping är en bra idé. Nej, det tycker jag inte. För att det finns väl inte möjligheter att ha hygieniska förhållanden, duschar och toaletter och sådana där saker där, det tror jag inte.

Entiteter

- Identifiera vilka "ämnen" som avhandlas och tagga materialet.
- Identifiera "sentiment" (positiv/negativ/neutral)?
- Identifiera spann som innehåller musik (i fall delar av materialet kan tillgängliggöras för en bredare publik).
- Koppla tablåinformation till sändningar.
- Dubletthantering