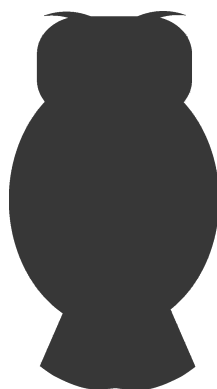

Strix

Användarhandledning och övningar
Språkbankens höstworkshop 2017

<https://spraakbanken.gu.se/strix>
sb-strix@svenska.gu.se

17 oktober 2017



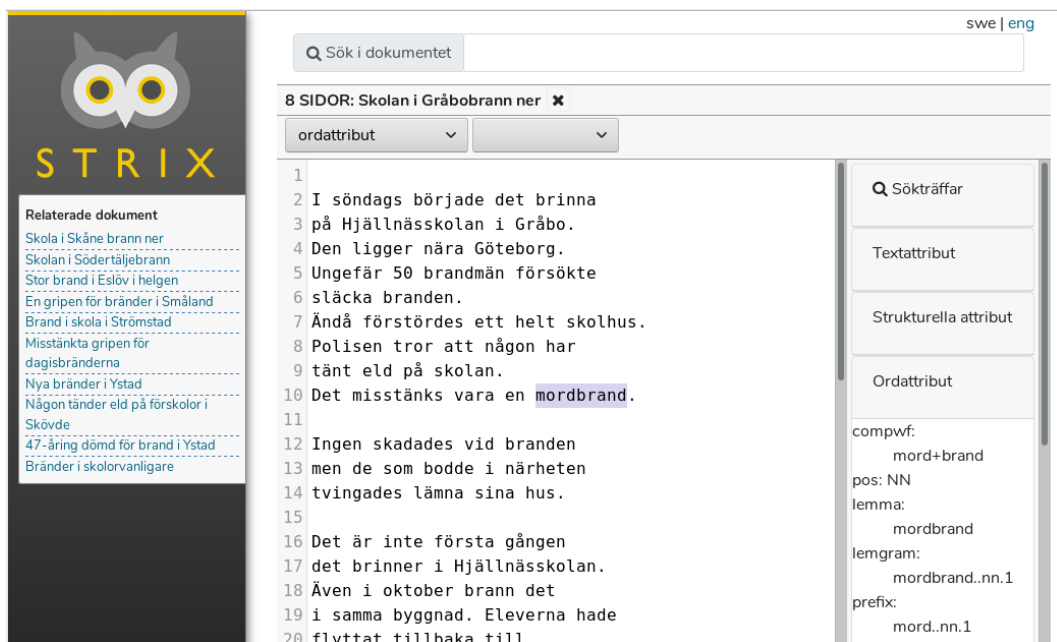
ÖVERSIKT

Strix är den del i Språkbankens infrastruktur som öppnar upp för dej som användare att få ett dokumentcentriskt perspektiv i jämförelse med till exempel Korp som ger ett meningscentriskt eller formmässigt perspektiv. Det ger möjlighet att svara på andra frågor än vad Korp kan mot i stora delar samma korpusar (textmaterial).

Med hjälp av egenskaper, både givna och beräknade, uttryckta som attribut kan det mesta av innehållet i Strix filtreras med stor precision. Med urval i form av samlingar kan olika texttyper filtreras fram, till exempel “anförande”. På samma sätt kan attribut kopplade till dokumentet, “textattribut”, eller delar av dokument “strukturella attribut” och på ordnivå “ordattribut” bidra till att få fram de bästa dokumenten utifrån någon kombination av attribut och andra filtertyper.

“RD” i de olika komponenterna står för “Riksdagens öppna data”.

ÖVNING 1 - TERMSÖKNING



The screenshot shows the Strix search interface. On the left is a sidebar with the Strix logo (an owl) and a list of related documents under the heading "Relaterade dokument". The main area displays search results for the query "mordbrand". The results are shown as a list of lines from a document, with line 10 containing the term "mordbrand" highlighted. On the right side, there is a panel for search results, showing the search term "Sökträffar" and various attributes: "Textattribut", "Strukturella attribut", "Ordattribut", "compwf: mord+brand", "pos: NN", "lemma: mordbrand", "lemgram: mordbrand..nn.1", and "prefix: mord..nn.1".

Illustration 1: Resultatvy

I en termsökning på “mordbrand” kan vi i, Illustration 1, se resultaten från Strix’ resultatvy för markerad träff med dess olika komponenter. I detta läge ser vi ett dokument från samlingen “8 sidor” med termen mordbrand markerad i texten på rad 10 av dokumentet. I anslutning ser vi:

- till vänster “Relaterade dokument”,

- till höger attributen på de olika nivåerna “text-, strukturellt och ordnivå”, samt “sökträfflistan”.

En termsökning om inget annat anges är en “lemgram-sökning” där alla ordformer i ett igenkänt böjningsmönster (paradigm) matchar. Detta gäller både om sökningen är en “fras-sökning” det vill säga flera termer i följd, skönsvärdet för bockrutan “i följd”, eller om det är någon av termerna i valfri ordning “nyckelordssökning”.

Notera att nu står det “sök i dokumentet” bredvid termfältet vilket gör att vi kan borra oss ner med hjälp av ytterligare filtreringar på till exempel ordattribut i dokumentet, så kallad “facettsökning”. Vill vi avföra detta dokument kan det enkelt klickas bort med krysset bredvid samling/dokumenttiteln “8 SIDOR: Skola i Gråbo brann ner”.

- 1.1** Gör en bred termsökning på termen “nazister” i “8 SIDOR”. Du får ca 126 dokument som resultat. Som du märker får du även träffar på andra ordformer som “nazist” osv. Vad kallas denna typ av termsökning?
- 1.2** Välj ett Blingbring-attribut, de är sorterade i fallande ordning. Välj ett som ger dej ett dokument kvar som resultat.
- 1.3** Välj SweFN-attributet från “visa fler filter”. Vilken ram har högsta värdet?

ÖVNING 2 - ATTRIBUT- / FACETTSÖKNING

Även utan termer i sökfiltret kan du söka på de olika attributen på samma sätt som ovan med facettsökningen i ett första resultat. Välj till exempel “Nordiska museets frågelistor”. Attributen är i bland ordnade efter “förekomst”, antalet inom parenteser, eller för till exempel tidsattribut “tidsordning” eller i “bokstavsordning”. Förutom att du kan välja olika typer av attribut kan du även välja flera attribut av samma typ till exempel om vi tar både “Nyodling” och “Skogsbruk” under attributet “Ämne” markeras båda dessa värden i listan. Då blir resultatet uppdaterat. Vill du utesluta ett valt värde klickar du på det igen och ser att resultatvyn direkt uppdateras för att reflektera de valda attributens värden.

- 2.1** Gör en facettsökning på samling “frågelistor”. Vilket år är attributet “Ämne” med värdet “Fotogenlampan” utskickad?

ÖVNING 3 - INTERAGERA MED VALDA SÖKRESULTAT

Om du väljer ett sökresultat i resultatvyn kan du interagera med detta. Alla ord kan klickas på och du får då upp ordattributen för dem. Du kan fortsätta filtrera med hjälp av attributen tillgängliga i rullgardinsmenyn under termfältet som nu benämns “sök i dokumentet” för att matcha (andra) termer.

3.1 Gör en bred termsökning och borra vidare med hjälp av “sök i dokumentet”. Välj ett av varje typ av attribut som finns tillgängligt av text-, ord-, och strukturella attribut.