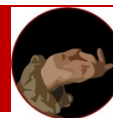


Lathund till Transkribus



Inledning

Vad är Transkribus?

Transkribus är en digital plattform som använder artificiell intelligens (AI) för att transkribera texter, d.v.s. överföra dem från en skriven form till en annan, till exempel från ett handskrivet brev till en digital text. Transkriptionen kan ha flera syften, men ett övergripande ändamål är att göra källtexten mer tillgänglig. Exempelvis går det snabbare att läsa en svårläst handskrift som transkription. När transkriptionen går hand i hand med digitaliseringen av ett dokument, tillkommer dessutom andra fördelar: det som tidigare bara var inskannade *bilder* av ett dokument, blir plötsligt *texter*, där orden kan sökas, taggas, bearbetas och analyseras med hjälp av olika digitala verktyg.

Hur fungerar det?

I korthet använder sig Transkribus av olika ”modeller” för att läsa olika typer av text, till exempel olika skrivstilar på olika språk. För att fungera, måste dessa modeller tränas. Man utgår då från en transkription av ett urval texter från det material man vill transkribera, så att maskinen någorlunda lär sig att läsa det. Den kan då automatiskt transkribera ett större urval från samma material. Antagligen kommer denna transkription fortfarande att innehålla mängder med fel; transkriptionen måste därför manuellt korrigeras och den nya datan matas in — maskinen lär sig på så vis av sina misstag och blir stegvis allt bättre på att återge texten korrekt. Den manuella korrekturen är därför inte bara ett oundgängligt steg i modellträningen, utan också för att få fram korrekt transkriberade texter.

Den här lathunden avser att ge en kort introduktion till hur man arbetar med att korrigera texter i Transkribus.

Få tillgång till samlingen

För att få tillgång till materialet vi kommer att arbeta med, behöver du **i förväg registrera ett användarkonto**.

- Gå till <https://readcoop.eu>.
- Klicka på *Sign in* längst upp till höger.
- Klicka på *Register* och fyll i dina användaruppgifter (namn, e-post och lösenord).

När du fått ett konto, behöver du **maila oss e-postadressen** du använde för att registrera sig. Vi ger dig sedan tillgång till samlingen.

Öppna Transkribus

Transkribus finns i en version för webbläsare (tidigare kallad *Transkribus Lite*) och som skrivbordsklient (*Expert Client*). Skrivbordsklienten uppdateras dock inte längre, och det är tänkt att webbläsarversionen skall få alla funktioner som idag finns i skrivbordsklienten. Det är webbläsarversionen vi använder oss av.

Öppna Transkribus i din webbläsare:

- Gå till <https://transkribus.eu/lite/>.
- Logga in med användaruppgifterna till ditt konto.

Arbeta med ett dokument

Vi arbetar endast med att korrigera *texten*. Där det finns fel i raderna eller deras ordning, korrigeras det av projektdeltagarna i efterhand.

- När du fått ett dokument att arbeta med, går du in i *Collections* i Transkribus. Välj samlingen "Gustavs Hand – Svenska dokument" och leta dig fram till ditt dokument. Du öppnar det genom att helt enkelt klicka på ikonen.
- Du kan börja arbeta med dokumentet genom att klicka på den sida du vill öppna. En bild av originalet visas till vänster, transkriptionen till höger.
- Korrigera texten genom att helt enkelt skriva in det korrekta värdet i transkriptionen. Använd normdokumentet för att se hur du gör med förkortningar med mera.
- Du byter sida m.h.a. pilarna längst ned på sidan (**Bild, nr 3**).
- När en sida eller ett dokument är klart trycker du på *Save* (diskettikonen högst upp till höger)! När hela dokumentet är färdigt, kan du gå tillbaka till samlingen (namnet står i menyn högst upp till vänster) och börja redigera ett nytt dokument.

Specialfunktioner

- Till vänster i dokumentvyn finns verktyg för att redigera rader och radordning (**Bild, nr 1**). Dessa verktyg behöver vi inte bry oss om, då vi endast redigerar texten.
- Till höger i samma vy finns verktyg för zoomning med mera (**Bild, nr 2**). Håll pekaren över ikonerna för att se vad de gör. Om du måste vrida en bild, kom ihåg att vrida tillbaka den till ursprungsläget innan du sparar!
- *Virtual keyboard* (**Bild, nr 4**) används för specialtecken. Här har projektdeltagarna lagt in ofta förekommande tecken. Klicka på dessa för att använda dem i texten. Om fler tecken behövs, kan projektdeltagarna lägga in dem.
- I inställningarna (**Bild, nr 5**) ändrar du t.ex. hur saker visas. Här kan du bl.a. välja att visa radnumren och hur stora siffrorna på raderna då skall vara. Det mesta här kan du dock låta vara som det är.

The screenshot shows the Transkribus web interface. On the left is a thumbnail of a handwritten document. On the right is the transcription of the document. The interface includes a top navigation bar, a left sidebar with various tools, a main viewing area, and a bottom toolbar.

Red boxes and numbers highlight the following elements:

- 1**: Points to the top-left corner of the document thumbnail.
- 2**: Points to the top-right corner of the document thumbnail.
- 3**: Points to the bottom-left corner of the document thumbnail.
- 4**: Points to the bottom-right corner of the document thumbnail.
- 5**: Points to the bottom-right corner of the document thumbnail.

The transcription on the right is as follows:

REGION 1

Fast jag har skrivit för några dagar sedan
igenom
lilla
Bror så villia idag för sista gången
skrifva i från Spa som jag lämnar vid ganska
gått hymeur och ännu bättre hälsa i morgon
jag har väl medfått igenom deBeche som
jag eij väntade men eij gienom Clas Bruse
som jag onskade edert Bref jag tallar eij här
om denna besynnerliga skickning som skiet
tvert emott ofver tio reitererade ordres jag
har där om redan skrivit med lilla Bror och
vil där om eij mera tala förän min hemkomst
jag undfår nu edert bref af d. 5 sept: uti vilket
ni berättar Gref Wrangels hastiga död och
Troils val til Lindköpings stift jag vet doch eij
vem som står mera på förslaget men jag gör
endast den anmärkning at för 78 Ricksdag
alla Biscops val föllo på obehageliga Personer och
efter Ricksdagen hava de alla utfallit efter
önskan. Cronans Consideration har derfore