



# Spørgsmål om ophavsret – den islandske erfaring

Sigrún Helgadóttir

Árni Magnússon instituttet for islandske studier

Leksikografisk afdeling

Spåkteknologisk infrastruktur

Nordiskt seminarium vid Wallenberg center i Göteborg

26. oktober 2006

Nordens språkråd



- Oversigt
  - gå i gennem lov om ophavsret
  - 5 forskellige korpusser
  - oversigt over det som man allerede har gjort med hensyn til at sikre licens til brug af tekster i det islandske korpus.



# BLARK (Basic Language Resource Kit)

## Tekstkorpusser

- Et tekstkorpus er en struktureret samling af tekstprøver i elektronisk form der dokumenterer et sprog sådan som det skrives i en bestemt tid.
- Adgang til tekster i elektronisk form
- kan ikke lægge ophavsretsværnet materiale ud på internettet uden ophavsmandens samtykke.
- Metoder til at fremskaffe licens fra indehavere af ophavsretten

.....



# Lov om ophavsret i de nordiske lande

(dansk lov brugt som eksempel)

## Ophavsret

- § 1. Den, som frembringer et litterært eller kunstnerisk værk, *har ophavsret til værket*, hvad enten dette fremtræder som en i skrift eller tale udtrykt *skønlitterær eller faglitterær fremstilling*, som ...
- § 2. Ophavsretten medfører, med de i denne lov angivne indskrænkninger, *eneret* til at råde over værket ved at fremstille eksemplarer af det og ved at gøre det tilgængeligt for almenheden i oprindelig eller ændret skikkelse, i..



# Lov om ophavsret i de nordiske lande

(dansk lov brugt som eksempel)

Hvem er ophavsmanden?

- § 7. Som *ophavsmand* anses, når ikke andet er oplyst, den, hvis navn eller alment kendte pseudonym eller mærke på sædvanlig måde er påført eksemplarer af værket eller opgives, når det gøres tilgængeligt for almenheden.
- *Stk. 2.* Er et værk udgivet, uden at ophavsmanden er angivet i overensstemmelse med stk. 1, kan udgiveren, hvis denne er nævnt, og ellers forlæggeren handle på ophavsmandens vegne, indtil denne bliver angivet på et nyt oplag.



# Lov om ophavsret i de nordiske lande

(dansk lov brugt som eksempel)

## Andre paragraffer som har betydning

### Oversættelser

- § 4. Den, som *oversætter*, omarbejder eller på anden måde bearbejder et værk, herunder overfører det til en anden litteratur- eller kunstart, har *ophavsret* til værket i denne skikkelse, men kan ikke råde over det på en måde, som strider mod ophavsretten til det oprindelige værk.

### Offentlige aktstykker

- § 9. Love, administrative forskrifter, retsafgørelser og lignende offentlige aktstykker er ikke genstand for ophavsret.



- Rettighedshaverorganisationer som tager sig af aftalelicenser om fotokopiering.
  - Bonus Presskopia
  - COPY-DAN
  - Fjölís
  - Fjølrit
  - Kopinor
  - Kopiosto
  - Samikopiija



- **Korpusser i forskjellige lande**
- **Hvordan man har fået licens til at bruge tekster i korpusser fra ophavsrettens indehavere**
  - BNC (British National Corpus) (engelsk)
  - ANC (American National Corpus) (amerikansk engelsk)
  - Korpus 2000 (dansk)
  - Oslo korpuset av taggedede norske tekster (norsk)
  - SUC (Stockholm Umeå Corpus) (svensk).





- BNC (British National Corpus)
  - 3000 tekster – 100 millioner ord.
  - skrive til ophavsrettens indehavere
  - aldrig hele tekster som er værnet af ophavsret
  - søgbar i konkordansform på projektets hjemmeside
  - købe brugerlicens og få hele korpuset på disketter eller købe „subscription service” og få adgang gennem internettet
  - man betaler kun behandlingsomkostninger



- ANC (Americ National Corpus)
  - under opbygning – anden udgave har nu 22 millioner ord
  - lægge ind tekster gennem projektets webside (upload)
  - ophavsmanden sender e-post til projektlederen med standardiseret tekst som giver projektet licens til at bruge teksten „for the purposes of linguistic education, research, and development“.



- **Korpus 2000 (dansk)**

- 28 millioner ord fra ca. 110.000 forskellige tekster skrevet i perioden 1998-2002
- Korpus 2000 er et citatkorpus som defineres på websiden således:
  - „Et **citatkorpus** er et tekstkorpus, som først er splittet op i enkelte sætninger, som herefter er blandet i tilfældig rækkefølge. Det indeholder altså præcist det samme sproglige materiale, som det oprindelige tekstkorpus, men sætningerne kommer blot i vilkårlig rækkefølge, så det ikke længere er muligt at rekonstruere de oprindelige tekster.“
- Denne fremgangsmåde er nødvendig af ophavsretlige grunde



- Oslo korpuset av taggedede norske tekster
  - bokmålsdelen, indeholder omtrent 18,5 millioner ord
  - nynorskdelen indeholder omtrent 3,8 millioner ord
  - teksten til korpuset blev hentet fra tekster som allerede var tilgængelige for internal brug inden for universitetet i Oslo
  - når korpuset skulle blive åbnet for søgning på webben har man skrivet til alle institutter som havde bidraget med tekster
  - i brevet har man sagt at hvis man ikke fik svar skulle det betragtes som samtykke
  - adgang er med brugernavn og password



- SUC (Stockholm Umeå Corpus)
  - udviklet 1990–1996
  - første version indeholder 500 filer med omtrent 2065 ord hver fil
  - version 2 er søgbar på internettet gratis i konkordansform
  - der findes ikke ret meget skrevet om hvordan licens for brug af tekster i SUC er blevet fremskaffet undtagen at „legal agreements“ har været optegnet og at man måtte udelade nogle tekster på grund af at man ikke fik licens til at bruge dem.



- Islandsk korpus
  - 25 millioner ord, hvert ord tagget med morfosyntaktisk tag og lemma
  - tekster skrevet 2000 og senere
  - tekster hentet fra Ordbogens tekstsamling
  - komplettere Ordbogens tekster med tekster fra forskellige genrer
  - søge licens til brug af tekster fra rettighedshavere



- Hvordan søger man licens til at bruge en tekst?
  - Alle tekster er værnet af opahavsret uanset man ved hvem ophavsmanden er
  - Blogs – begrænse valget til blogs hvor ophavsmand er kendt, sende e-post
  - Blogs – man giver licens hvis det ikke koster for megen umage
  - Postlister – abonnere med samtykke fra postlistens webmaster
  - Taler som findes i tekstform på internet – sende e-post
  - Webpladser – offentlige institutter og firmaer har ophavsret til tekst som offentliggøres på deres webplads uden forfatterens navn



- Hvordan søger man licens til at bruge en tekst?
  - Nyheder i tv og radio –chefredaktøren for den statslige radio- og tv-stations nyhedsbureau har fået tilladelse fra reporterene til at aflevere teksterne til brug i ordbogens tekstsamlinger
  - Aviser – lave kontrakter med avisernes redaktioner, redaktionerne har i de fleste tilfælde sikret ophavsret til alle tekster som publiceres i aviserne
  - Bøger og tidsskrifter – sende forfatteren brev eller e-post





- **Konklusion I**

- Rettighedshavere er positive over for selve projektet
- Man er imod den idé at give offentlige institutter licens til at bruge tekst uden betaling uanset værdien af de enkelte projekter imens det offentlige ikke vil lave en aftale om generel brug af digitaliserede tekster.



## • **Konklusion II**

- Rettighedshavere behøver detaljerede og klare oplysninger om projektet
  - hvordan skal teksten bruges, hvordan er adgang til tekstsamlingen og korpuset
  - man inkluderer aldrig hele ophavsretligt værnede tekster i korpuset
  - tekster kan ikke rekonstrueres fra korpuset.
- Det må være enkelt at sende erklæringen om samtykke
  - per e-post, fax, adresseret og frankeret konvolut