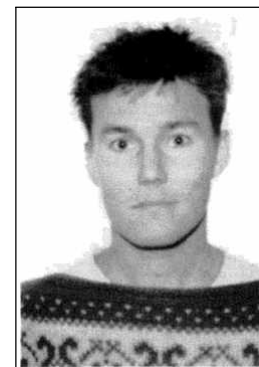


Lag dine egne søkeskjermer til biblioteksdatabasene på World Wide Web



Av Tore Brattli, Universitetsbiblioteket i Tromsø

Stadig flere biblioteksdatabaser blir nå tilgjengelige på World Wide Web (WWW) med de mulighetene det gir til brukervennlig litteratursøking. En av disse nye mulighetene er at de litteratursøkende nå kan lage sine egne brukervennlige og skreddersydde søkeskjermer, uten at de trenger å være dataeksperter.

Det er i hovedsak to typer søkeskjermer som det åpner seg muligheter for. Den første er de tradisjonelle søkeskermene med et eller flere felt som søket kan skrives inn i. Det er nå praktisk mulig å lage sine egne varianter av denne typen søkeskjermer, både når det gjelder hvilke felt som skal være med og hvilket søkespråk som skal benyttes. Den andre typen er lenking av søk til hypertekst, eller det som også kalles predefinerte søk. Dette åpner for spennende muligheter som f.eks. søking via emneordslister eller klassifikasjonstabeller, og bibliografier der hvert dokument er lenket til sin post i databasen.

Selv om artikkelen i stor grad omhandler biblioteksdatabasen BIBSYS, er eksemplene likevel med små modifikasjoner overførbar til BIBLIOFIL-bibliotekenes databaser eller andre databaser på WWW.

Et tilbakeblikk

BIBSYS, som startet opp i 1972, er et felles biblioteks-system for alle universitetene, nasjonalbiblioteket og en rekke høyskoler i Norge. WWW er et distribuert informasjonssystem som kom i gang rundt 1990 for at forskerne i miljøet rundt partikkelfysikklaboratoriet CERN lettere skulle kunne dele informasjon. Systemet har senere spredd seg med imponerende fart ut til hele verden, spesielt etter at NCSA lanserte Mosaic, den første WWW-leseren med grafisk brukergrensesnitt, i januar 1993. I november 1993 ble BIBSYS tilgjengelig på WWW etter at Edb-sentret ved Universitetet i Tromsø lagde en søkeskerm som kunne kommunisere med BIBSYS sin Gopher-tjener. BIBSYS selv lanserte søking på WWW i april 94. I dag, et år senere mottar BIBSYS ca. 60.000 søk per måned, eller et sted mellom 20 og 50 prosent av den totale eksterne søkemengden fra WWW. I mars 1995 fikk Asker bibliotek sin biblioteksdatabase tilgjengelig på

WWW som det første av folkebibliotekene i Norge. Etter det har flere kommet til, også av fylkesbibliotekene.

Hva trengs for å lage egne grensesnitt?

Både verktøyet og kunnskapene som trengs for å lage sine egne søkeskjermer burde være overkommelige å skaffe seg for de som allerede har tilgang på WWW.

Til enkle lokale søkeskjermer trengs det en teksteditor og litt kunnskap i HTML (HyperText Markup Language). HTML er språket som beskriver sidene i WWW. Skal søkeskermene bli tilgjengelig for resten av verden, må de legges ut på en WWW-tjener. For å lage mer avanserte søkeskjermer må det litt programmering til. Programmeringsspråk som Perl eller C er godt egnet til dette. Til emneordslister, klassifikasjonstabeller og bibliografier er det i tillegg praktisk med et PC-databaseprogram. Det meste av programvaren som trengs er enten gratis eller nesten gratis.

Overføring av søket til databasen

Et søk i en database på WWW overføres vanligvis som en del av URL-en (Uniform Resource Locator). En URL adresserer ressurser på WWW, og kan sees på som en utvidet utgave av DOS/UNIX metoden for å adressere filer.

Fordelen med å ha litt kjennskap til URL-ens oppbygging er at det gjør oss uavhengige av spesielle søkeskjermer. Vi kan lage egne søkeskjermer, søke i andre felt enn det som databasen legger opp til, eller søke direkte via URL-en uten å gå veien om søkeskjermer. Det gir flere muligheter for søking, i tillegg til at søkingen lettere kan tilpasses lokale behov. URL-ene kan stort sett deles inn i to typer. Den første typen, som brukes til å adressere filer, er vanligvis på formen:

```
http://www-tjener/katalog/filnavn.html
```

der fila `filnavn.html`, som ligger i katalogen `katalog` på WWW-tjeneren `www-tjener`, benytter overføringsprotokollen `http` når den skal sendes over nettet. Den andre typen, som brukes til å kjøre programmer, er ofte på formen:

`http://www-tjener/katalog/program?
parameter=verdi`

der programmet program på WWW-tjeneren `www-tjener` mottar parameter `parameter` med verdi `verdi` og benytter overføringsprotokollen `http` for å returnere eventuelle resultater. Det er parameterdelen til URL-en som brukes for å overføre søket til databasen.

Koder for spesialtegn

Når søkestrengen overføres til databasen som en del av URL-en, sendes noen av tegnene over i kodet form. Her følger en oversikt over noen av disse kodene:

Tegn	URL-kode	Tegn	URL-kode
æ	%E6	/	%2F
ø	%F8	?	%3F
å	%E5	+	%2B
(%28	,	%2C
)	%29	space	+

Det er lett å oppfatte URL-ene som vanskeligere enn det de egentlig er på grunn av kodene.

Søkespråket i pubsøk

I BIBSYS er URL-en til et pubsøk på formen:

```
http://search.bibsys.no:5001/pub.action?  
type=søketype&bool=logikk&lang=språk&  
base=database&fo=forfatter&bd=basisdel&  
de=dewey&sn=ISSN
```

URL-en kan deles opp i tre deler:

1. `http://search.bibsys.no:5001/pub.action?`
2. `type=søketype&bool=logikk&lang=språk&
base=database&`
3. `fo=forfatter&bd=basisdel&de=dewey&
sn=ISSN`

der parameterdelene (1), (2) og (3) overføres til BIBSYS sin WWW-tjener for søk `search.bibsys.no:5001`.

Parameterdel (1) forteller BIBSYS at søket skal tolkes som et pubsøk, (2) spesifiserer søketype (`search`, `scan`), logikk (`and`, `or`), språk (Norsk, Engelsk) og database (BIBSYS, UBTØ, ...), mens (3) spesifiserer søkbare felt som `forfatter`, `basisdel`, `Dewey-klassifikasjon`, osv.

I parameterdel (3) kan det legges inn søking i andre felt enn det som BIBSYS legger opp til, for eksempel `UDK-klassifikasjon`, `trykkeår` eller `serie`.

Eksempel: Søk i BIBSYS etter `ibsen,h?` som forfatter og `vildand?` som fritekst gir URL-en:

<http://search.bibsys.no:5001/pub.action?type=search&bool=and&lang=N&base=BIBSYS&fo=ibsen%2Ch%3F&bd=vildand%3F>

Søkespråket i gensøk

I BIBSYS gensøk er URL-en til et søk på formen:

1. `http://search.bibsys.no:5001/gen.action?`
2. `lang=språk&`
3. `st=_database_søkestreng`

der parameterdel (1) forteller BIBSYS at søket skal tolkes som et gensøk, (2) spesifiserer språk mens (3) inneholder eventuell spesifisering av database i tillegg til selve søkestrengen. Søkestrengen skrives i vanlig gensøkspråk. Den eneste forskjellen er at kommandoen `finn` ikke benyttes, og at noen av spesialtegnene skrives på kodet form i URL-en.

Eksempel: Søk i database UBTØ etter Henrik Ibsens skuespill `Vildanden`. I vanlig gensøk kan søket formuleres som:

```
ibsen,h?/fo*(vildanden+villanden)/bd
```

Når søket skal sendes som en del av URL-en får vi:

http://search.bibsys.no:5001/gen.action?lang=N&st=UBT%D8_ibsen%2Ch%3F%2Ffo%28vildanden%2Bvillanden%29%2Fbd

Adressering av poster

For å hente poster i BIBSYS brukes objektid som adressering. URL-en er på formen:

```
http://search.bibsys.no:5001/  
Språk.TypeObjektid.html
```

der parameterdelen spesifiserer hvilken type informasjon som skal returneres og hvilket språk informasjonen skal presenteres på. Vanlige koder for `type` er `b` (bibliografisk), `k` (klassifikasjon), `e` (eksemplar) og `m` (marcformat).

Eksempel: Bibliografisk informasjon fra posten med objektid `930618416` skrevet ut på norsk gir URL-en:

<http://search.bibsys.no:5001/N.b930618416.html>

Tradisjonelle søkeskjermer

Den enkleste måten å lage nye søkeskjermer til en biblioteksdatabase på er å ta utgangspunkt i en eksisterende søkeskerm. Fjern først det som er unødvendig og legg deretter til det som mangler. Ved å lagre endringene er den nye søkeskjermeren ferdig og klar til bruk.

Eksempel: Lag en enkel BIBSYS søkeskerm tilpasset UDK-klassifikasjon:

1. Last BIBSYS-søkeskjermeren inn i en WWW-leser. (URL: <http://www.bibsys.no/search/pubn.html>)
2. Få WWW-leseren til å vise HTML-koden som beskriver søkeskjermeren.
3. Kopier koden over i en enkel teksteditor (f.eks. Notepad i Windows)

4. Fjern alt bortsett fra:

```
<HTML>
<TITLE>Søk BIBSYS</TITLE>
<FORM METHOD=GET ACTION=
"http://search.bibsys.no:5001/pub.action">
<PRE>
Forfatter: <input name="fo" size=60>
Tittel   : <input name="ti" size=60>
Fritekst : <input name="bd" size=60>
Dewey    : <input name="de" size=60>      (1)
<input type="submit" value="Send">
</PRE></FORM>
</HTML>
```

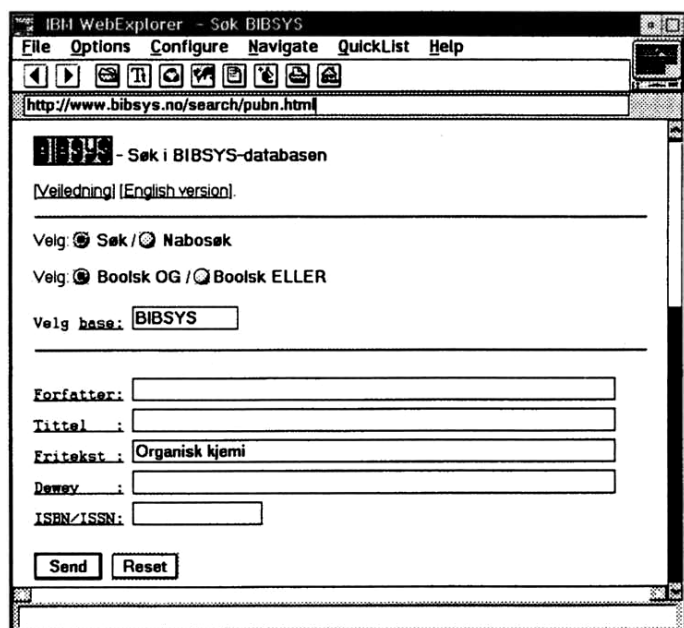
5. Endre linje (1) til

```
UDK-klass: <input name="ud" size=60>
```

6. Den nye søkeskjermen kan nå lagres som f.eks.:
C:\www\bibsys.htm

7. Søkeskjermen er nå ferdig og klar til bruk ved å laste inn WWW-siden med URL:

<file:///C:/www/bibsys.htm>



Eksempel fra tradisjonelle søkeskjerner – BIBSYS søkeskjerm

På denne måten er det mulig å lage søkeskjermene slik vi selv vil ha dem, så lenge vi holder seg innenfor søkespråket til BIBSYS. Endringer i søkespråket kan gjøres ved å bruke CGI-skript. Et CGI-skript er et dataprogram som kan kommunisere med WWW via CGI (Common Gateway Interface). CGI-skript blir vanligvis skrevet i programmeringsspråkene Perl eller C, der Perl absolutt er å anbefale for de som ikke er så stødige i programmering.

Ved å bruke CGI-skript er det mulig å lage søkeskjerner som kan ta imot andre søkespråk enn det som BIBSYS benytter. En slik søkeskjerm må først sende søket til et CGI-skript for oversetting til BIBSYS-språk, før det kan sendes til BIBSYS. BIBSYS utfører søket og returnerer resultatet til WWW-leseren som sendte søket. Det er også mulig å få BIBSYS til å sende resultatet til et CGI-skript for bearbeiding, før det sendes videre til WWW-leseren.

Predefinerte søk

Med et predefinert søk menes et søk der søkestrengen ligger ferdig lagret som en del av URL-en, slik at søket kan utføres ved å klikke på tilhørende hypertekst. Predefinerte søk egner seg bra for søk som utføres mange ganger, eller når mengden av hva som er mulig eller interessant å søke etter er begrenset.

Eksempel: En bibliotekar i Tromsø med ansvar for Ex.phil. (Examen philosophicum) får ofte spørsmål etter bøker om førsokratisk filosofi i forbindelse med semesteroppgavene. Et slikt søk kan legges inn på ei WWW-side som et predefinert "klikke" søk i BIBSYS. Først utføres søket

```
førsokrat? thales anaximander pythagoras
parmenides anaxagoras
```

i WWW-pubsøk med base UBTØ og logisk ELLER. Kopier trefflistens URL

<http://search.bibsys.no:5001/pub.action?type=search&bool=or&lang=N&base=UBT%D8&fo=&ti=&bd=f%F8rsokrat%3F+thales+anaximander+pythagoras+parmenides+anaxagoras&de=&sn=>

fra Web-leseren over til en teksteditor. Legg nødvendig HTML-kode og forklarende tekst til URL-en slik at vi får

```
<HTML>
<A HREF = "URL">Bøker om førsokratisk filosofi
i Tromsø</A>
</HTML>
```

Ved å lagre dette som ei WWW-side kan den brukes hver gang vi trenger å søke etter førsokratisk filosofi i BIBSYS.

Litteratursøking ved å bruke emneordslister

Det er nå mulig å lage sine egne systemer for litteratursøking i biblioteksdata-baser på WWW med utgangspunkt i emneordslister. Ved å bruke mulighetene som hypertekst i WWW gir, er det stort sett bare fantasien som setter grenser for hvordan emneordslister kan presenteres for brukerne. Alfabetisk, alfabetisk permutert eller hierarkisk over et eller flere nivåer er noen av mulighetene. Se-henvisninger, definisjoner og klassifikasjoner kan også integreres i systemet. Hyperteksten gjør at brukerne kan "vandre" rundt i emneordshierarkiet på

en enkel og oversiktlig måte. Ved å lenke emneordene til søk i databasen, kan systemet bli et meget bra alternativ til de tradisjonelle søkeskjermene.

I praksis kan et slikt system lages ved å legge inn emneordlista i en PC-database, der hver post består av et emneord med tilhørende definisjoner, se-henvisninger og en kode som beskriver tilhørigheten til de andre emneordene.

Ved å eksportere databasen (emneordlista) som ei tekstfil til WWW-tjeneren, kan et CGI-skript generere både den hierarkiske strukturen og lenkingen mot biblioteksdata-basen ut fra tekstfila.

BIBSYS vil sannsynligvis komme med HUMORD-emneord på WWW i løpet av sommeren 1995. NTUB-/MESH-emneord vil sikkert også bli tilgjengelig med tiden. Så for de som er utålmodige, eller de som bruker andre emneordslister, er det bare å sette i gang på egen hånd.

Litteratursøking ved å bruke klassifikasjonstabeller

Bruk av klassifikasjon i litteratursøk har tradisjonelt vært forbeholdt de få. Klassifikasjonstabellen er sjelden tilgjengelig for biblioteksbrukere, og selv om den skulle være det, er det de færreste som har tålmodighet til å sette seg inn i bruken av den. Problemet med tabeller i bokform er at både oversikten og strukturen forsvinner når hierarkiske strukturer skal presenteres i en dimensjon. Ved å legge ut klassifikasjonstabellen som hypertekst på WWW, og lenke hver klasse til søk i BIBSYS, blir søking via klassifikasjon tilgjengelig for folk flest.

Klassifikasjonstabeller på WWW kan både lages, presenteres og brukes på stort sett samme måte som for emneordslister.

Eksempel: Mathematics Subject Classification (MSC). <http://www.ub.uit.no/cgi-bin/msc.pl>

I denne hypertekstutgaven av MSC blir vanligvis en klasse skrevet ut som overskrift, etterfulgt av ei liste med alle underklasser på nivået under. Det er mulig å bevege seg nedover i tabellen ved å klikke på en av klassifikasjonene i lista, mens det motsatte skjer ved å klikke på klassifikasjonen i overskrifta. Søk i BIBSYS gjøres ved å klikke på deskriptorene. Et klikk på en deskriptor i lista utfører eksakt søk på tilhørende klassifikasjon, mens et klikk på deskriptoren i overskrifta gir trunkert søking. På den måten ivaretas søking både med og uten trunkering.

Bibliografier med bruk av BIBSYS

BIBSYS kombinert med WWW kan med fordel brukes til å lage bibliografier over bøker registrert i BIBSYS. Noen av problemene med bibliografier på papir er registrering, oppdatering og distribusjon. Ved å bruke WWW er det nok å registrere tittel og objektid. Tittelen lenkes til sin post i BIBSYS ved hjelp av objektid. Den bibliografiske informasjonen tar BIBSYS seg av, både når det gjelder registrering og oppdatering. Distribusjonen tar WWW seg av, gratis og til hele verden.

Lista over titler og objektid må vi selv oppdatere, og dette gjøres greiest ved å legge disse inn i en helt vanlig PC-database. Etter oppdateringer i PC-databasen kan titlene med objektid eksporteres som HTML kode til WWW-tjeneren og legges ut som ei WWW-side.

En litt mer avansert variant er å eksportere innholdet i PC-databasen som ei tekstfil til WWW-tjeneren. Ved å skrive et CGI-skript som kan søke i tekstfila har vi lagd vår egen søkbare database på WWW.

Eksempel: Nye bøker i informatikk ved UBTØ de siste 3 måneder. Last inn WWW-siden:

<http://www.ub.uit.no/cgi-bin/nybok.pl?emne=Informatikk&fra=-1&til=3>

På denne WWW-siden ligger ei informatikkbok helt fra den ble satt i bestilling til det er mer enn tre måneder siden den ble klar til utlån. Det gir brukerne til enhver tid full oversikt over hvilke bøker som er i bestilling, til registrering i biblioteket og klar til utlån. Kombinert med muligheten for bestilling av lån i BIBSYS blir dette en kraftig tjeneste.

Litteratur til videre lesing

L Lemay, Teach yourself Web publishing with HTML in a week (Sams publ., 1995)

R Schwartz, Learning Perl (O'Reilly & Associates, 1993)

B Schatz og J Hardin, NCSA Mosaic and the World Wide Web: Global Hypermedia Protocols for the Internet, Science 265 (12 Aug 1994)

Tore Brattli er universitetsbibliotekar på Universitetsbiblioteket i Tromsø med ansvar for fagene matematikk, statistikk, informatikk og kjemi

Adresse: Universitetsbiblioteket i Tromsø, avdeling for realfag, medisin og helsefag, 9037 Tromsø

Telefon: ++47 77 64 54 68 (kl. 8-15)

Telefax: ++47 77 64 54 75

E-mail: tore@ub.uit.no