

"Sæt-niogtredive-hvafter-en-fisk?"

Hvorfor skal vi have Z39.50, og hvorfor vedkommer det mig?

Alle taler om den

En hel del har den

Nogen gør noget ved den

Kun få forholder sig til den

Der bliver talt en hel del om Z39.50 og dens velsignelser og forbandelser i disse år. Efterhånden begynder man også at se resultater af al den snak og planlægning.

Desværre er det næsten udelukkende blandt systemfolk hos systemleverandørerne og i bibliotekernes it-afdelinger, diskussionerne og implementeringen foregår.

Mit formål med dette indlæg er at gøre "almindelige" bibliotekarer opmærksomme på, at det i høj grad er noget, man bør forholde sig til og blande sig i. Det er nemlig de fremtidige søgemuligheder, der fastlægges her.

Hvad er Z39.50?

Ifølge Biblioteksstyrelsen er "Z39.50 (= ISO 23950) en standard for netværkssammenkobling af informationssystemer af den type, der anvendes i biblioteker og informationscentre".

Altså er det en protokol, der gør en *klient* (f.eks. en Aleph500-søgeklient, som bibliotekaren benytter til at komme i kontakt med diverse baser) og et *target* (f.eks. DanBib-basen på en DBC-server som modtager søgninger fra diverse klienter) i stand til at tale sammen. Protokollen består af mange elementer, hvoraf jeg kun vil omtale nogle enkelte.

Strukturen er en maskinstruktur, hvorpå der kan bygges "pæne grænseflader" (i f.eks. HTML til en WEB-side).

Det betyder, at man meget sjældent ser selve "Z39.50-trafikken", da der normalt vil være lagt en grafisk brugergrænseflade ovenpå, som man som slutbruger benytter til at udfylde diverse søgekasser og udvælge definerede rullegardinpunkter.

Når man så trykker på send/find-knappen, vil det, man har udfyldt, blive efterbehandlet af lokalsystemets søgeklient og normalt blive omdannet til en søgesyntaks, som Z39.50 er god til at håndtere, nemlig den såkaldte *RPN-syntaks* (Reverse Polish Notation).

Det udfyldte vil så blive sendt til target (den ønskede base) sammen med mange andre Z39.50-elementer, der indeholder f. eks. formatønske, om det er søgning eller skanning og mange andre tekniske parametre.

Det kan godt lade sig gøre at sende en CCL-syntaks, men den er vanskeligere at tolke korrekt i et Z39.50-miljø.

Jeg vil senere komme nærmere ind på, hvorfor dette er nyttigt at vide.

Hvem fastlægger den danske anvendelse?

Z39.50-protokollen er meget åben og har mange fortolkningsmuligheder. For at sikre at danske Z39.50-miljøer bestående af klienter og targets kan tale ordentligt sammen, har Biblioteksstyrelsen nedsat en gruppe, der hedder danZIG-gruppen. Denne gruppe beslutter, hvordan vedtagelser i det internationale forum ZIG implementeres i danske Z39.50-systemer.

ZIG har fastlagt nogle internationale retningslinier i et dokument, der hedder "Bath-profilen".

De danske retningslinier er fastlagt i "danZIG-profilen", som desuden indeholder en oversigt over, hvordan man tolker (eller i fagsprog – mapper) Biblioteksstyrelsens "Praksisregler for søgeveje", så man får dannet de ovenfor nævnte RPN-streng på en sådan måde, at target (den fremmede base) forstår dem og opfatter dem korrekt.

Hvad består en Z39.50-forbindelse af?

Z39.50-strukturen ser sådan ud:

Klient Port på target

Når klienten sender en brugers søgning oversat til en RPN-streng afsted, modtages søgningen på targets server af den port, som er angivet i klientens "ordre" (normalt port 210).

Porten indeholder "pegepinde" til én eller flere baser (såkaldte backends).

En "pegepind" indeholder definitioner, som blandt andet består af oplysninger om tilladte formater og hvilke af basens registre, der rammes af hvilke RPN-streng.

Hvis klientens RPN-streng rammer en gyldig "attribut-kombination" (se nedenfor), sender porten søgningen videre til den base, klienten har ønsket i sin forespørgsel.

Basens "maskiner" udfører søgningen og returnerer svaret via porten til klienten. Brugeren vil nu kunne se svaret på sin skærm.

Hvad er "attribut-kombinationer"?

Hvis man vil vide, hvordan en RPN-streng opbygges, er det nødvendigt at fordybe sig lidt i attributter (ikke dem, som man normalt tænker på, men dem, der er defineret i Z39.50-profilen).

Nu gælder det om at holde tungen lige i munden.

En RPN-streng indeholder et *attribut-sæt* bestående af op til 6 *attribut-typer*.

Hver attribut-type har flere værdier, der hver især angiver en egenskab for det/de søgeord, der sendes afsted.

Værdien udtrykkes som et tal, hvis betydning er forudbestemt i Z39.50-definitionerne.

Det vil sige, at hvert søgeord har 6 tilknyttede egenskaber, der kan justeres efter behov ("knapper" man kan skrue på).

De 6 *attribut-typer* ("knapper") er:

- 1: use-attribut (obligatorisk) - angiver hvilken type ord der kommer (f.eks. forfatter, titel, ISBN)
- 2: relation-attribut - angiver relationen til use-attr. (f.eks. større end, lig med, mindre end)
- 3: position-attribut - angiver placering i dokument/register (f.eks. først i felt/delfelt, hvor som helst)
- 4: struktur-attribut (obligatorisk) - angiver strengens struktur (f.eks. ord, frase, årstal, normaliseret)
- 5: trunkering-attribut - angiver trunkering (f.eks. ingen, højre, venstre, begge)
- 6: Completeness-attribut - angiver hvor meget, der skal bruges (f.eks. complete field, incomplete field)

Om use-attributten (attr 1) er en forfatter eller en titel o.s.v. angives som tidligere nævnt med en talværdi. F.eks. er forfatter = 1003, titel = 4 og ISBN = 7. En use-attribut-angivelse for en titel vil så se sådan ud: 1:4, og tilsvarende for de øvrige attribut-typer

Det vil altså sige, at et attribut-sæt, der indeholder værdierne: 1:4 2:3 3:3 4:2 5:1 6:1 "hus"

betyder 'titel' 'lig med' 'hvor som helst i felt/delfelt' 'ordsøgning' 'ingen trunkering' 'ukomplet felt/delfelt' 'søgetermen hus', hvilket gerne skulle resultere i en søgning på titelordet hus som ord i et titel-ord-register.

Som rigtig RPN-streng vil den se sådan ud: @attr 1=4 @attr 2=3 @attr 3=3 @attr 4=2 @attr 5=1 @attr 6=1 "hus"

Uanset om man forstår ovenstående "til bunds" kan man se, at antallet af kombinationsmuligheder vil være meget stort.

Derfor besluttet det ved implementering/opsætning af klienten blandt andet

- hvilke attribut-sæt-kombinationer der skal sendes afsted
- for hver udfyldt "søgekasse"
- for hvert rullegardinvalg
- for hver CCL-søgestreng
- hvilket format svaret skal returneres i
- og meget andet.

Når man implementerer target (porten til basen), beslutter man blandt andet

- hvilke attribut-sæt-kombinationer, det er logisk at tillade
- hvilke registre hver attribut-sæt-kombination skal "ramme"
- hvilke tegn der er tilladt i en søgestreng
- hvilke formater der tilbydes, og hvordan de ser ud
- og meget andet.

Hvad skal det bruges til?

Z39.50 har mindst én meget nyttig anvendelse, som er det hele værd. Med denne protokol kan man *søge* i baser, der har en Z39.50-adgang *uden at forlade sin egen brugergrænseflade*.

Det vil sige, at man kan udvælge nogle relevante baser og sende sin søgning afsted én gang. Man får så et svar tilbage fra hver af de valgte baser *i ens egen brugergrænseflade*. Man behøver altså ikke at kende 27 forskellige søgesprog og 35 forskellige måder at komme videre på.

Hvorfor skal jeg "blande mig"?

Der er to kernepunkter, der har betydning for, om man får et optimalt udbytte af søgning i "fremmede" baser:

Det første er, om klienten sender søgningen/skanningen afsted som "gode" RPN-streng.

Det andet er, om basen har "rigtige" registre i forhold til, hvad man forventer.

For at få den bedste implementering er det vigtigt, at brugerne af klienten og targets baser er i dialog med systemfolkene, da det er hos de professionelle brugere, søgeekspertisen findes. Det er de professionelle brugere, der ved, hvilke registre der er gode at bruge til hvad, og hvilke søgemuligheder det er nyttigt at have i grænserfladen.

Det er også vigtigt at forholde sig til, hvad basens registre indeholder. "Praksisregler for søgeveje" har forsøgt at fastlægge, hvad hvert enkelt register (hver enkelt søgekode) skal dække, så man ved, hvad man kan forvente, når man søger i det. Desværre er disse regler ikke slået rigtigt igennem - ofte på grund af lokale hensyn og lokale traditioner.

Hvis samsøgning i flere baser skal kunne gennemføres på et detaljeret niveau, er det nødvendigt at diskutere på bibliotekerne, om et evt. lokalt afvigende registerindhold er så vigtigt, at det skal fastholdes, eller om man vil omlægge til Praksisreglerne. Især defaultregisterindholdet varierer meget fra base til base.

Læs mere:

På adressen <http://www.bs.dk/> under emnet danZIG findes en udførlig beskrivelse af Z39.50 i Danmark og mange links til internationale dokumenter. På samme adresse findes under Publikationer "Praksisregler for søgeveje"

På adressen <http://www.nlc-bnc.ca/bath/bp-current.htm> findes den internationale Bath Profile