

FRITEKSTSØGNING

Fritekstsøgning kan her defineres som søgning efter datanumeriske elementer, der ligger i de søgeregistre, der danner basis for den pågældende database. Dette skal forstås som, at 'fri-tekst søgning' ikke nødvendigvis er fuldstændig identisk fra database A til database B. De standard søgeregistre, der afsøges, vil dog altid i det mindste være dannet af elementer i titel og abstract registrene (men heller ikke dette er entydigt, fx er der ikke abstract i SveMed databasen). Ofte indgår der også hele eller dele af emnekoder, forfattere m.m. i standard fritekstsøgning. Fritekstsøgning kan typisk kvalificeres/afgrænses man kan fx begrænse til søgning på titel eller abstract.

Sprogbrug

Fordelen ved fritekstsøgning er, at man bruger forfatterens eget ordvalg. Desværre er forfatterne ikke nødvendigvis enige om betegnelser, forkortelser osv. Det kan derfor være hensigtsmæssigt at inkludere alle de af forfatterne benyttede fagudtryk, hvis man ønsker en høj recall (genfinding) af sin søgning – altså hvis man ønsker at finde mest mulig relevant litteratur. Der er desuden problemer med brug af lægelatin versus amerikansk-henholdsvis britisk-engelsk, som kan drille alvorligt i sundhedsvidenskabelig fritekstsøgning. Det kan derfor være hensigtsmæssigt at søge på både britisk-engelske og amerikansk-engelske stavemåder. PubMed oversætter selv mellem fag-engelske udtryk i meget høj grad (men der kan ikke søges på latin!).

Indekseringsdybden

Optimal fritekstsøgning kan også kræve viden om opbygningen og indholdet af de bibliografiske databaser. De klassiske videnskabelige databaser, som PubMed, Web of Science m.fl., indekserer hovedsageligt kun forfatter(e), titel og

abstract fra de originale artikler. Hertil eventuelt også emne- og citationsdata, afhængigt af databasen. Nyere søgemaskiner som Google Scholar indekserer også maskinelt – om end delvist – selve indholdet af artiklernes *fulde tekst*. Dette muliggør en langt grundigere søgning efter mere specielle tekst- eller datastrengene. Problemet er, at omfanget af denne automatiske indekseringsproces holdes skjult for brugerne – i modsætning til de klassiske databaser, hvis indekseringsproces er klart og fuldstændigt deklareret.

Johan A. Wallin, redigeret af Mette Brandt Eriksen 30.7.2015