

Relevans: kvalitet eller kvantitet? - forsøg på en begrebsmæssig afklaring

Af Birgit Heie

Abstract

Indledningsvis fastslås relevansbegrebets betydning for BDI-informationsvidenskaben. Herefter undersøges betydningen af det engelske *relevance*, det danske *relevans* og de synonymer der har været brugt. Dette munder ud i en grundlæggende forståelse af relevans som et forhold der indebærer en fortolkning. Ud fra en fastlæggelse af hvilke relevansforhold der kan være tale om i IR, beskrives de forskellige forhold kort med hensyn til hovedproblemstillinger, metoder og forskningsmæssig betydning. Herefter præsenteres fire teorier om relevans: psykologisk relevans, situationel relevans, hermeneutisk relevans og faglig relevans, og ligheder og forskelle mellem de forskellige teorier trækkes op. Endelig diskuteres anvendelsen af relevans som mål ved evaluering af informationssystemer.

1. Indledning

Relevans er et af BDI-informationsvidenskabens grundbegreber. Begrebet er vævet dybt sammen

med informationsvidenskabens teori og praksis, og de forskellige definitioner og teorier der gennem tiden har været fremsat, afspejler derfor de forskellige forskningsmæssige holdninger og synsvinkler der har præget informationsvidenskaben. Relevans har tætte relationer til en række andre grundbegreber, f.eks. information, viden og emne, og man kan næppe afklare relevans uden også at komme ind på sammenhængen til disse begreber.

Relevansbegrebets vigtighed for BDI-informationsvidenskaben viser sig især på to områder. For det første er relevans siden 50'erne blevet brugt som et mål for IR-systemers effektivitet, dvs. man evaluerer IR-systemerne i forhold til relevansen af de resultater man opnår, og begrebet indgår implicit eller eksplicit i en lang række forskningsprojekter. For det andet indgår *relevans-feedback* i forskellige sammensatte partiel match systemer, således at man på baggrund af en brugers konkrete feedback på søgeresultatets relevans efter et første gennemløb, modificerer søgespørgsmålet eller på anden måde anvender vurderingen ved supplerende søgninger.

I denne artikel vil jeg forsøge at indkredse relevansbegrebet nærmere. Jeg har ønsket at danne mig et overblik over begrebets og problemstillingerens mangfoldighed, primært ud fra en teoretisk synsvinkel. Jeg har således ikke efterforsket primærlitteratur om resultater af empirisk forskning i relevans, men koncentreret mig om litteratur der forholder sig til definitioner, afgrænsning og oversigt. Mit udgangspunkt har været at få afklaret hvad relevans er, og hvilke roller det har spillet i IR. Desuden har jeg været interesseret i at undersøge hvor teoridannelserne omkring relevans befinder sig i dag, og om disse teorier giver baggrund for en fortsat anvendelse af de relevansbaserede mål recall/precision ved evaluering af informationssystemer.

Da den største del af primærlitteraturen er på engelsk, vil jeg indledningsvis kort diskutere forholdet mellem det engelske *relevance*, det danske *relevans*, og de synonymer der har været anvendt. Relevans kan forstås som en fortolkning af et forhold mellem to elementer, og i IR kan der være tale om en række forskellige relevansforhold. Jeg vil forsøge at belyse disse forskellige forhold og undersøge hvad begrebet betegner i de forskellige forståelser, og hvilke følger har det fået for teorier og eksperimenter og for udvikling og evaluering af IR-systemer.

Herefter vil jeg se nærmere på nogle nutidige forsøg på dels at definere relevans kvalitativt, altså anskue begrebet teoretisk i sammenhæng med andre vigtige grundbegreber, og jeg vil diskutere de ligheder og forskelle der er mellem disse opfattelser. Endelig vil jeg diskutere brugen af relevans som kvantitativt evalueringsmål og se på et par forsøg på at operationalisere begrebet i lyset af den kritik der har senest har været rejst heraf.

2. Hvad betyder relevans?

En indkredsning af den normale danske betydning af ordet *relevans* giver følgende resultat: *tilknytning til hovedsagen; betydningsfuldhed, væsentlighed i en given sammenhæng; gyldighed.*² Hovedbe-

tydningen af det engelske *relevance* er: *bearing upon or connected with the matter at hand; pertinent.*²

Det engelske *relevance* svarer til det danske *relevans*, men har yderligere bibetydningen *anvendelig*, der antyder et nytteperspektiv.³ Umiddelbart synes en oversættelse af *relevance* med *relevans* dog ganske dækkende.

I en del af den engelsksprogede litteratur om emnet er relevansbegrebet imidlertid søgt splittet op i forskellige delbegreber, der så har fået deres eget ord, eller det er erstattet med en anden mere eller mindre synonym betegnelse.⁴ De væsentligste af disse betegnelser er: *pertinence, usefulness, utility og satisfaction.*

Med hensyn til *pertinence* er vi ikke godt stillet på dansk, en ordbog giver følgende oversættelse: *relevans, forbindelse med den foreliggende sag*⁵. Som ovenfor nævnt forklares også på engelsk *relevance* med bl.a. *pertinence*. Så her er der blot tale om at man på engelsk har flere synonymer til rådighed. *Pertinence* har især været brugt specielt om relevans af f.eks. dokumenter i forhold til informationsbehov,⁵ men der findes dog også andre forsøg på definitioner. Buckland (1983) foreslår således *pertinence* anvendt om emnemæssig eller anden beslægtethed. Så faglig entydighed er der altså heller ikke om dette ords anvendelse. Froehlich argumenter mod en professionel skelnen mellem *relevans* og *pertinence*: "Relevance judgements do not conform to the professional research distinction between relevance and pertinence" (1994, side 128).

Usefulness og *utility* - på dansk *anvendelighed* og *nytte* - giver det særlige aspekt at informationen skal bruges af nogen, og at den skal vurderes i forhold til sin brugbarhed. Men definitionerne kan gøres så brede at det kan være vanskeligt at se fordelene ved at bruge disse begreber i forhold til det mere anvendte *relevans*.⁶

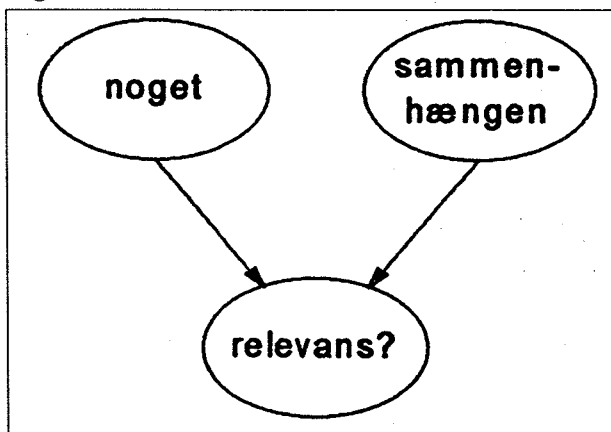
Satisfaction - på dansk *tilfredsstillelse* - knytter sig tæt til begrebet *behov*. Det giver en drejning hen

imod det mere subjektive, men begrebet har i øvrigt været anvendt i nogenlunde samme betydning som *pertinence*.

Man kan selvfølgelig vælge et eller flere af de anførte synonymmer for at pointere et bestemt aspekt, eller for at foretage en professionel skelnen mellem forskellige relevansforhold. Men der har ikke været faglig konsensus på området, og relevans står stadig tilbage som den centrale, mest brugte og bredest favnende betegnelse.

Trods adskillige forslag mangler der dog stadig en definition der er bred enighed om. Schamber et al. beklager dette og ønsker en afklaring:

Figur 1: Relevans



“Information scientists must finally establish a full theoretical and empirical understanding or definition of relevance as a fundamental concept, so that the discipline can move on to other matters” (Schamber et al., 1990, side 756).

Mangetydigheden kan imidlertid også ses som en fordel:

“In sum, the ordinary language use of “relevance” is quite appropriate for making judgments about the output or success of information systems, because it can function in the many ways in which various users make relevance judgments” (Froehlich, 1994, side 132).

Accepterer man mangetydigheden, stiller det imidlertid store krav til redegørelse for hvorledes begrebet anvendes i hvert enkelt tilfælde, og en overordnet teoretisk forståelse af relevansbegrebet i IR, som Schamber et al. efterlyser, kan være nødvendig som ramme eller fixpunkt for de forskellige måder at bruge begrebet på i forskning og udvikling.

Relevans er som nævnt et grundbegreb inden for IR. IR drejer sig nemlig ikke primært om at finde hvilken som helst information, men om at finde *relevant* information. *Relevans* er en karakteristik af den information man ønsker at finde - omvendt har man traditionelt målt processens succes på graden af relevans.

Relevans, det at noget har betydning eller er væsentligt i en sammenhæng, drejer sig om *forholdet* mellem to elementer i en *kommunikationsproces*, nemlig mellem dette “noget” og “sammenhængen”, det kan konkret være mellem f.eks. dokumenter og en forespørgsel efter information. Se figur 1. Relevans er ikke en iboende egenskab, og man kan ikke bare sige at noget er relevant; det kan kun være relevant i kraft af sit forhold til noget andet.

Relevans forudsætter en *fortolkning* af forholdets elementer, der *vurderes* eller *bedømmes* ud fra en bestemt synsvinkel eller bestemte *kriterier*, således at nogle egenskaber eller karakteristika vurderes som væsentlige, andre som uvæsentlige. Dette forudsætter igen et *subjekt* (eller flere) der foretager vurderingen, og evt. et subjekt som vurderingen foretages i forhold til; det kan selvfølgelig være den samme (f.eks. brugeren) eller det kan være forskellige personer. Andre spørgsmål kan tilføjes, fx: I hvilken *situation*, det vil blandt andet sige med hvilket *formål*, med hvilke *metoder* og under hvilke *omstændigheder* eller *betingelser* bliver vurderingen foretaget? På hvilket *tidspunkt*?

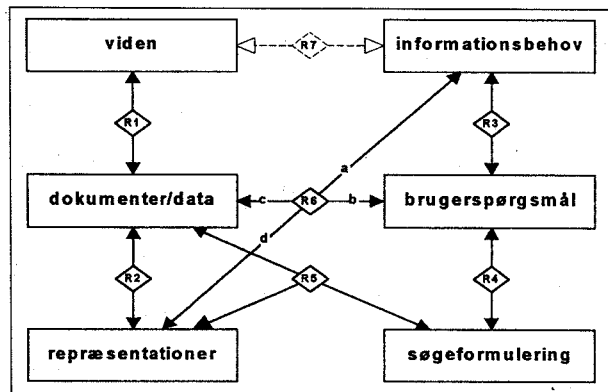
I det følgende afsnit vil jeg tage udgangspunkt i forskellige relevansforhold og sætte dem i relation til IR-forskningens forskellige traditioner. Jeg vil nærmere undersøge hvilke elementer der indgår i

de forskellige forhold, samt hvilke af de andre variable der særligt kendetegner relevansforholdene.

3. Relevansforhold

Nedenstående figur 2 er et forsøg på at få et overblik over de væsentligste svar der kan gives på spørgsmålet: Hvad skal være relevant i forhold til hvad? - altså en slags klassifikation af relevansforhold.

Figur 2: Relevansforhold



Til venstre i figuren ses de elementer der indgår i informationssystemerne: viden, dokumenter og repræsentationer. Altså det "noget" der skal vurderes. Til højre de elementer der særligt har med brugerne at gøre, og som udgør den sammenhæng "noget" skal vurderes i forhold til: informationsbehov, brugerspørgsmål (*request*) og søgeformulering (*query*). Men også mellem de forskellige elementer i henholdsvis venstre og højre side kan man tale om relevansforhold, som så kommer til at ligge til grund for, eller forud for, relevansforhold mellem brugersiden og systemsiden.

De følgende afsnit bygger især på de oversigter jeg har fundet hos Saracevic (1975), Schamber et al. (1990) og Froehlich (1994).

3.1 Viden, dokumenter og repræsentationer

I figur 2 ses nederst til venstre de enheder der findes i IR-systemerne: dokumenter og repræsentationer. Dokumenterne indeholder potentiel informati-

on i form af data. Dokumenterne kan forekomme i deres fulde udstrækning, men som oftest er de i IR systemerne repræsenteret i en bestemt form, normalt en kombination af bibliografiske data, klassifikation og indeksering. En repræsentation bygger på en fortolkning og vil altid være udtryk for en reduktion af visse af dokumentets aspekter, ofte af betydeligt omfang, samtidig med at andre aspekter bliver forstærket. Spørgsmålet bliver derfor i hvor høj grad repræsentationerne er relevante for de dokumenter de repræsenterer (i figuren angivet som R2) og hvorvidt forskellige metoder kan bruges til at repræsentere dokumenter på en hensigtsmæssig måde. Forholdet behandles inden for faglige områder som klassifikationsteori og emneteori.

Under dette relevansforhold gemmer der sig endnu et (figurens R1), nemlig forholdet mellem viden og repræsentation af viden, f.eks. mellem en fagligvidenskabelig viden, de dokumenter der repræsenterer denne viden, og det udvalg af dokumenter der i et givet system repræsenterer den faglige viden. Og det er jo dybest set denne viden man søger gennem informationssystemerne, ikke dokumenter eller data i sig selv.

Er det overhovedet rimeligt at kalde disse forhold relevansforhold, eller er der blot tale om en vurdering af kvaliteten af forskellige former for repræsentation? Jeg mener godt en diskussion af relevans kan rumme også disse forhold, og at det kan være rimeligt at tale om relevans i forbindelse med repræsentationer af henholdsvis viden i dokumenter og af dokumenter i f.eks. poster. Det er dog ikke i første række det enkelte konkrete forhold mellem viden og dokument der er interessant, men det mønster af relevansforhold der tegner sig inden for et fagligt område. Tilsvarende drejer det sig ikke om repræsentationen af det enkelte dokument, men om de muligheder det pågældende informationssystems repræsentationspraksis giver.

De to relevansforhold dækker det Saracevic (1975) kalder The Subject Knowledge View og The Subject Literature View - begge med fokus på emnemæssig relevans, inden for henholdsvis en viden-

skabelig-faglig viden og en faglig litteratur. Sara-cevic anfører The Subject Knowledge View som det centrale og grundlæggende relevansforhold, og han beklager at der ikke forskningsmæssigt har været ofret tilstrækkelig opmærksomhed på det. Hjørland (1995, side 360) tilslutter sig dette synspunkt. Han mener at videnskabsteori, filosofi og de videnskabelige fags grundlag, kan bidrage til en belysning.

Substansen i disse relevansforhold er emnemæssig beslægtethed, men også andre faktorer kan have indflydelse: metode, tradition, synsvinkel eller paradigme, institutionelt tilhørsforhold, prestige osv. Spørgsmålet er i hvor høj grad disse faktorer kommer til udtryk jo højere man kommer i repræsentationshierakiet, og om de reduktioner repræsentationerne er udtryk for, i tilstrækkelig grad tilgodeser brugernes forskelligartede behov i forbindelse med relevansvurderinger.

3.2 Informationsbehov, brugerspørgsmål og søgeformulering

Brugersiden er til højre i figuren. I figuren har jeg valgt kun at nævne tre begreber: informationsbehov, brugerspørgsmål (som det fremsættes over for systemet) og søgeformulering (det formaliserede søgespørgsmål); men flere kan trækkes frem, f.eks. brugerens forudsætninger (herunder vidensstrukturer), problemområder, arbejdsopgaver m.fl. I denne sammenhæng er det mest interessant at forstå informationsbehov som det bevidste og formulerede informationsbehov, svarende til Q3 hos Taylor (1968); mens brugerspørgsmål svarer til Q4, nemlig det ønske om information som brugeren fremsætter over for informationssystemet. Også her kan man tale om relevans, nemlig om i hvor høj grad brugerspørgsmålet er et relevant udtryk for informationsbehovet (angivet som R3). Her kommer også problemer som kompromisser og etiketteffekt ind i billedet.⁷

Desuden kan man tale om hvorvidt søgeformuleringen, som normalt er en formel-logisk algoritme, er et relevant udtryk for brugerspørgsmålet (figu-

rens R4). Her er der igen tale om repræsentation, idet spørgsmålet, fremsat i talesprog og måske udledt af formidleren gennem et længere interview, bliver omsat til det meget præcise og derfor meget begrænsede søgesprog.

Kernen i dette relevansforhold er formidlerens forståelse af brugeren, dennes baggrundsviden og arbejdsområde, det konkrete problem og informationsbehov. Det medfører en interesse for de metoder formidleren (menneskelig eller maskinel) anvender til dels at afdække informationsbehovet, dels at omsætte det i et konkret søgespørgsmål.

Dette relevansforhold har bl.a. været i fokus ved udvikling af formidlingsmekanismer i ekspert-IR-systemer. Problemet er her hvorledes ekspertsystemet (et edb-program) "oversætter" brugerens spørgsmål til en søgeformulering, evt. ved først at foretage et struktureret interview med brugeren for at "forstå" informationsbehovet, f.eks. ved hjælp associationsnetværk eller stereotyper i en eller anden form, og om man overhovedet kan eller skal forsøge at udvikle programmer til dette. Dette problem inddrager også diskussionen om kunstig intelligens, bl.a. hvorledes edb-programmer håndterer naturligt sprog.

Også her kan man spørge om det er rimeligt at tale om relevansforhold. Personligt mener jeg at en sådan betragtning kan være med til at kaste lys over disse forhold. Hvordan søger brugere at give relevant udtryk for deres informationsbehov, og hvordan forsøger formidlere at trænge bag om et kompromisset udtryk? Er en søgeformulering relevant for et givet brugerspørgsmål? Giver et bestemt informationssystem søgefaciliteter overhovedet mulighed for at udforme en relevant søgeformulering? Traditionelt har vægten ligget på især at give mulighed for angivelse af emnemæssig relevans, men - som undersøgelser af relevansvurderinger viser - en lang række andre kriterier har (mindst) lige så stor betydning.

3.3 Repræsentationer og søgeformulering

De relevansforhold der især har haft betydning i IR-forskning, er de forskellige forhold der kan opregnes mellem systemsiden (til venstre i figur 2) og brugersiden (til højre). En række forskellige kombinationer er teoretisk mulige, men især to har været fremherskende: Forholdet mellem repræsentationer, dokumenter/data og søgeformulering, og forholdet mellem søgeresultatet som helhed, informationsbehov og brugerspørgsmål.

I traditionel eller systemorienteret IR-forskning har interessen samlet sig om relevans mellem en given søgeformulering og repræsentationerne for et sæt udsøgte dokumenter (figurens R5). Relevans er her altid forstået som emnemæssig beslægtethed, og relevansvurderingen er forsøgt objektiviseret. Man har først og fremmest været interesseret i at anvende relevansvurderinger som konstanter til at måle og sammenligne effektiviteten af forskellige indekserings- og søgesystemer, dvs. systemernes evne til at matche søgeformulering og repræsentationer.

Ved eksakt match forstås relevans som en absolut angivelse: et dokument kan enten være relevant eller det kan være ikke-relevant i forhold til en søgeformulering. Ved partiel match drejer det sig derimod om graden af overensstemmelse mellem de termer der indgår i søgeformulering'en, og de termer der indgår i dokumentrepræsentationerne. Termerne er normalt vægtede efter bestemte principper og overensstemmelsen mellem søgeformuleringen og det enkelte dokument udregnes ved hjælp af matematiske lighedsmål. I partiel match bliver relevans således skalérbar; der kan være tale om større eller mindre relevans.

Relevans målt på denne måde har været taget som et direkte udtryk for IR-systemernes effektivitet. Ved at anvende testdatabaser hvor dokumenterne i forvejen var relevansvurderede i forhold til bestemte søgeformuleringer, kunne man sammenligne effektiviteten af forskellige IR-systemer, bl.a. ved at udregne recall og precision. Recall drejer sig om at udsøge et så stort antal af de relevante dokumen-

ter som muligt, precision⁸ om at udsøge så få af de ikke-relevante dokumenter som muligt. Hermed drages også begrebet ikke-relevans ind.

Ved disse effektivitetsundersøgelser drejer det sig om målbarhed og sammenligning, metoderne er matematiske og statistiske, og idealet er klassisk empirisk-rationalistisk naturvidenskab. Også relevansdomme blev forsøgt passet ind i dette ideal. Man tilstræbte en objektivitet f.eks. ved at spørgsmålene var konstruerede til formålet ud fra udvalgte dokumenter (source-dokument test), og ved at relevansbedømmelserne blev foretaget af en gruppe af fagfolk, oftest studerende og forskere. På denne måde forsøgte man at reducere antallet af variable, således at det kun var systemernes evne til at finde relevante og udelade ikke-relevante dokumenter der blev undersøgt. Målet var at finde alment gyldige lovmæssigheder.

Dette relevansforhold har været det fremherskende i informationsvidenskabens første årtier - 40'erne og 50'erne - og det ses stadig brugt ved evaluering af IR-systemer i testdatabaser, på trods af massiv kritik. Allerede fra starten viste der sig voldsom uenighed i de konkrete relevansdomme, der på den måde alligevel viste sig som variable.⁹ Uenigheden om relevansdommene førte til kritik af denne tradition og dens ensidige læggen vægt på emnemæssigt beslægtethed som kriterium. Samtidig kritiseredes forsøgene for at være for urealistiske ved ikke at afspejle brugernes behov og relevansopfattelse.¹⁰

3.4 Søgeresultat, informationsbehov og brugerspørgsmål

I takt med et ønske om en mere brugerorienteret forskning opstod et krav om at undersøge rigtige relevansvurderinger i forhold til ægte informationsbehov hos levende brugere i fungerende informationssystemer. Relevans afgrænses til relevans af søgeresultater (repræsentationer eller data) i forhold til brugerens spørgsmål (R6, b-c/d i figur 2) eller informationsbehov (R6, a-c/d). Synsvinklen lægger vægt på brugeren som individ, som subjekt, og

relevans forstås i forhold til den enkelte bruger. Dette bringer vurderingerne i centrum med vægt på de faktorer der påvirker dem,¹¹ og på de kriterier der ligger til grund for vurderingerne.¹² I de senere år har der været øget opmærksomhed om tidsdimensionen og søgeprocessens dynamiske karakter; det at brugerens spørgsmål og ofte også informationsbehov ændrer sig gennem søgeprocessen,¹³ bl.a. i kraft af at en sådan proces samtidig er en indlærings- og erkendelsesproces.

Metoderne til at afklare individuelle relevansvurderinger er fortrinsvis kvalitative, psykologiske og beskrivende; man får fastlagt en række faktorer og kriterier der kan spille ind, og dermed problematiseres den mere kvantitative tilgang i den systemorienterede forskning. Idealet er imidlertid stadig at nå frem til om ikke objektive, så dog generaliserbare resultater. Det kræver en balancegang mellem på den ene side case-studies med en begrænset gyldighed, f.eks. af bestemte brugergrupper eller fagspecifikke informationssystemer, og på den anden side undersøgelser med en så generel tilgang at resultatet bliver indholdsløst. Afdækningen af forholdets subjektive og variable karakter rejser desuden en ny række af problemer ved anvendelsen af relevans som evalueringsmål for IR-systemer. Dette behandles nærmere nedenfor i afsnit 5.

3.5 Helhedsbetragtning

Et af formålene med IR-systemer kan ses som opfyldelse af informationsbehov ved udsøgning af relevante data eller dokumenter (R6), og man kan derfor argumentere for at dette relevansforhold er af overordnet karakter. Underforstået findes imidlertid endnu et relevansforhold, antydnet som R7 i figur 2: det er i virkeligheden viden man søger når man henvender sig til et informationssystem med et informationsbehov, såvel repræsentationer som dokumenter er at betragte som midler. Alligevel er også de øvrige ovennævnte relevansforhold (R1-R5) væsentlige at tage i betragtning, idet man i praksis ikke kan knytte en direkte forbindelse mellem et informationsbehov (der jo - omend formuleret for brugeren selv - ikke er fremsat for IR-systemet)

met) og dokumenter eller data fra et informationssystem. Vejen går gennem fremsættelse af brugerspørgsmål og (i hvert tilfælde i nærmeste fremtid stadig) søgeformulering, ligesom der ikke foreløbig er tegn på at man i større udstrækning kan søge direkte i store dokument-samlinger og dermed undgå anvendelse af repræsentationer. På den måde bliver R1-R5 bestemmende for det overordnede resultat, R6. Er f.eks. en søgeformulering ikke relevant for et brugerspørgsmål, er det urealistisk at forestille sig at søgeresultatet vil være relevant for det tilgrundliggende informationsbehov. Omvendt vil en relevant match mellem en søgeformulering og et sæt repræsentationer ikke i sig selv garantere for at brugeren får sit informationsbehov opfyldt ved de dokumenter sættet repræsenterer.

Saracevic påpeger gentagne gange i sin forskningsoversigt at ingen af de enkelte relevansopfattelser er udtømmende, og han sammenfatter afslutningsvis:

“There is no, and there cannot be any one specific, view of relevance, for there does not exist any one system of relevance in communication. Different systems of relevances may involve some different factors, but they are coupled in such a way that they can hardly be considered without other systems of relevance” (Saracevic, 1975, side 339).

4. Relevans som kvalitet

Jeg vil i det følgende nærmere undersøge fire forskellige forsøg fra begyndelsen af 90'erne på at give en kvalitativ definition på relevans i en mere omfattende teoretisk sammenhæng, nemlig Harter (1992), Schamber et al. (1990), Froehlich (1994) og Hjørland (1993). Jeg vil herefter sammenligne de forskellige teorier og se nærmere på deres opfattelse af relevans som et individuelt eller et kollektivt anliggende, som et subjektivt eller objektivt fænomen samt i hvilken grad de lægger vægt på emnemæssig relevans. Endelig vil jeg diskutere hvorvidt teorierne faktisk er i modstrid med hinanden, eller om de blot beskæftiger sig med relevans på forskellige niveauer.

4.1 Psykologisk relevans - Harter

Stephen P. Harter (1992) tager udgangspunkt i begrebet *psykologisk relevans* som det er udviklet af Sperber og Wilson i 1986. Sperber og Wilson anvender begrebet inden for verbal kommunikation, og de har udviklet et helt begrebssystem til deres teorier. Harter forsøger at overføre deres teorier og begreber til informationssøgning.

Sperber og Wilson skelner imellem *manifest viden*, dvs. det man er i stand til at vide; og *mentale repræsentationer*, dvs. det man rent faktisk véd. En persons *kognitive rum* (environment) er det sæt af fakta (ikke nødvendigvis sande i absolut forstand) og antagelser der er manifest for personen, dvs. det omfatter også mentale repræsentationer som jo bl.a. er karakteriseret ved at man er i stand til at vide dem.

En persons *kontekst* repræsenterer de antagelser om verden personen har på et givet tidspunkt, og kontekst er således et dynamisk begreb. Sperber og Wilson går ud fra den grundlæggende antagelse at en persons *kognitive mål* på et givet tidspunkt er med det mindst mulige besvær at opnå maksimal relevans mellem en given information og konteksten. Informationen kan enten være andre persons antagelser der kommunikerer, eller det kan være fænomener der på en eller anden måde påvirker personens kognitive rum. Konteksten er ikke givet på forhånd, den bliver valgt i forhold til det kognitive mål.

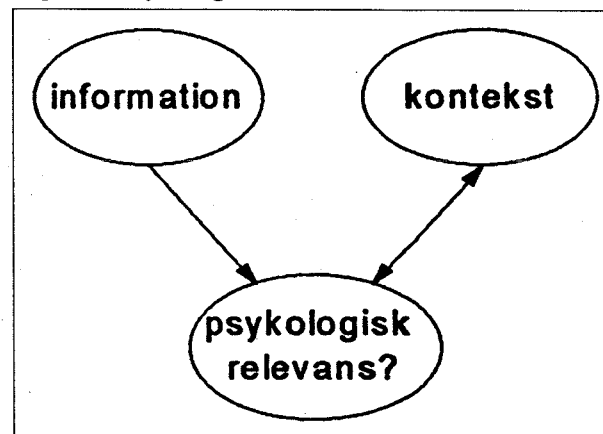
Der er tre måder hvorpå antagelser eller fænomener kan påvirke en kontekst:

- de kan gå i forbindelse med konteksten og frembringe nye antagelser,
- de kan medføre at man opgiver tidligere antagelser fra konteksten, eller
- de kan styrke eller svække antagelser i konteksten.

En bestemt information er *relevant* for en person, hvis den er relevant for mindst én af de kontekster der er mulige for en person på et givet tidspunkt.

Den er relevant i den udstrækning dens virkning på konteksten er stor, eller i den udstrækning den indsats der skal til for at behandle den, er lille. Konteksten vælges ud fra det kognitive mål at maksimere relevansen, dvs. at opnå den bedste balance mellem kognitiv udbytte og kognitiv indsats. Se figur 3.

Figur 3: *Psykologisk relevans*



Harter hævder en endog meget tæt relation mellem relevans og det Buckland (1991) kalder information-as-process:

“Relevance and information (as-process) can usefully be regarded as the same, or more accurately, as two very closely related aspects of the same experience. (...) The act of informing is the act of making something relevant. To be found relevant is to be information. To be informative is to be relevant. A necessary and sufficient condition for relevancy is informativeness” (Harter, 1992, side 611).

I forhold til IR ser Harter *informationsbehov* som den indledende kontekst for en person der i et informationssystem søger hjælp til et problem. Tendensen til at man altid vil søge at optimere relevansen ved at tilpasse konteksten, gør imidlertid informationsbehovet til en flydende, konstant foranderlig størrelse - enhver relevant information vil pr. definition forandre konteksten - og dermed informationsbehovet. Derfor kan en person nok forsøge at beskrive sit informationsbehov, men det

er illusorisk at beskrivelsen kan være udtømmende. Harter finder det derfor næppe umagen værd at gøre for meget ud af at beskrive informationsbehov andet end overfladisk, og han foretrækker i øvrigt betegnelsen kontekst som mere præcis.

Informationsbehovet indeholder ikke kun emnemæssige aspekter, men også alle mulige andre antagelser der vedrører den pågældende kontekst. Harter kalder emnemæssig relevans (on the topic)¹⁴ for *svag relevans* ud fra den forudsætning at emnemæssig relevant information er forudsigelig, og at den derfor ikke kan have den store forandrende virkning på konteksten. Den som søger information er interesseret i forandringer i sit kognitive rum, og derfor vil information der i større grad påvirker konteksten, have en *stærk relevans*.

“Thus we come to the perhaps surprising conclusion that, in information retrieval, references on the topic may be less important than relevant references not on the topic - references that allow the making of new intellectual connections or cause other cognitive change” (Harter, 1992, side 612).

Af samme grund finder han at citationssøgning (cited reference searching) i højere grad end emneordssøgning er et udtryk for psykologisk relevans.

Hvis der ikke er noget absolut informationsbehov, kan man heller ikke sige at søgeprocessen er motiveret af et enkelt mål. Snarere er det en dynamisk proces, hele tiden styret af de forandringer der sker i konteksten. Harter finder belæg for dette bl.a. i de teorier der er udviklet af Swanson i 1977 (informationssøgning som en trial-and-error proces), af Bates i 1989 (informationssøgning sammenlignet med bærplukning) og af Ellis i 1989 (en række forskellige adfærdsmønstre i informationssøgning), og han foreslår psykologisk relevans som en forenende og forklarende teori.

Det mest kontroversielle punkt ved Harters teori er at han så kraftigt nedtoner emnemæssig relevans, uden at han redegør for hvilken andet kriterium

eller hvilken anden forbindelse der skal være mellem informationen og konteksten før man kan tale om stærk relevans - kun at der skal være effekt på konteksten. Desuden definerer Harter ikke emne (topic) nærmere, og dette begreb er afgørende for en forståelse af emnemæssigt beslægtethed som aspekt af relevans.

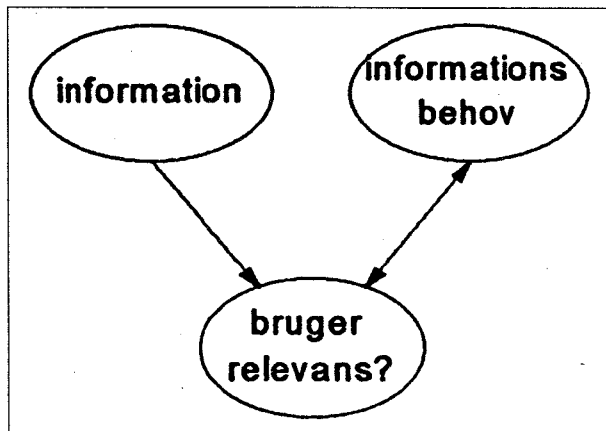
4.2 Situationel og dynamisk relevans - Schamber, Eisenberg og Nilan

Schamber, Eisenberg og Nilan (1990) konkluderer ud fra deres gennemgang af både traditionelle og alternative synspunkter: “Regardless of the perspective, the ultimate goal is to fulfill the information needs of users” (side 770), og de foreslår et dynamisk, situationelt relevansbegreb med fokus på brugeren.

Brugeren forstås som aktiv og lærende i en kontekst¹⁵ af konstant forandring, hvor hun eller han søger at finde mening¹⁶ på baggrund af en forståelse af fortid, nutid og fremtid (herunder mål og hensigter). *Informationsbehov* opstår hos brugeren som en opfattelse af den situationelle kontekst som uklar, usammenhængende eller usikker, samtidig med en forventning om at information kan hjælpe med til at finde mening. *Information* som mening eller betydning bliver konstrueret af brugeren i dennes bevidsthed som en fortolkning af de både indre og ydre faktorer som kan hjælpe med til at skabe sammenhæng i den situationelle kontekst.

Relevans er således en indre begivenhed, noget der sker i brugerens bevidsthed. *En relevansdom* er brugerens evaluering af information (fra både indre og ydre kilder) i forhold til informationsbehovet i en bestemt situation på et bestemt tidspunkt. Se figur 4. *Relevanskriterier* er de værdier der ligger bag en relevansdom, og som brugeren benytter til at bestemme kvaliteten af forholdet. Kriterierne er skabt af brugeren og knyttet til denne i tid og rum. På denne måde er relevans pr. definition bundet til en flerdimensionel usikkerhed: relevans betyder noget forskelligt for forskellige mennesker, og det betyder noget forskelligt for det samme menneske på forskellige tidspunkter.

Figur 4: *Situational relevans*



Informationssøgning forstås som en sekventiel, iterativ læreproces. Informationssystemerne og brugerne ses ikke som to adskilte enheder, men systemerne er en del af brugerens kontekst som de opfattes af hende eller ham, og som de indgår i læreprocessen.

For Schamber, Eisenberg og Nilan er det relevanskriterier der er interessante at udforske, dels i sig selv, men dels selvfølgelig også for at kunne forbedre design af informationssystemer. Undersøgelser af et så subjektivt og så tids-, steds- og situationsafhængigt et begreb kræver undersøgelser af så realistiske situationer som muligt: rigtigt brugere med ægte informationsbehov. Brugere interviewes enten fortløbende under processen eller efterfølgende, og fortæller i egne ord om sekvenser af deres søgning og evaluering. Informationssøgning må forstås som en helhed og undersøges som en helhed. Ifølge Schamber, Eisenberg og Nilan giver de "snapshots" tidligere relevansundersøgelser har taget, kun delvise sandheder.

Trods de individuelle variationer som et strengt subjektivt relevansbegreb medfører, giver begrebet om at finde eller skabe mening en overordnet referenceramme, der kan give mulighed for at finde generaliserbare mønstre, og Schamber, Eisenberg og Nilan anfører at der findes resultater der understøtter dette (1990, side 771-772).

Sammenfattende konkluderer de:

"1. Relevance is a multidimensional cognitive concept whose meaning is largely dependent on users' perceptions of information and their own information need situations.

2. Relevance is a dynamic concept that depends on users' judgements of the quality of the relationship between information and information need at a certain point in time.

3. Relevance is a complex but systematic and measurable concept if approached conceptually and operationally from the user's perspective." (Schamber et al., 1990, side 774).

4.3 Hermeneutisk relevans - Froehlich

Thomas J. Froehlich (1994, side 130ff) forsøger at syntetisere forskellige opfattelser af relevans. Han anfører at relevansbedømmelse som en udtalt forudsætning har været opfattet som en fortolkningshandling, bl.a. derfor mener han at den filosofiske hermeneutik, især som den kommer til udtryk hos Gadamer, kan være et brugbart grundlag for en forståelse af relevans.

Froehlich tager udgangspunkt i relevans som en såkaldt *naturlig kategori*,¹⁷ og ikke som et cartesiansk, klart definerbart videnskabeligt begreb. En naturlig kategori er karakteriseret ved en *prototype*. Kategoriens medlemmer vurderes ud fra *familieligheder* med denne prototype, og medlemmerne kan i forhold til denne være mere eller mindre typiske. Opfattelsen af en bestemt prototype opstår som en kulturel erfaring, og man kan ofte finde en stor grad af enighed om prototypen inden for forskellige sociale eller kulturelle grupperinger, f.eks. hos mennesker med den samme uddannelse eller de samme arbejdsopgaver. Froehlich mener at de forskellige forsøg på at tvinge relevansbegrebet ind i en snæver definatorisk ramme, begrænser en fornuftig anvendelse af begrebet.

Prototypen for relevans er ennemæssig beslægtet-

hed¹⁸, men der findes en lang række andre typer af, eller kriterier for, relevans. Grupper af brugere med et opgave- eller andet fællesskab vil normalt være fælles om en række kriterier med indbyrdes familieligheder: ikke alle bruger de samme kriterier, men mange deler mange kriterier. Den prototypiske emnemæssige beslægtethed kan optræde alene, som det eneste kriterium, men kombineres som regel med andre kriterier, og kombinationen kan give helt andre resultater. Froehlich kalder derfor emnemæssig beslægtethed for *polyvalent* (med et billede fra kemi). Emnemæssig relevans er således et ustabil kriterium - det kan optræde alene eller det kan forbinde sig med andre kriterier og dermed få et helt andet indhold eller måske helt tilsidesættes af disse andre kriterier. Relevansbedømmelse må forstås som en dynamisk og dialektisk proces: relevansdommene tildeler søgeresultaterne relevans, samtidig med at resultaterne påvirker relevansdommene ved at give anledning til redefinering eller ændring af informationsbehov og kriterier.

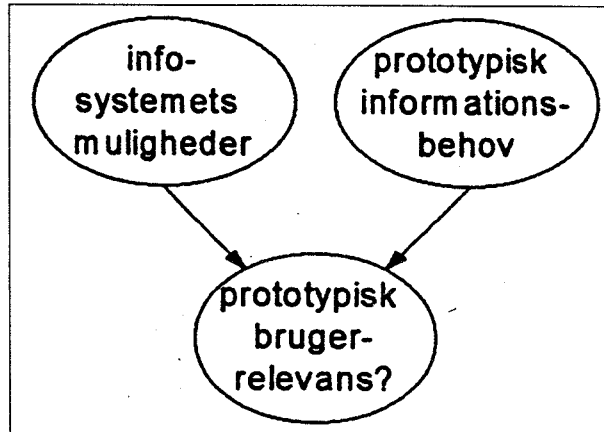
Froehlich anfører at relevans kan analyseres dels som en helhed indeholdende tre gensidigt forbundne hovedaspekter, og dels ud fra hver af tre hovedaspekter for sig: brugerne, deres behov og disses repræsentationer (R3 og R4 i figur 2); informationssamlingen og dens representation og/eller erstatning (R1 og R2 i figur 2); og formidlingen enten direkte gennem systemet eller gennem en formidler gennem systemet (R5 og R6 i figur 2).

Brugerne fortolker deres behov for sig selv, for formidlerne og for systemet i forhold til deres *horisont*¹⁹, *kontekst*²⁰ og *undersøgesobjekt*. Froehlich mener at fokus for et specialiseret informationssystem er - ikke den enkelte bruger - men derimod *den prototypiske bruger* inden for den eller de discipliner, professioner eller arbejdsfællesskaber som systemet betjener. Se figur 5.

“The case of individualization is overstated, for while end-users are the ultimate judge of the output of information retrieval systems, they are not absolutely autonomous judges, because they

bring with them the norms, mores and socializations they acquired as they became inculturated into a field (...)” (Froehlich, 1994, side 127).

Figur 5: Hermeneutisk relevans



En analyse af brugen af systemet bør udforske dialektikken mellem den prototypiske brugers faglige baggrund og hendes eller hans informationsbehov i forhold til den faglige struktur, f.eks. professionelle eller disciplinære standarder.

I *informationssystemet* fortolkes det enkelte værk i forhold til systemets emneområde: skal værket overhovedet med eller ej? Desuden indebærer behandlingen af værket i forhold til systemet i høj grad fortolkning, f.eks. ved indeksering. Også dette sker i forhold til en idé om systemets prototypiske brugere.

Endelig fortolker den menneskelige eller maskinelle *formidler* systemet og dets muligheder i forhold til brugerens informationsbehov som det udtrykkes i enten brugerspørgsmål eller søgeformulering.

4.4 Faglig relevans - Hjørland

Birger Hjørland argumenterer for at man inden for et videnskabeligt domæne eller paradigme kan opnå en vis grad af velargumenteret konsensus og dermed objektivitet:

“Et sammenhængende teoretisk grundlag for en

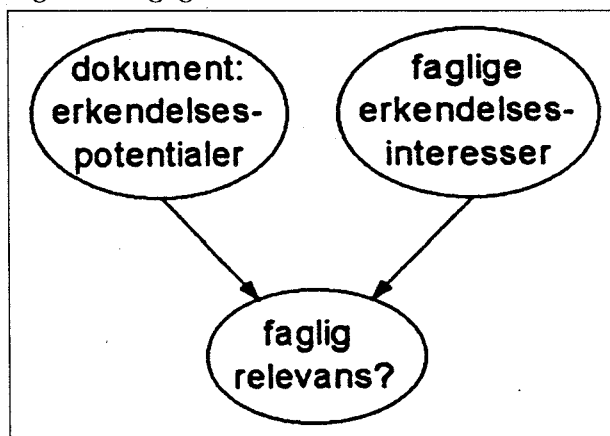
videnskab implicerer, hvilke informationer der især er relevante for den pågældende videnskab, og hvilke informationer der er mindre relevante eller eventuelt helt irrelevante” (Hjørland, 1993, side 149).

Hovedemnet for Hjørlands doktorafhandling fra 1993 er emneteorien, men *emne*²¹ er beslægtet med relevans og afhandlingen indeholder en række synspunkter vedrørende relevans:

“Vor analyse peger altså på, at begrebet “emne” er beslægtet med begrebet “relevans”: emneanalysen analyserer, prioriterer og udvælger relevante egenskaber ved dokumenter i forhold til mere specifikke eller mere almene erkendelsesbehov” (Hjørland, 1993, side 86).

Hjørland bekender sig angiveligt til en objektiv pragmatisme, og hans relevansbegreb er tilsvarende ikke centreret om brugerne. En objektiv relevansvurdering er således en vurdering af relevansen af et dokumentets erkendelsespotentialer (dvs. informative eller videnskabelige muligheder) i forhold til bestemte erkendelsesinteresser, nærmere bestemt i forhold til den kollektive viden inden for en eller flere videnskabelige discipliner. Se figur 6. Med et citat af Wilson fra 1993: “(..) the main way in which information from outside affects a speciality is by being recognized by the group as being impersonally, “objectively” relevant”.²²

Figur 6: Faglig relevans



Erkendelsesmæssige potentialer er dels “potentialer for videnskabelige fremskridt, d.v.s. perspektivet er menneskehedens samlede videnskabelige fremskridt” (1993, side 79), dels de muligheder der er for at anvende den pågældende information til en række formål. Hjørland fastslår at “dokumenter indeholder uendelig mange potentielle emner. En emneanalyse indebærer en stillingtagen til hvilke egenskaber af et dokument, der skal synliggøres” (ibid, side 72). Og “Et dokument har ikke eet sandt emne. Det har forskellige erkendelsesmæssige potentialer, som prioriteres fra faglige synsvinkler” (ibid, side 40). Analysen skal ske i en kontekst²³, det vil bl.a. sige ud fra de erkendelsesinteresser der ligger i et givet system.

“Men analysens sigte er ikke at være “subjektiv”, d.v.s. farvet af f.eks. den person, der foretager analysen, men mest mulig rammende, “objektiv”, d.v.s. mest mulig i overensstemmelse med sit objekt (...)” (Hjørland, 1993, side 39).

At foretage en objektiv relevansvurdering vil derfor - på samme måde som en objektiv emneanalyse - være at vurdere et dokument eller en dokumentrepræsentation i forhold til faglige erkendelsesinteresser. En emneanalyse skal således foretages “ud fra de spørgsmål, som man indenfor informations-systemet ser det som sin opgave at kunne besvare” (Hjørland, 1993, side 90). På samme måde må relevansvurderingerne foretages i relation til netop disse spørgsmål, der kan ses som de faglige erkendelsesinteresser et givet informationssystem betjener. Det er således også de spørgsmål der er væsentlige ved evaluering af et system.

Hjørland erkender selv at hans teori ikke er let at omsætte i handling, og at især identifikation af videnskabelige potentialer, som jo er et centralt begreb, er svær:

“Hvilken procedure viser vejen til dette? Ikke nogen let, operationaliserbar, objektiv” metode, men snarere en moden vurdering byggende på almen videnskabelig indsigt” (Hjørland, 1993, side 88).

Det kræver selvfølgelig - som også Hjørlands emneanalyse gør det - et overordentligt dybtgående kendskab til den eller de pågældende faglige discipliner eller domæner: til det faglige grundlag, til videnskabsteori m.m. Jo mere inde i disse forhold en person er, jo bedre vil vedkommende kunne foretage objektive relevansvurderinger.

"(...) filosofisk, videnskabsteoretisk, metodologisk og konceptuel afklaring [er] vejen til støtteklarhed over relevanskriterier og dermed over "informationsbehov" (Hjørland, 1993, side 194).

Det er interessant at bemærke at Hjørland her sætter relevanskriterier før - og dermed over - informationsbehov. I alle andre teorier er det omvendt. Hjørland stiller sig stærkt kritisk over for en psykologisk forståelse af informationsbehov.²⁴ "Et informationssystemets opgave er ikke blot at tilfredsstille behov, også at skabe eller udvikle behovene" (Hjørland, 1993, side 173). Han skelner i øvrigt imellem individuel og kollektiv erkendelsesudvikling; på det individuelle niveau tilnærmer man sig gradvis, f.eks. gennem uddannelse, det kollektive niveau; den kollektive udvikling derimod er synonym med disciplinens faghistorie.²⁵

4.5 Fælles træk og forskelle

Står disse fire teorier i modstrid med hinanden? Kun til en vis grad. De to af teorierne, Schamber et al. og Harter, beskæftiger sig med relevans i et mikroperspektiv, på et psykologisk-kognitivt niveau, mens de to andre, Froehlich og Hjørland, befinder sig på et socialt niveau med henholdsvis et hermeneutisk og et sociologisk-videnskabsteoretisk udgangspunkt, dvs. de ser relevans i et makroperspektiv²⁶. De to perspektiver udelukker ikke hinanden, men belyser begrebet i forskellige dimensioner. Fælles for de fire teorier er desuden en mere eller mindre eksplicit afstandtagen til det traditionelle, lighedsbaserede relevansbegreb, jf. afsnit 3.3. Alle erkender de behovet for en redefinering, baseret på en kvalitativ forståelse af relevansbegrebet i en større teoretisk sammenhæng.

En af de mest iøjnefaldende forskelle mellem teorierne ligger i opfattelsen af brugerne. I psykologisk relevans og situationel relevans er brugerens rolle central, han eller hun er aktiv og medskabende - ja, relevans kan overhovedet ikke tænkes uden at brugeren er inde i billedet. Relevans er således grundlæggende et subjektivt begreb, det eksisterer i og med et subjekt der fortolker og vurderer. Både Schamber et al. og Harter tager udgangspunkt i den individuelle bruger og vil ud fra empiriske undersøgelser forsøge at finde generaliserbare mønstre i relevansbedømmelserne. Harter ønsker brugernes kognitive forandringer nærmere undersøgt i forbindelse med begrebet om psykologisk relevans, mens Schamber et al. er særligt interesserede i at kortlægge relevanskriterier.

Også Froehlich interesserer sig særligt for relevanskriterier, men han mener dog at det individuelle aspekt tillægges for megen vægt; man bør i stedet fokusere på den prototypiske bruger al den stund at brugerne er medlemmer af kulturelle og sociale fællesskaber og dermed dømmer ud fra en forståelseshorisont de til en vis grad er fælles om. I forhold til dette går Hjørland et skridt videre; for ham drejer det sig ikke primært om brugerne som enkeltpersoner, men som kollektive udøvere af en videnskab eller en faglig disciplin der i sig selv har bestemte, objektive erkendelsesinteresser. Hjørlands tildeler således, som den eneste af de nævnte, relevansbegrebet objektivitet.

Både Harter og Schamber fremhæver relevansbegrebet som dynamisk: relevans er ikke statisk, men udvikler sig i takt med brugerens kontakt med informationssystemet og søgningens resultater, og også Froehlich karakteriserer processen som dialektisk. For Hjørland består det dynamiske i den kollektive erkendelsesudvikling inden for det enkelte fagområde, og dermed også udviklingen i erkendelsesinteresser og relevanskriterier.

Det er desuden tydeligt at jo mere der fokuseres på den enkelte bruger, desto mindre betydning tillægges kriteriet emnemæssig beslægtethed. Således fraskriver Harter, som det ene yderpunkt,

næsten fuldstændig kriteriet betydning, mens Hjørland, som det andet, bygger sin teori op omkring en forståelse af netop emnebegrebet. Punktet forekommer at være meget centralt, og man kunne ønske en mere indgående forståelse af dette kriterium også på det subjektive niveau, f.eks. en nærmere undersøgelse Froehlich's idé om polyvalens.

For Harter er selve det kognitive mål at skabe relevans - dvs. at få udefra kommende antagelser og fænomener til at harmonere med personens kontekst for at opnå den størst mulige effekt. I denne proces er det ligeså ofte konteksten der vælges eller tilpasses, som det er de udefra kommende påvirkninger der vælges til eller fra. Schamber et al. taler tilsvarende om at skabe mening, og den relevante information er netop den der hjælper til at skabe mening i konteksten. - På trods af at de bruger forskellige ord, synes det i høj grad som om det er de samme forhold de to synspunkter refererer til.

At Froehlichs og Hjørlands teorier er beslægtede, påpeger de selv. Froehlich (1994, side 131) henviser således til Hjørlands teorier, men han påpeger dog forskellen mellem Hjørlands fokus på faglige discipliner i sig selv og hans egen idé om at fokusere på den prototypiske bruger. Omvendt citerer Hjørland og Albrechtsen (1995) den netop nævnte artikel af Froehlich som støtte for deres synspunkter.

På trods af lighederne og på trods af at de forskellige teorier i nogen udstrækning kan passes ind i det samme "verdensbillede", vil der dog være stor forskel på hvordan man i praksis ønsker at prioritere kræfter og midler til forskning og udvikling; om man f.eks. finder der skal sættes på kvalitative, psykologisk orienterede brugerundersøgelser, eller om det er vigtigere at kortlægge domæner og videnskabelige discipliners faglige grundlag. De forskellige perspektiver og niveauer burde dog kunne berige hinanden og samlet føre til en større empirisk og teoretisk indsigt i området, en indsigt som bl.a. kan benyttes ved udvikling af informationssystemer.

5. Relevans som kvantitet? - en diskussion om metoder

Relevans har som tidligere nævnt traditionelt været brugt som målestok, bl.a. til at evaluere informationssystemer og IR-mekanismer som f.eks. indekseringsprog. Diskussionen om relevans som kvantitet er således også en diskussion om metoder: hvordan undersøger og evaluerer man informationssystemer?

Bag ved anvendelsen af relevans som evalueringsmål ligger der mere eller mindre implicit en række antagelser, bl.a. følgende:

- Informationssystemers mål er at tilfredsstille brugernes informationsbehov ved at finde relevant information.
- Informationssystemers kvalitet kan måles på hvor megen relevant information de giver mulighed for at finde (recall), og hvor stor del af den fundne information der er relevant (precision).
- Relevans kan måles kvantitativt.

Jeg vil i det følgende se lidt nærmere på disse antagelser, med afsæt i bl.a. den kritik der har været rejst af Ellis (1996) og Harter (1996).

Informationssystemers mål er at tilfredsstille brugernes informationsbehov ved at finde relevant information. Sætningen virker indlysende indtil det trivielle, men også uklar: relevant i forhold til hvad? Det er ikke ligegyldigt om man tager udgangspunkt i f.eks. en psykologisk-kognitiv definition af relevans, eller om synsvinklen er hermeneutisk eller sociologisk-videnskabsteoretisk.

Man kan i øvrigt stille spørgsmål ved om denne målsætning er rammende og udtømmende. Buckland anfører således et citat fra den østrigske ærkehertug Johann (1782-1859) som et stadigt gyldigt formål for informationssystemer:

»give sense to its collections, make learning easier, stimulate curiosity, help avoid pure memorizing wich is so detrimental to indepen-

dent thinking and self-reliance, and fill more and more the harmful gap between formal concepts and intuition, theory and practice» (Buckland, 1991, side 36).

Også Hjørland anfører - som tidligere nævnt - at målet for informationssystemer er bl.a. at skabe og udvikle informationsbehov (1993, side 173). Disse formuleringer angiver et langt bredere formål, og en evaluering i forhold hertil lægger op til metoder med inspiration i pædagogik snarere end i naturvidenskab.

Informationssystemers kvalitet kan måles på hvor megen relevant information de giver mulighed for at finde, og hvor stor del af den fundne information der er relevant.²⁷ Er recall/precision et fyldestgørende mål for kvaliteten af et informationssystem? Det giver mulighed for at sætte kvaliteten på en formel og udtrykke den i enkle forhold. Har man først fået udledt tallene, har man et forjættende enkelt middel til at karakterisere og sammenligne kvaliteten af informationssystemer på tværs af fagområder, geografi, tid, brugerkræds osv.

Ellis stiller dog spørgsmål ved om det i det hele taget er relevans man måler, og han sætter begreberne på plads:

“(...) the problem of measurement in information retrieval research is not really that of measuring relevance but of employing relevance as a measure, or more accurately, as a criterion of measurement. The ratios employed as measures, such as recall, precision, and fallout, are not measures of relevance, they are measures derived from the use of relevance as a criterion of measurement” (Ellis, 1996, side 25).

Recall/precision-forholdet er en stærk forenkling af den mangfoldighed af kvaliteter et informationssystem kan have, fx: Er det brugervenligt (og da: for hvilke typer af brugere og hvilke typer af informationsbehov)? Let at manøvrere i? Giver det mulighed for at bruge forskellige søgemetoder? Letter det relevansvurderingen for brugeren (hvilke rele-

vanskriterier understøtter det)? Og med Buckland: Stimulerer det nysgerrighed og selvstændig tænkning? Og Hjørland: Skaber og udvikler det informationsbehov?

Relevans kan måles kvantitativt. Hvis målet skal baseres på relevans, skal begrebet være operationelt, og det skal kunne kvantificeres: det skal have en vis grad af pålidelighed, dvs. det skal kunne gøres stabilt, uafhængigt og konsistent. Men det skal samtidig afspejle den teoretiske definition af relevans. Kan dette overhovedet lade sig gøre?

For at svare på dette kan man enten fastholde at man vil tage udgangspunkt i en kvalitativ definition af relevans som subjektivt, eller man kan tage udgangspunkt i en forståelse af begrebet som objektivt.

5.1 Subjektivt udgangspunkt

Med udgangspunkt i psykologisk og situationel relevans kan man vælge at fastholde det individuelle og subjektive: at en relevans er et forhold mellem en brugers informationsbehov og et dokument i en given situation, på et givet tidspunkt. For at opnå et operationelt begreb kan man så forsøge at foretage en *normalisering* af brugerens oplevelse af relevans, dvs. udlede en slags gennemsnit af flere brugeres relevansvurdering, f.eks. til brug i testbaser. Ørslykke Møller og Kidmose (1995, side 51-52) har et konkret og teknisk set ret detaljeret forslag til hvordan dette kan gribes an.

For at opnå dels en realistisk situation, dels konsistens i forhold til tidsdimensionen, foreslår de “en litteratursøgning, der sker ud fra en skriftligt stillet opgave på en uddannelsesinstitution, (...) som kunne gentages senere og stadig være realistisk i forhold til brugerne” (ibid, side 51-52). På baggrund af denne litteratursøgning udarbejdes en testpulje af såkaldt tilnærmet objektivt relevante dokumenter, dvs. “dokumenter, som på tværs af brugerne i undersøgelsen indeholder eksempelvis fem relevanskriterier eller mere” (ibid, side 51).

Analyserne af relevansbedømmelsen skal være trindelte og de skal foregå på forskellige stadier i litteratursøgningsprocessen. Ørslykke Møller og Kidmose foreslår højtænkning, interview og observation via video som baggrund for analysen. Ved undersøgelser af informationssystemer skal man sammensætte brugergrupper der såvidt muligt matcher hinanden, og brugergrupperne skal desuden være relativt store og homogene.

En sådan undersøgelse vil imidlertid være temmelig dyr at foretage, idet den kræver kvalitative undersøgelser (højtænkning, interview, og videooptagelser) for en relativt stor brugergruppe. Dette skal ses i forholdt til udbyttet, som angiveligt er af begrænset gyldighed. Det er heller ikke umiddelbart indlysende hvordan man kan anvende resultaterne fra kvalitative undersøgelsesmetoder til kvantificering. Metodisk savnes der også en afklaring af hvordan man sammensætter brugergrupper der matcher hinanden; det ville kræve en meget nøje afdækning af sammenhæng mellem identificerbare karakteristika ved brugerne, de relevanskriterier de benytter, og måden de benytter dem på.

Harter kalder denne form for normalisering for *pooling*, og han kritiserer metoden for ikke at udnytte de værdifulde informationer om individuelle forskelle i søgning og relevansvurdering der opnås undervejs:

“Pooling and averaging retrieval results is an important component of the traditional IR evaluation model. However, it is ill-suited to a user-oriented research agenda, because it obliterates the environmental differences among users, queries, and their associated relevance judgments. Equally important, a pooled approach to evaluation erases differences in how the systems being compared deal with these variations” (Harter, 1996, side 46).

Ellis kritiserer selve grundlaget for anvendelsen af relevans hvis man tager udgangspunkt i en kognitiv opfattelse:

“But the problem would then become how to measure changes in knowledge structures in response to retrieved information and it is not clear how this can be done, if it can be done at all. But if relevance is maintained as the measure, then there is a schism between the theoretical assumptions of the overall research design and the form of measurement adopted in the evaluation” (Ellis, 1996, side 33).

5.2 Objektivt udgangspunkt

Swanson (1986) argumenterer for at skelne mellem subjektiv og objektiv relevans. Den subjektive relevans er et mentalt fænomen, det er en vurdering foretaget suverænt af den enkelte bruger, og forholdet er “between a document and a person” (side 390) - dvs. personens informationsbehov. Den subjektive relevans hører hjemme i fungerende informationssystemer, hvor det gælder om at tilfredsstille brugernes behov: “(...) whatever the requester says is relevant is taken to be relevant; the requester is the final arbiter” (side 390).

Swanson finder imidlertid at der er behov for en objektiv målestok når det gælder design og test af informationssystemer, dvs. en objektivisering af relevans. Han foreslår at man til det formål tager udgangspunkt i brugerens spørgsmål som det er fremsat. Et brugerspørgsmål der er fremsat, f.eks. skriftligt, kan på denne måde objektiviseres; det er gjort til objekt for fortolkning af andre end den som har fremsat det, og det ligger derfor åbent for argumentation og dermed for relevansvurderinger i forhold til dokumenter - præcis på samme måde som dokumenterne er objektiviserede.

Indholdet af både forespørgslen og dokumenterne kommer på den måde til at tilhøre det Popper kalder Verden 3: den objektive viden. Med et citat fra Wilson anfører Swanson at relevans kan være til stede uden at man i en given situation er i stand til at indse det, dvs. at relevans på denne måde får en form for objektiv eksistens i forhold til et brugerspørgsmål. Swanson arbejder altså tre med objektiviserede elementer: brugerspørgsmålet, dokumen-

tet og relevansvurderingen, alle tre åbne for fortolkning og argumentation.

"(...) perhaps we should be suspicious of a definition that declares as unarguable something that people argue about all the time - that is, about whether some document is or is not relevant to a particular request. That such disputes occur suggests that there must be another kind of relevance - a kind that one can argue about, that is open to criticism, and that therefore must have some validity or meaning independent of the requester - thus relevance that is objective rather than subjective" (Swanson, 1986, side 391).

Også Hjørlands teori indeholder tre objektiviserede elementer: erkendelsepotentialer, erkendelsesinteresser og relevans mellem disse. En objektiv analyse af disse vil være en analyse der er "mest mulig i overensstemmelse med sit objekt" (Hjørland, 1993, side 39), og ikke farvet af den person der foretager analysen.

Spørgsmålet er dog om man kan komme uden om det subjektive element i fortolkning og vurdering - selvom det er hensigten. Nogen - hvor veluddannede og velmenende de end måtte være - skal fortolke og vurdere, og det må nødvendigvis ske ud fra deres forudsætninger og baggrund, altså ud fra deres kontekst og horisont i vid betydning. Det drejer sig altså ikke blot om at en eller flere personer skal kunne aflæse objektive eller mekaniske måleinstrumenter som i naturvidenskaberne, men om at denne eller disse personer *er* måleinstrumentet.²⁸

Denne form for objektivisering blev bl.a. forsøgt anvendt i ASTIA-Uniterm-testene i 1953²⁹, hvor to grupper vurderede dokumenternes relevans i forhold til virkelige brugerspørgsmål. Undersøgelsen mislykkedes angiveligt pga. spørgsmålet om relevansvurdering:

"It is significant that not only were [the two groups] not able to agree on their relevance assessments, but also that there was no decision

procedure they could appeal to resolve their differences (...)" (Ellis, 1996, side 25).

Det subjektive element er her flyttet fra brugeren og brugerspørgsmålet til de personer der foretager relevansfortolkningen. Man kan argumentere om fortolkningen, men den kan ikke tænkes løsrevet fra et subjekt der fortolker, og kan der ikke opnås enighed blandt flere fortolkere, er der ingen objektiv procedure til at løse dette.

5.3 Subjektivitet kontra målbarhed

På baggrund af de kvalitative forståelser af relevans der er redegjort for i afsnit 4, må der stilles spørgsmål ved om recall/precision, udledt på baggrund af relevansvurderinger, kan anvendes ved evaluering af informationssystemer. Relevans er for mangfoldigt og variabelt et begreb, og selv mere objektive tilgange undgår ikke det subjektive element i fortolkning og vurdering.

Ellis konkluderer således:

"The dilemma of measurement has remained intractable even given the different ontological and methodological assumptions of the archetypal, probabilistic, and IR-ES [Information Retrieval-Expert System] approaches. The reason for this intractability deriving from: (1) The nature of the subject matter of the field; (2) the nature of relevance judgment; and (3) the nature of cognition and knowledge" (Ellis, 1996, side 33).

Og senere:

"However, while it is feasible to describe or analyze changes in knowledge in response to new information qualitatively, to attempt to do the same quantitatively seems to have no tenable theoretical or practical foundation and represents a similar unsustainable research goal" (Ellis, 1996, side 33).

Det betyder ikke at man må opgive at evaluere

informationssystemer og at sammenligne evalueringerne, men at man må udvikle mere målrettede, komplekse, varierede og kvalitative metoder der i højere grad tager hensyn til og stemmer overens med den teoretiske forståelse af informationsøgning, herunder af relevans.

6. Konklusion og sammenfatning

Relevans kan grundlæggende forstås som et forhold der indebærer en fortolkning, og forholdet kan defineres mellem en række forskellige grundelementer i IR. Alt afhængigt af hvilke elementer der indgår i forholdet, og hvordan man vælger at forstå og vægte fortolkningsprocessen, implicerer man forskellige videnskabelige tilgange til begrebet og får dermed også vidt forskellige resultater af undersøgelser af relevans. De seks forskellige relevansforhold jeg har identificeret og beskrevet ovenfor, afspejler en række af de centrale problemstillinger og forskningsområder i IR. Det er derfor vigtigt, både i forhold til teoretiske analyser, forskningsmæssige undersøgelser og anvendelse af relevansbegrebet til mere praktiske formål, at gøre sig helt klart hvilke relevansforhold og hvilke aspekter af begrebet, man behandler.

Forskellige nyere teoretiske tilgange til relevans behandler begrebet ud fra henholdsvis en psykologisk-kognitiv, en hermeneutisk og en sociologisk-videnskabsteoretisk synsvinkel. Disse teorier er dog ikke så væsensforskellige at de ikke kan opfattes som forskellige aspekter af det samme fænomen. Det kvalitative indhold af relevans i IR kan på baggrund af de fire teorier fastslås som et dynamisk, stærkt variabelt begreb, der på det psykologiske niveau er stærkt subjektivt, men på det sociale niveau kan nærme sig en vis objektivitet, samtidig med at kriteriet »emnemæssigt beslægtet« gør sig stærkere gældende.

Relevans har traditionelt været brugt som mål ved evaluering af informationssystemer. Dette kræver imidlertid en kvantificering af begrebet, der synes vanskeligt forenelig med begrebets kvalitative indhold. Vejen ud af dette kunne være at se nærmere

på de implicite antagelser der ligger til grund for anvendelsen af relevans til dette formål, og overveje om ikke evalueringsmetoder inspireret af pædagogik og psykologi kunne vise sig mere frugtbare.

Noter

1. Se f.eks. Politikens Nudansk ordbog (1991) og Gyldendals Fremmedordbog (1987).
2. Random House Webster's Dictionary (1992), her efter Froehlich (1994).
3. Gyldendals Engelsk-dansk ordbog (1988).
4. Som refereret af Saracevic (1975) og Schamber et al. (1990)
5. Som bl.a. refereret af Saracevic (1975) og Schamber et al. (1990).
6. Jf. f.eks. Coopers definition af utility som refereret af Saracevic (1975, side 334).
7. Se Ingwersen (1992, side 116 ff) og Taylor (1968).
8. Precision blev oprindeligt kaldt relevance (Saracevic, 1975, side 327 - note),
9. Se Ellis, 1990, side 2.
10. Se Ellis, 1996, side 27.
11. Faktorer der påvirker relevansdomme er især undersøgt i to større forskningsprojekter, begge fra 1967: Cuadra & Katter og Rees & Schultz.
12. Relevanskriterier har især været i fokus for nyere projekter, bl.a. Park i 1993 og Barry i 1994.
13. Se f.eks. Schamber et al. (1990).
14. Hjørland (1995, side 417) anfører at den almindelige oversættelse af topic er emne, men argumen-

terer for at en oversættelse med område (f.eks. forskningsområde) er mere præcis. Jeg finder dog at emnemæssigt beslægtet bedst dækker den brug af topicality som både Harter og Froehlich gør af begrebet.

15. Kontekst defineres her ikke nærmere, men begrebet har tydeligvis ikke præcis det samme indhold som i Harters definition. Politikens filosofi leksikon (1983) anfører at begrebet (foruden en tekstlig sammenhæng) betegner en sammenhæng som omfatter "alle relevante træk ved den situation, som den talende eller skrivende befinder sig i, dvs. alle træk, som er medbestemmende for, hvilket indhold en fremsat ytring har". Denne afgrensning harmonerer ikke helt med den måde Schamber et al. bruger begrebet på; her drejer det sig nærmest om sammenhæng i vid betydning.

16. Schamber et al. bygger her på en teori udviklet af Dervin i 1983.

17. Froehlich bruger begrebet naturlig kategori som det er udviklet af Rosch i 1978 og Lakoff i 1987.

18. Froehlich redegør her ikke nærmere for hvad han forstår ved "topicality" - emnemæssig relevans. Jf. note 14.

19. Horisont er i den moderne hermeneutik "rækken af de underforståede begreber og teorier, som farver min udlægning af det værende og den verden, jeg befinder mig i" (Politikens filosofi leksikon, 1983).

20. Froehlich's brug af kontekst svarer til definitionen angivet i Politikens filosofi leksikon (1983), se note 15.

21. Hjørland (1993) oversætter i det engelske resumé af afhandlingen emne med subject, og han angiver på denne måde en forskel til begrebet topic. Se også note 14. Det fører imidlertid for vidt her at forfølge disse begrebers betydningsvariationer yderligere.

22. Her efter Hjørland & Albrechtsen, (1995, side 401).

23. Her nærmest sammenhæng. Jf. note 15.

24. Jf. også side 172 ff (Hjørland, 1993).

25. Ibid, side 185.

26. Disse begreber bruges også af Hjørland (1995, side 360-361).

27. Se f.eks. Martyn & Lancaster (1981, side 172).

28. Jf. Ellis (1996).

29. Som refereret af Ellis (1996).

Litteratur

Buckland, M. (1983). Relatedness, relevance and responsiveness in retrieval systems. *Information Processing & Management*, 19(3), side 237-241.

Buckland, M. (1991). *Information and information systems*. New York, Praeger.

Ellis, D. (1990). *New Horizons in Information Retrieval*. London, Library Association.

Ellis, D. (1996). The Dilemma of Measurement in Information Retrieval Research. *Journal of the American Society for Information Science*, 47(1), side 23-36.

Froehlich, T. J. (1994). Relevance reconsidered - towards an agenda for the 21th century: introduction to special topic issue on relevance research. *Journal of the American Society for Information Science*, 45(3), side 124-133.

Harter, S. P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science*, 43(9), side 602-615.

Harter, S. P. (1996). Variations in Relevance Asses-

sments and the Measurement of Retrieval Effectiveness. *Journal of the American Society for Information Science*, 47(1), side 37-49.

Hjørland, B. (1993). Emnerepræsentation og informationsøgning: Bidrag til en teori på kundskabs-teoretisk grundlag. Göteborg, Bibliotekshögskolan & Centrum för biblioteks- og informationsvetenskap.

Hjørland, B. (1995). Informationsvidenskabelige grundbegreber (2. udgave). København, Danmarks Biblioteksskole.

Ingwesen, P. (1992). *Information Retrieval Interaction*. London, Taylor Graham.

Martyn, J. & Lancaster, F. W. (1981). *Investigative Methods in Library and Information Science: an introduction*. Arlington, Information Resources Press.

Saracevic, T. (1975). Relevance: A framework for the thinking on the notion in information science. *Journal of the American Society for Information Science*, 26(6), side 321-343.

Schamber, L., Eisenberg, M. B. & Nilan, M. S. (1990). A re-examination of relevance: Toward a dynamic, situational definition. *Information Processing & Management*, 26(6), side 755-776.

Swanson, D. R. (1986). Subjective versus objective relevance in bibliographic retrieval systems. *Library Quarterly*, 56(4), side 389-398.

Taylor, R. S. (1968). Question negotiation and information seeking in libraries. *Coll. & Res. Libraries*, 29, side 178-194.

Ørslykke Møller, C. & Kidmose, B. (1995). Relevant om relevans - teori og metodologi. *Biblioteksarbejde*, 16(45), side 37-53.