

Relevant om relevans

- teori og metodologi

Af Christian Ørslykke Møller og Birger Kidmose

Introduktion

Hensigten med herværende artikel er at præsentere relevansbegrebet i dets traditionelle form, samt at skitsere nogle af de tilgange til relevansbegrebet indenfor informationsvidenskab, som har været fremherskende i nyere tid. Listen omfatter emnemæssige, brugerorienterede, domæneanalytiske og hermeneutiske synsvinkler, som udgør nogle af de mest citerede forskeres bud på problematikken. Vi vil studere resultaterne af de seneste empiriske undersøgelser, ligesom vi vil diskutere grundlæggende metodiske problemer i relevansforskningen.

Relevansbegrebet viser sig ved nærmere eftersyn at være behæftet med langt større definitionsproblemer end man umiddelbart skulle tro. En del af definitionsproblematikken har rod i diskussion om kontekstafhængighed og om, hvorvidt relevans er situationelt betinget. Vi vil forsøge at trække nogle

linier, ligesom vi vil give et kortfattet bud på en problematisering af de enkelte retninger. Vi vil bestræbe os på at forholde os kritisk, uden dog at gå mere i dybden end artiklens rammer tillader. Endelig vil vi kort opridse rammerne for vores eget syn på relevansbegrebet, samt søge at opridse rammerne for et brugbart undersøgelsesdesign.

1. "The Systems View"/"Topicality"

Relevans er og har altid været den grundlæggende forudsætning for evaluering af IR-systemers effektivitet. En egenskab, som ifølge Schamber (1994) gør det til det centrale begreb i informationsvidenskaben. I de tidligste undersøgelser opfattede man relevans som et meget håndgribeligt fænomen, som noget måleligt, der kunne bedømmes objektivt. Denne opfattelse, som vi og andre med os har valgt at betegne "The Systems View" (Schamber 1990), er helt fundamental, og dens karakteristika

og mangler er kilden til al senere diskussion om relevans.

"The Systems View" kendetegnes primært ved at opfatte relevans som et udtryk for relationen mellem en forespørgsel og en dokumentrepræsentation. Forholdet mellem et dokument emne og en forespørgsel er ofte blevet beskrevet som et spørgsmål om, hvorvidt et dokument er "on the topic", hvorfor man traditionelt har refereret til denne opfattelse som "topicality". Der opereres med enten exact eller partial match mellem termer i forespørgsel og dokument. Det forudsættes, at brugerens forespørgsel er et præcist udtryk for dennes informationsbehov, og modsvarende, at dokumenttermer er et tilstrækkeligt udtryk for dokumentets emnemæssige indhold. Såfremt dette er tilfældet, vil det være muligt at tale om en umiddelbart observerbar, objektiv relevans, som kan bruges som måleredskab ved systemevaluering. Barry udtrykker det således:

"The system matches subject terms that comprise document representations and queries. Evaluation is then based on the systems ability to retrieve references to documents that are relevant within the users current information need situation (Barry, 1994, s. 150)".

Ifølge Schamber (1994) kan man tale om "The Systems View", når ens studier er centreret om tekniske input/outputfunktioner og interne matchingfunktioner *uden hensyntagen til menneskelig bedømmelse*. Det er netop i dette forbehold, vi finder nøglen til de problemer, som "The Systems View" er behæftet med. Park siger:

"Topical relevance is context-free and is based on fixed assumptions about the relationship between a topic of a document and a search question, ignoring an individuals particular context and state of needs. It is an undimensional view of users' information problems, disregarding the changing nature of the individuals information problem and its subsequent impact on the search (Park, 1994, s. 136)".

2. Brugerorienterede synsvinkler

2.1 Baggrunden for den brugerorienterede tilgang

Tidlige eksperimenter henledte forskeres opmærksomhed på, at relevansbedømmelse kunne være influeret af andre faktorer end de rent emnemæssige. Denne erkendelse førte ifølge Barry (1994) til omfattende forskningsprojekter. Cuadra & Katter og Rees & Schultz udførte begge midt i tresserne en række eksperimenter, som anses for de mest centrale blandt forskere som eksempler på "brugeropmærksomme" forsøg. Som baggrund for undersøgelserne skriver Cuadra & Katter (1967), at relevansbedømmelser hidtil havde været opfattet, som om de oprandt fra en "sort kasse", hvis hemmeligheder ikke interesserede nogen, så længe den udfyldte sin funktion tilfredsstillende. D.v.s. så længe den forsynede forskerne med nogle resultater, som kunne fungere som måleredskab ved systemevaluering. Og videre, at denne opfattelse ville have været meget succesfuld, hvis bare man kunne bevise, at relevansbedømmelser ikke lod sig påvirke af andre variable end forholdet mellem dokument og forespørgsel. "Desværre" var der, som Cuadra & Katter også gør opmærksom på, allerede på daværende tidspunkt, et voksende antal beviser for, at det modsatte er tilfældet.

Tilsammen identificerede Cuadra & Katter og Rees & Schultz op imod 80 forskellige faktorer, som alle havde indflydelse på relevansbedømmelse hos den enkelte bruger. Resultaterne af undersøgelserne bliver ifølge Barry (1994) opsummeret af Rees & Saracevic i en række udgangspunkter, som efter vores mening faktisk viser sig at være dækkende for størstedelen af nyere relevansdiskussioner og tilgange:

Disse antagelser blev startskuddet til centrale diskussioner om hvilke faktorer, som burde indgå i et brugbart relevansbegreb. Debatten har dels drejet sig om en indkredsning af de faktorer, som indvirker på brugeres relevansbedømmelse og dels om, hvorledes man kan operationalisere disse faktorer med henblik på systemdesign og -evaluering.

"A sharp distinction can be made between relevance to a question and relevance to the need underlying a question;

Only the user himself can judge the relevance of the document to this need; i.e., the relevance judgment is subjective;

For [a] user a relevance judgment will change; i.e., it is not constant;

Documents 'inform' various people differently, i.e., the cognitive set has to be taken into account in relevance judgment;

Relevance is a comparative rather than a qualitative concept; i.e., it is not inherent to a document; and

Various types of judgments exist because of the different purposes for which information is required"

Rees & Saracevic I: Barry, 1994, s. 150

2.2 Imod en brugerorienteret, kognitiv tilgang

Schamber (1994, s. 7) bruger betegnelsen "The Information View" om den retning indenfor relevansforskning, som følger efter "The Systems View" og som markerer overgangen til en brugerorienteret tilgang.

Ifølge Chamber opfattes relevans under denne synsvinkel som en subjektiv bedømmelse af det konceptuelle forhold mellem en forespørgsel og et dokument. "The Information View" er kendetegnet ved, at en intermediær kan relevansbedømme på vegne af brugeren. Tankegangen er, at der findes en fælles offentlig viden, som både intermediæren og brugeren har adgang til og som danner basis for en tilnærmet objektiv enighed om, hvad der er relevant. Vi ser senere Hjørland (1992) tale om en fælles viden, som dog hos ham er begrænset til det enkelte videnskabelige domæne. Problemet ved denne retning i relevansforskningen er, at kun brugeren har adgang til sin helt private viden, som

influerer på dennes vurderinger. Variationen i den enkelte brugers vidensstruktur medfører, at en tilnærmet objektivitet, og dermed enighed, mellem bruger og intermediær ikke er mulig (Schamber, 1994, s. 8). Vi mener endvidere, at det er problematisk, at der under "The Information View" ikke tages tilstrækkeligt hensyn til den enkelte brugers sociale og faglige kontekst.

Under den brugerorienterede paraply er der en del relevansopfattelser, som tager udgangspunkt i den enkelte brugers vidensstruktur og "problem-space", før og under mødet med informationssystemet. Disse synspunkter karakteriserer vi i det følgende som kognitive, idet der fokuseres på brugerens kognitive struktur.

Ingwersen (1991) repræsenterer en pragmatisk og brugerorienteret tilgang til relevansbegrebet. Hans mediatormodel (1991, s. 204) fortolker vi som et eksempel på "The Information View", idet systemet ved hjælp af evaluerings- og feedback-faciliteter og brugermodellering foretager relevansbedømmelse på vegne af brugeren. I forhold til en teoretisk relevansdefinition forholder Ingwersen sig pragmatisk. Han siger :

"Hence, relevance is ultimately a value of pragmatic nature, linked to the individual user's problem space and state of knowledge (Ingwersen, 1991, s. 54)".

Med dette udgangspunkt må Ingwersens synspunkt karakteriseres som kognitivt, samtidig med, at relevansbegrebet hos Ingwersen er baseret på de emnemæssige træk ved dokumentrepræsentation og relevansen måles i graden af overensstemmelse mellem dokumentet og det aktuelle "information requirement". Relevansbedømmelsen er hos Ingwersen centreret omkring det problematiske i, at IR systemer aldrig kan opnå 100% recall, men må acceptere, at der er en "Dark Matter" vi aldrig kan få fuld erkendelse af. Denne drejer sig om mængden af potentielt relevant information, som efter alle solemærker at dømme befinder sig et sted i informationssystemet.

"The searcher, the IR system, and the IR researcher, 'does not know what he does not retrieve' - and will not know it (Ingwersen, 1991, s. 54)".

Denne usikkerhed er efter Ingwersens mening et universelt princip i IR. Fra det kognitive synspunkt kan "Dark Matter" antage en explicit såvel som en implicit form, i.e. explicit relevant information kan unddrage sig genfindning p.g.a. dårlige søgemuligheder eller lign. Endvidere kan information være implicit tilstede i et dokument og muligheden for at "læse" informationen ud af dokumentet afhænger af modtagerens vidensstruktur. Det er vores opfattelse, at "Dark Matter"-problematikken ikke er af signifikant natur i forhold til spørgsmålet om relevansbedømmelse, idet vi senere vil argumentere for et multidimensionalt, dynamisk relevansbegreb, hvor et begreb som "Dark Matter" bliver af mere perifer betydning. Baggrunden herfor er, at hvis man antager at informationsbehovet er dynamisk og konstant skiftende, vil "Dark Matter" ligeledes ændre sig konstant, og vil i praksis være umulig at sige noget endegyldigt om.

Ingwersen fokuserer på den individuelle brugers kognitive struktur. Det er her, han tager sit udgangspunkt, mens han samtidig åbner op for en række andre relevansfaktorer, bl.a. ser han ingen modsætning mellem sin relevansopfattelse og indtænkning af Schamber et al.'s multidimensionale relevansbegreb, som vi senere vil redegøre yderligere for.

2.3 Fokus på situationelle og dynamiske faktorer

Den situationelle tilgang leder hen mod en dynamisk opfattelse af relevans og begrebet bliver dermed multidimensionalt. Tilgangen betoner informationssøgeprocessens og den enkelte brugers informationsbehovs kontekstafhængighed. Den afføder en pragmatisk tænkemåde, da problemerne med at operationalisere sin relevansbedømmelse bliver stadigt større, jo flere dimensioner ved begrebet, der skal tages højde for.

Schamber (1994) beskriver overgangen fra "The Information View" til "The Situation View" som glidende, hvor den vigtigste skelnen er, at indenfor "The Situation View" er det kun brugeren, der kan foretage en adækvat relevansbedømmelse af information. Relevansbedømmelsen sker med udgangspunkt i brugerens aktuelle problem situation. I forbindelse med denne overgang støder vi på en regn af begreber, som hver især bidrager til at nuancere (og forplumre) billedet af relevansbegrebet. Schamber(1994) og Hjørland(1992) nævner en række: *Pertinence* som betegner en opsplitning af relevansbegrebet i dels relevans, som her indikerer forbindelsen mellem dokumentrepræsentation og brugerspørgsmål, og dels pertinence, som udtrykker forbindelsen mellem dokumenterne og informationsbehovet. Et andet begreb er *Utility*, som er udtryk for en pragmatisk opfattelse, som koncentrerer sig om dokumentets anvendelighed. Nært beslægtet hermed er *usefulness*, som ligeledes betegner dokumentets nytteværdi for brugeren og *satisfaction*, som opfatter behovstilfredsstillelse som nøgleordet. Fælles for denne højt kvalificerede begrebsnuancering (forvirring?) er et forsøg på at tilnærme sig et pragmatisk, brugerorienteret relevansbegreb i frustration over, at de resultater man nåede frem til under "The Systems View" betød store problemer i forhold til operationalisering i den meget pragmatiske og teknologiorienterede informationsvidenskab.

Det er vigtigt på dette tidspunkt at gøre opmærksom på, at de ovenfor skitserede brugercentrerede tilgange *ikke repræsenterer en afstandtagen til det emnemæssige aspekt ved "The Systems View"*. Der er tale om en udvidelse af relevansbegrebet, samt en anden prioritering af hvilke områder indenfor relevans, der skal fokuseres på. De emnemæssige dimensioner ved relevans er stadig vigtige, især i forhold til de brugerdefinerede relevanskriterier, men mere herom senere, når vi ser på de nyeste resultater indenfor relevansforskningen. Efter vores mening viser den situationelle tilgang sig at være et af de vigtigste træk ved den brugerorienterede tilgang, da det medfører, at man i forbindelse med

relevans må inddrage dynamiske og kontekstuelle faktorer.

Et godt eksempel på en decideret kognitiv tilgang, som er stærkt fokuseret på situationelle og dynamiske faktorerers indflydelse på brugerens informationsbehov, og som drager temmelig vidtrækkende konsekvenser, er den opfattelse, som præsenteres af Stephen P. Harter (1992), som kalder sit synspunkt "psykologisk relevans". Harters teori bygger på de to amerikanske psykologer, Sperber & Wilsons ideer om relevans i forbindelse med verbal kommunikation på hverdagsplan. Deres opdagelser og konklusioner bliver af Harter implementeret direkte i IR, og herudfra danner han sit meget dynamiske og situationelt orienterede relevansbegreb. Et af Harters udgangspunkter er en ret utraditionel stillen spørgsmål ved den antagelse, at det som en bruger er interesseret i og vil være mest tilfreds med, er et dokument omhandlende det emne, han formulerede i sin forespørgsel. Harter påstår, at denne antagelse er helt uden empirisk grundlag. En påstand som vi iøvrigt tillader os at betvivle, som vi senere vil komme ind på.

Harters bud på brugerens behov er, at brugeren ønsker information, som forårsager en ændring i dennes nuværende kognitive struktur. Information som opfylder denne betingelse er, ifølge Harters adopterede relevansopfattelse at betragte som relevant.

Vi kan kort skitsere Harters teoris betydning for informationssystemer ved at begynde med hans citat fra Wilson & Sperber:

"A phenomenon is relevant to an individual to the extent that the contextual effects achieved when it is optimally processed are large (p. 153) (Wilson & Sperber, 1986 I: Harter, 1992, s. 606)".

Fænomen er her at forstå som dokument eller dokumentrepræsentation og helt centralt: Den enkeltes informationsbehov opfattes som dennes til ethvert tidspunkt givne kognitive stade/struktur

eller kontekst. Det essentielle er, at Harter fokuserer på den *nuværende* kontekst (ell. kognitive stade/struktur), i stedet for på informationsbehovet som det så ud *ved starten* af den aktuelle søgning. Vi ser, at der er tale om en ekstremt dynamisk relevansopfattelse, idet brugerens informationsbehov er i konstant bevægelse, efterhånden som søgningen skrider frem i.e., en dokumentrepræsentation forårsager en kontekstuel ændring og søgningen, og dermed informationsbehovet, skifter retning. Harters holdning indebærer, at brugerens oprindelige forespørgsel umuligt kan være et adækvat udtryk for dennes informationsbehov. Harter siger:

"My interpretation of information need suggests that we will need to be satisfied with a superficial statement of an evolving information problem, recognizing that such a statement must be grossly inadequate to predict what documents will and will not be found relevant by any given individual at a given time (Harter, 1992, s. 607)".

Interessant er det at se, at Harter ser meget kritisk på folk som Belkin og Taylor, som hver for sig taler om et informationsbehov, der stammer fra et oprindeligt kognitivt stade; Belkin om ASK, Taylor (1968) om Q₁ og Q₂. For Harter ophører det oprindelige kognitive stade/informationsbehov med at være interessant i det øjeblik den første kontekstuelle ændring indtræffer i brugerens kognitive struktur. Harter nævner selv, at hans relevansbegreb er nært beslægtet med den definition, som Buckland anfører for information-som-proces; information er selve processen at blive informeret - at finde noget relevant.

Opsummerende kan det siges, at Harters teori om relevans som en kontekstuel ændring i en brugers kognitive struktur karakteriserer et gennemført dynamisk relevansbegreb, som er i evig forandring, som er af stærkt subjektiv og kun subjektiv natur, og som tager højde for næsten enhver tænkelig brugerbestemt faktor, som kunne tænkes at influere på relevansbedømmelsen.

Mulighederne for at operationalisere psykologisk relevans bremses af en række måske uoverstigelige hindringer, som vi senere vil diskutere i vores afsnit om metodologi.

Harters synspunkt støttes af T.K. Park (1992,-1994), der lig Harter understreger kompleksiteten og dynamikken i den enkelte brugers relevansbedømmelse. Park opstiller en model (se afsnittet om metodologi), som viser hvilke brugerkriteriekategorier og hvilke kontekster, der indvirker på den enkelte brugers relevansfortolkning. Park går langt videre end den traditionelle opfattelse af relevans, som vi finder i "The Systems View":

Hos Park er relevans et flydende begreb, som kun kan observeres på et bestemt tidspunkt, for en bestemt bruger, i en bestemt kontekst. Den enkelte brugers relevansbedømmelse bliver dermed i høj grad et spørgsmål om fortolkning på flere niveauer.

Park repræsenterer således, i lighed med Harter (1992) en kognitiv og brugerorienteret synsvinkel, som gør opmærksom på vigtigheden af dynamiske og psykologiske faktorer.

Swanson (1988, s. 395) kommer ind på det problematiske ved at formalisere de situationelle parametre i forbindelse med en brugers informationssøgning, fordi der er en betragtelig afstand mellem opfattelser på et operationelt henholdsvis teoretisk niveau. Swanson argumenterer for at skelne mellem subjektiv og objektiv relevansforståelse. Han bestrider, at forholdet mellem en forespørgsel og et dokument nødvendigvis er en subjektiv afvejning, og at spørgeren er den eneste, som kan relevansbedømme et dokument. Vi citerer:

"But 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, s. 391)."

Objektiv relevans hænger ifølge Swanson sammen med Poppers begreb "objective knowledge", der hører hjemme i dennes "verden 3". Swanson mener, at der *objektivt* set er en relation mellem en forespørgsel og et dokument, i en situation hvor dokumentet er et direkte svar på et, i forespørgslen explicit formuleret, spørgsmål.

Swanson understreger, at eksistensen af objektiv relevans hverken er indlysende, målbar eller bevislig. Vi mener derfor, at Swansons advokaten for objektiv relevans er af spekulativ karakter, med tanke på teoriens afsæt i Poppers kritiske rationalisme.

I en reviewartikel fra 1990 konkluderer Schamber, Eisenberg og Nilan, at relevans har udviklet sig til et multidimensionalt, kognitivt, dynamisk og komplekst begreb (Schamber et al., 1990, s. 774). De er dog klar over, at de ikke er nået til en endelig konklusion omkring relevans og nævner selv flere forbehold, hvad angår forskningsmæssig sammenhæng. *Et* forbehold tager de i forhold til, hvorledes eksisterende informationssystemers funktioner og faciliteter, som influerer på den enkeltes informationsadfærd, risikerer at tvinge forskningen ind på en for snæver vej. Udover dette forbehold opstiller de en liste af forskningsspørgsmål, som det efter deres mening vil være betimeligt at undersøge fremover (Schamber et al., 1990, s. 773). Vi skal ikke her gengive listen i sin helhed, men den stiller hovedsageligt spørgsmål, der går på at afklare hvilke kriterier brugere anvender ved relevansbedømmelse, og hvorfor, samt hvilke karakteristika ved dokumentrepræsentationen, som er inkluderet i disse kriterier, og om der er en prioritering blandt disse. Denne forskningsdagsorden er udgangspunktet for Barry (1994) og Park (1993). Vi kommer ind på disse i

denne artikels afsnit om den nyeste forskning og forskningsmetodologi.

2.4 Den hermeneutiske tilgang

Thomas Fröhlich tilføjer en mere videnskabssteoretisk/filosofisk vinkel til det stadigt voksende kompleks af tilgange til forståelsen af relevans. Den prototypiske kerne i en brugers relevansbedømmelse er ifølge Fröhlich emnet, men udenom kernen træder en række andre relevanskriterier i funktion. Fröhlich (1994) mener, at man bør tage udgangspunkt i hermeneutikken for at forstå, hvorledes kontekstuelle faktorer former relevansbedømmelsen og skaber forskellige fortolkninger. Fröhlich støtter sig til forståelse af hermeneutik som den fremlægges af Gadamer, og Fröhlich anbefaler, at man anser relevansbedømmelse som en interpretativ handling:

"One of the unstated assumptions about the nature of a relevance judgment is that it is an act of interpretation, an act in which the end-user assesses the document or document surrogates in terms of his need in relation to his task, his background, his subject matter, the context, his goals, the norms of his discipline, etc. (Fröhlich, 1994, s. 130)".

Ifølge Fröhlich er slutbrugerens interaktion med IR-systemer og -processer sammensat af flere fortolkningsprocesser. For det første den proces, hvor brugeren fortolker sit eget informationsbehov over for sig selv, intermediæren eller IR-systemet. En proces, der er præget af den "horisont", som brugeren bringer med sig til informationssystemet, og som påvirker hans relevansvurdering. Fröhlich foreslår, at man i denne sammenhæng fokuserer på prototypiske brugere indenfor det enkelte domæne. For det andet sker ydermere en fortolkning af dokumentet og emnet for at placere et dokument i en samlingen først og fremmest i accessions- og indekseringsøjemed. Og endelig den, systemet iboende "horisont", hvad angår selve systemets "arbejdsmetode", hvor der sker en programmeret fortolkning når brugerens forespørgsel, der kan være fritekst eller deskriptorer bruges af IR-sy-

stemet i genfindingsøjemed. Fröhlich (1994) har den holdning, at hvis man ser relevansbedømmelsen som en proces af fortolkende rammer, som udledt fra en gruppe af prototypiske brugere med en endelig række af relevanskriterier, vil man få en langt bedre forståelsesramme, når IR-systemer skal udvikles og designes.

2.5 Den domæneanalytiske tilgang

Hjørland tilslutter sig en kontekstafhængig tilgang, men i modsætning til den del af IR, der tilslutter sig den brugerorienterede synsvinkel, fokuserer Hjørland (1992) på den retning, som af Saracevic karakteriseres som "The Subject Knowledge View", hvor vægten lægges på den faglige videns natur, struktur og omfang. Hjørland placerer denne teori i en domæneanalytisk ramme:

"Selvom det altid er den enkelte brugers subjektive opfattelse, der gør sig gældende ved relevansvurdering, må idealet være en så objektiv definition som muligt. I fagenes grundlag, i filosofien og videnskabssteorien beskæftiger man sig på det overordnede plan med de spørgsmål, der har betydning for givne informationers forskningsmæssige relevans. Det er den højeste grad af objektivitet, man kan nå frem til (Hjørland, 1992, s. 160)".

Hjørland anerkender, at den kognitive vinkel på informationsvidenskaben kan være nødvendig, men pointerer, at det teoretiske udgangspunkt eller fokus er væsentlig anderledes og at den kognitive synsvinkel ikke anerkender vigtigheden af "the knowledge-producing communities" og deres betydning for den menneskelige erkendelse (Hjørland & Albrechtsen, 1994). Hjørland er altså af den opfattelse, at forudsætningen for at træffe bevidste valg med størst mulig grad af objektivitet afhænger af den enkeltes indsigt i fagets grundlagsproblemer. Vi konkluderer herud fra, at Hjørland er af den opfattelse, at der indenfor et givet domæne eksisterer litteratur, som, groft sagt, må være objektiv relevant. Det skal understreges, at der ikke er tale om, at noget er absolut objektivt

relevant, men om en høj grad af tilnærmet objektivitet

I henhold til "Dark Matter"-problematikken, som vi møder den hos Ingwersen, mener vi, at hvis man som Hjørland antager en tilnærmelsesvis objektivt relevant, afgrænset vidensmasse indenfor domænet, må "The Dark Matter" være reduceret til et minimum. Ingwersen (1991) kritiserer Hjørland for at antage, som Ingwersen kalder det, en gudlignende supervisor med totalt overblik over domænets kollektive erkendelse; For en gud kan der ikke eksistere Dark matter.

Et centralt problem ved Hjørlands meget domænefikserede opfattelse er, at det er svært at se, hvorledes man kan generalisere fra videnskabelige domæner til ikke-videnskabelige sammenhænge

Tendenserne i de senere års relevansdebat opsummeres af Schamber (1994) i de følgende punkter. Relevans bliver betragtet som værende:

"Subjective, depending on human (user or nonuser) judgment and thus not an inherent characteristic of information or a document;

Cognitive, relating to the individual users' information problems;

Multidimensional, influenced by many factors;

Dynamic, constantly changing over time; and

Measureable, observable at a single point in time

Schamber, 1994, s.6

2.6 Diskussion

Vi er et langt stykke ad vejen enige i ovenstående opsummering, men det er vores opfattelse, at der er enkelte yderligere hensyn, som må tages, for at

vi kan komme videre rent teoretisk. Vi har et forbehold overfor, hvorvidt en måling af relevansbedømmelse for en bestemt bruger, på et bestemt tidspunkt er et tilstrækkeligt udgangspunkt for en generalisering, som i praksis kan anvendes til design og evaluering af informationssystemer. Vi er af den opfattelse, at det vil være umuligt at tage højde for det utal af brugerkræfter den slags målinger vil producere, når man skal designe og evaluere informationssystemer, samt udforme katalogiserings- og indekseringspraksis. Schamber (1994) opregner på basis af tidligere litteratur om relevans 80 brugeridentificerede relevanskræfter, og der er på det seneste kommet flere til, eksempelvis, som tidligere nævnt, de kræfter vi finder hos Barry (1994) og Park (1994). Både Schamber, Barry og Park giver udtryk for problemer med at generalisere fra den enkelte bruger. De foreslår, at vejen frem er en syntese af resultaterne af hidtidige og fremtidige undersøgelser, herunder især undersøgelser, der sigter på at rangordne relevanskræfter. Schamber, Barry og Park støtter en multi-dimensional tilgang, som inddrager kontekstuelle variable, men som i sidste ende fokuserer kraftigt på den enkelte bruger.

Vi mener ikke, at problemet ved den store mængde af ikke-generaliserbare relevanskræfter kan løses udelukkende ved en rangordning. Det er nødvendigt, ikke alene at rette blikket mod den enkelte bruger, men også mod synsvinkler, der kan skabe et bredere fundament for relevansbegrebet. Et relevansbegreb, som i højere grad lader sig operationalisere. Her tænker vi på relevanskræfter, der går på tværs af brugergrupper, og som er afhængige af underliggende teknologiske/systemafhængige og fagvidenskabelige faktorer.

Vi mener, at en teknologisk/systemafhængig vinkel er nødvendig, idet brugerens muligheder for relevansbedømmelse er afhængige af de teknologiske rammer, som informationssystemet sætter. Når informationssøgesystemer og dokumentrepræsentationer er betingede af de forhåndenværende tekniske muligheder, er det naturligt at nye tekniske faciliteter kan skabe nye former for relevanskræfter.

er, som ikke før har været tænkelige. Derfor er en fuldstændig styring af design og udvikling af informationssystemer baseret på relevanskriterier, som er udledt gennem brugerundersøgelser i forhold til eksisterende systemer ikke mulig. Den nye teknik skaber kontinuerligt nye måder at interagere med systemet på, og den teknologiske udvikling vil dermed altid være en dark horse med hensyn til at kortlægge relevanskriterier. Det er vores holdning, at det i realiteten ikke er muligt at identificere den endelige række af relevanskriterier, som flere forskere indenfor feltet lægger op til. Omend vi mener denne vinkel er en dark horse, er det efter vores opfattelse vigtigt at holde sig den for øje, fordi den understreger, at nok har den enkelte bruger subjektive forudsætninger for relevansbedømmelse, men han er også "opdraget" af eksisterende informationssystemer, der har begrænset og dermed i længden formet brugerens opfattelse af, hvad der er relevant.

De ovenfor skitserede problemer taget i betragtning, mener vi, det er nødvendigt at nå frem til en indgangsvinkel, som inddrager overordnede faktorer, der går på tværs af brugere og systemer. Et redskab hertil kan være en fagvidenskabelig/domænespecifik vinkel, der inddrager en langt større kontekst, og især er vigtig for informationssystemer i akademiske og videnskabelige miljøer.

Vi har tidligere rejst en kritik af Hjørlands (1992) domæneanalytiske tilgang, som fortrinsvis går på hans grundliggende antagelse om domænets kollektive erkendelsesmasse, som basis for en tilnærmet objektiv relevansbedømmelse. Det betyder ikke, at vi er blinde overfor den sammenhæng, som brugerens relevansbedømmelse har med den fagvidenskabelige kontekst. I enhver litteratursøgningsproces, strækkende sig over kortere eller længere tid, vil terminologien indenfor domænets fremherskende paradigmer få gradvis større indflydelse på litteratursøgningen, og således påvirke søgningen og relevansbedømmelsen i domænespecifik retning. Endvidere opfatter vi eksempelvis citationspraksis og de herfor underliggende motiver som værende et anliggende for en domænespecifik analyse, og vi

postulerer, at citationspraksis er af signifikant betydning for brugerens relevansbedømmelse. Det sæt af skrevne, og ikke mindst uskrevne, regler, som gør sig gældende, når der citeres, indvirker på relevansbedømmelsen af mindst to årsager. Den første er, som Harter (1992) kort berører, at citationer i dokumenter, influerer på relevansbedømmelse, idet de er et udtryk for en, dokumentet iboende, historisk relevans.

"The act of citing is the statement of an historical relevance relationship, captured for all time in a published article (Harter, 1992, s. 614)".

Der er tale om en faktor, som påvirker læseren, når han optræder i rollen som bruger, og i litteratursøgningsprocessen skal vurdere, hvad han anser for at være relevant.

Endnu en årsag, til at citationspraksis er relevant i forhold til relevansbedømmelse, er at den veksler fra domæne til domæne og dermed er stærkt kontekstafhængig. Den enkelte brugers domænetilhørsforhold vil således, mener vi, have afgørende indflydelse på dennes relevansbedømmelsesnormer.

Vores postulat er, at en liste over relevansegen-skaber, som udover de af Schamber (1994) nævnte (se ovenfor) også inddrager de her skitserede, vil være et brugbart redskab i en relevansdiskussion, som ønsker at bevæge sig videre end det nuværende stade. Vi mener ikke, at vi med vores tilføjelser har gjort listen endelig, men at vi har gjort opmærksom på en række af overvejelser, som set i sammenhæng kan danne fundament for et frugtbart arbejde.

3. Forskning og metodologi

3.1 De første undersøgelser

Ifølge Ellis (1990) kan IR-forskning spores tilbage til 1953, hvor der i både England og USA blev udført tests med det formål at måle effektiviteten af Uniterm-indeksering i forhold til konventionelle indekseringssystemer.

Ifølge Ellis (1990) var testresultaterne ikke entydige, men de problemer, som manifesterede sig under processen, fik stor betydning for senere eksperimenter.

Kort sagt viste problemerne i det amerikanske forsøg sig ved uenighed om relevansbedømmelse. To grupper; en fra ASTIA og en fra Uniterm indekserede hver for sig ASTIAs beholdning på 15000 dokumenter. Systemernes effektivitet skulle så måles ud fra genfindingsgraden af relevante dokumenter i forhold til givne spørgsmål. Dette var, ifølge Ellis (1990) første gang relevans dukker op som målestok i systemtests. Testen brød sammen p.g.a. de to gruppers manglende enighed om, hvilke dokumenter, der var relevante for hvilket spørgsmål. Problemet viste sig at være det vigtigste udkomme af testen, idet det henledte forskernes opmærksomhed på essentielle problemer ved at anvende relevansbedømmelse som en objektiv størrelse.

Formålet med den engelske test var af samme karakter, nemlig at vurdere to indekseringssystemers effektivitet i forhold til hinanden. Metoden var dog en anden; for at omgå problemet med at afgøre hvilke dokumenter, som var relevante, udvalgte man en testpulje på 200 dokumenter. Ud fra testpuljen konstruerede man 40 forespørgsler, som blev brugt som søgegrundlag. Effektiviteten skulle så måles på genfindingsgraden af netop det specifikke dokument som den enkelte forespørgsel var konstrueret ud fra. Ellis gør opmærksom på to iøjnefaldende fejlkilder:

"(a) Only a single figure performance value was obtained (based on retrieval of the source document) ; no corresponding figures were obtained for the retrieval of non-source documents, whether those might be considered 'relevant' or 'irrelevant'

(b) A term-based system might be said to be favored over a conceptbased if the derivation of the artificial requests was influenced by the

terms used in the source documents (Ellis, 1990, s. 3)".

Eksperimenterne blev fulgt af en lang række nye tests. De mest kendte er Cranfield I og II. Cranfield I forsøgte at tage højde for en del eventuel kritik ved at lade en række informationsspecialister i forskellige organisationer udvælge, hvilke dokumenter de fandt relevante i forhold til 100 af 1200 konstruerede forespørgsler. De 41 spørgsmål, hvortil der var foreslået flest relevante dokumenter, blev så brugt som søgeforespørgsler i testen, der blev foretaget i en bestand på 18000 dokumenter - indekseret ved brug af fire forskellige indekseringssprog. Et meget vigtigt resultat af Cranfield I var, at man her for første gang blev opmærksom på den, lige siden bredt accepterede, regel om, at der eksisterer et omvendt proportionalt forhold mellem recall and precision.

Størstedelen af kritikken mod Cranfield I gik, som i de tidlige eksperimenter, på brugen af kildedokumenter, altså specifikke dokumenter fra bestanden som grundlag for konstruktionen af forespørgsler.

Cranfield II var metodisk en udbygning af Cranfield I, men der var stadig problemer med enighed om relevansbedømmelse. Samtidig var det stadig kritisabelt, at testen var foretaget på laboratoriebasis, og derfor nødvendigvis var meget konstrueret i forhold til en virkelig søgesituation. Cleverdon, som havde forestået Cranfield-forsøgene ofrede resultaterne og kritikken en del overvejelse, som uundgåeligt ledte ham til at konkludere, at der måtte være andre faktorer end det udtrykte forhold mellem forespørgsel og dokumentrepræsentation, som spillede ind på relevansbedømmelsen, og at denne således vil ændre sig afhængigt af metode og kontekst. Ellis citerer Cleverdon for en, i forhold til herværende artikel, ret essentiel konklusion:

"...the matter of relevance assessment is, without doubt, the most difficult intellectual problem - in fact one of the very few remaining problems - in the evaluation of information

retrieval systems (Cleverdon, 1966 I: Ellis, 1990, s. 17)".

Som tidligere nævnt førte disse erkendelser til omfattende projekter med det formål at identificere faktorer, som har indflydelse på relevansbedømmelse. Cuadra & Katter og Rees & Schultz udførte de for datiden mest signifikante undersøgelser af relevanskriterier. Deres vigtigste konklusioner var, at relevansbedømmes individuelle karakteristika; deres behov, deres viden, deres holdninger og hele kognitive struktur/stade var af afgørende betydning. Endvidere påpegede Cuadra & Katter, at det efter deres mening er metodologisk umuligt at måle relevansforholdet mellem en brugers informationsbehov og et dokument, p.g.a. dette behovs dynamiske og uhåndgribelige karakter. Deres resultater ledte dem således henimod at foreslå dels, at den fremtidige forskning skulle koncentrere sig om udvikling af bruger- og situationsmodeller, og dels at adækvat relevansbedømmelse kun er mulig for en stor gruppe af relativt homogene relevansbedømmere. Schamber et al. (1990) understreger vigtigheden af Cuadra & Katter og Rees & Schultz's undersøgelser. Undersøgelserne har sat dagsordenen for relevansforskningen helt op i vor tid. Schamber et al. tager dog et par forbehold, som vi finder meget væsentlige, og som vi er enige i:

"At this point we can say we agree with Rees and Schultz's and Cuadra and Katter's priorities for cognitive directions for future research. But we disagree with their contentions that it is methodologically impossible to measure the internal process of information evaluation, or to do so among heterogeneous groups of individuals. We also feel that in order to arrive at a full understanding of the meaning of relevance, researchers must focus on the perceptions of end-users in real information need situations, not on subject experts or search intermediaries, or on any judges in artificial test situations (Schamber et al., 1990, s. 763)".

3.2 Nyere empiri

Den nyere relevansforskning har som sit udgangspunkt konklusionerne fra de banebrydende forsøg i 1960'erne. Konklusionerne medførte et skift fra de traditionelle laboratorieforsøg mod mere kvalitative undersøgelsesmetoder. Den rationalistiske tankegang, der prægede "The Systems View" er afløst af en erkendelse af, at relevans er langt mere kompleks end traditionelt antaget. De nyeste empiriske undersøgelser af relevans sigter mod at afdække de grundliggende relevanskriterier, som slutbrugere anvender i deres relevansbedømmelse. Det brugerbaserede relevansbegreb kræver andre undersøgelsesmetoder, der kan fange temaer og mønstre ved den enkelte slutbrugers relevansbedømmelse, når man støtter en vidtfavnende, multi-dimensional opfattelse af relevans.

Barry er et eksempel på en af de nyeste empiriske undersøgelser af relevansbedømmelse, efter at diskussionen om relevans og relevansbedømmelse stod stille gennem den sidste del af 70'erne og begyndelsen af 80'erne. Det grundlæggende spørgsmål i Barrys undersøgelse var: Hvilke kriterier indvirker på brugeres afgørelse af, hvorvidt der eksisterer en forbindelse mellem information i dokumenter og brugeres informationsbehovssituation.

Hendes valg af metodologi var begrundet i et ønske om at skabe et så realistisk undersøgelsesmiljø som muligt, for at give brugerne/undersøgelsespersonerne mulighed for at evaluere information i ægte aktuelle behovssituationer. Barry udvalgte sine forsøgspersoner ved at annoncere med gratis on-linesøgning til brugere, som var villige til at deltage i undersøgelsen. Hun endte med en gruppe bestående af 18 studerende og lærere fra forskellige faggrupper på et universitet. Forsøgspersonernes formål med informationsøgningen spændte fra opgaveforberedelse for de yngste årgange af universitetsstuderende til doktorafhandlinger og videnskabelige artikler. Der blev først foretaget en foreløbig søgning for hver af forsøgspersonerne baseret på en skriftligt formuleret forespørgsel. Søgningerne dannede basis for præ-søgeinterviews

med forsøgspersonerne, hvorefter den endelige onlinesøgning blev udført. 15 dokumenter blev tilfældigt udvalgt fra søgningen. Hvert af disse dokumenter blev præsenteret for forsøgspersonerne med en række dokumentrepræsentationskarakteristika. Endelig blev hver forsøgsperson også præsenteret for tre tilfældigt udvalgte (af de 15) dokumenter i fuldtekstformat. Gennemsyn af de enkelte poster foregik i tilfældig rækkefølge og forsøgspersonerne blev bedt om at afmærke de dokumentrepræsentationskarakteristika, som fik dem til at

reagere. Efter endt gennemsyn blev forsøgspersonerne bedt om, for hver afmærkning at uddybe, hvorfor de havde reageret. Undersøgelsen fortsatte med yderligere uddybning, indtil forsøgspersonerne ikke længere kunne angive nye kriterier for relevansbedømmelsen, eller indtil tiden løb ud (der var afsat max. to timer til interview af hver forsøgsperson). Barry identificerede 23 relevanskriterier, som hun efterfølgende samlede i syv overordnede grupper. En gengivelse af Barrys kategoriskema ses herunder.

Grouping	Criterion categories
Criteria pertaining to the information content of documents	Depth/scope Objective accuracy/validity Tangibility Effectiveness Clarity Recency
Criteria pertaining to the user's previous experience and background	Background/experience Ability to understand Content novelty Source novelty Stimulus document novelty
Criteria pertaining to the users belief and preferences	Subjective accuracy/validity Affectiveness
Criteria pertaining to other information and sources within the information environment	Consensus External verification Availability within the environment Personal availability
Criteria pertaining to the sources of documents	Source quality Source reputation/visibility
Criteria pertaining to the document as a physical entity	Obtainability Cost
Criteria pertaining to the users situation	Time constraints Relationship with author

Barrys skema over de 23 relevanskategorier og deres inddeling i overordnede grupper (Barry, 1994, s.154)

Barry mener, at hendes undersøgelse støtter den antagelse, at relevansbedømmelsen påvirkes af andre faktorer end de rent emnemæssige, nemlig den række af situationelle faktorer, som ses i modellen ovenfor.

Vi mener, at et problem ved Barrys undersøgelse og konklusion er, at undersøgelsen rent metodisk er et statisk billede af hvilke relevanskriterier hendes svarpersoner angav. Problemet ligger, efter vores mening i, at en søgeproces er dynamisk, jvf. Harter (1992), der lægger stor vægt på, at der

løbende under søgningen sker kontekstuelle ændringer i brugerens kognitive struktur, som medfører en udvikling i brugerens informationsbehov, hvilket i praksis betyder, at Barry burde have interviewet sine studerende løbende under søgningen og stillet spørgsmål hver gang en relevansvurdering blev foretaget. Vores postulat er, at de samme relevanskriterier sandsynligvis ville optræde også under mere dynamiske undersøgelsesbetingelser, men at der skabes et mere reelt grundlag for at vurdere/rangordne relevanskriterierne. Endelig kan det ikke udelukkes, at nye kriterier kunne dukke op hos samme gruppe studerende, ligesom vi ikke kan udelukke det eksakt samme resultat, som oprindeligt blev nået.

Barry understreger, at den endelige række af relevanskriterier, hun mener at have identificeret kun kan siges at gælde for netop den gruppe i den situation og, at man ikke ud fra så forholdsvis spinkelt et grundlag kan generalisere til andre brugergrupper i andre situationer. Vi mener, at Barry har ret, når hun siger:

"It would not be possible to examine every conceivable type of user in every situation, examining all types of information, within a single study (Barry, 1994, s. 157)".

Der, hvor Barry, efter vores opfattelse, arbejder på et problematisk grundlag i sin undersøgelse, er, når hun siger, at hendes forskning er baseret på den grundlæggende antagelse, at der eksisterer en endelig række af relevanskriterier på tværs af brugergrupper og situationer. Vi mener, at det er umuligt, at opnå sikkerhed om en endelig række af relevanskriterier. Ikke alene er det umuligt at afgøre, hvorvidt en sådan grænse er nået, vi mener også, at først og fremmest den teknologiske udvikling; i.e. udviklingen i søgefaciliteter, interfaces etc. betyder, at rækken af både identificerbare og ikke-identificerbare relevanskriterier vil ændre sig. Set i forhold til implementering af resultaterne fra en undersøgelse som Barrys, er diskussionen om en endelig række af relevanskriterier en forholdsvis teoretisk diskussion, som man ved systemdesign og

-evaluering er nødt til at sætte sig delvist udover. Vi er derfor helt på bølgelængde med Barry, når hun foreslår, at det i første omgang handler om at syntetisere resultaterne af lignende undersøgelser, for herigennem at nå et skridt videre i udviklingen af den fremtidige forskningsdagsorden.

3.2.2 Park (1993)

Park (1993) lavede en undersøgelse, hvis formål var øge forståelsen for, hvad relevans er fra brugerens perspektiv. Hun støtter det synspunkt, at relevans helt og holdent er afhængig af brugerens situation og psykologiske stade. Hun valgte en naturalistisk dataindsamlingsmetode, da hun mente, at denne metode er velegnet til at beskrive brugeres adfærd i relevansbedømmelsesprocessen.

"The data collection method for this research was based on the belief that the naturalistic inquiry method is appropriate to reflect the underlying meaning of relevance in users' context because of its' potentially complex and subjective nature (Park, 1993, s. 325-326)".

Deltagerne i undersøgelsen var brugere fra et akademisk miljø, både lærere og studerende, som havde fået foretaget en litteratursøgning. Deltagerne var frivillige og blev valgt målrettet og ved tilfældig udtrækning. Grunden var den, at Park mente, at relevans fra brugerens perspektiv varierer afhængig af brugergruppe og den enkeltes informationsproblem. Ti brugere blev udvalgt og herefter interviewet med udgangspunkt i en liste af åbne spørgsmål. Spørgsmålene gik på brugerens problemfelt, forventninger til søgeresultatet og brugerens adfærd relateret til udvælgelsen af dokumentrepræsentationer fra søgningen. Et print af søgningen blev brugt til at spore den enkeltes tankeproces i forhold til hver enkelt dokumentrepræsentation i søgningen.

Analysen af interviewene gik ud på at identificere gennemgående temaer i den enkeltes relevansfortolkning. Båndene blev tolket og disse tolkninger blev forelagt deltagerne i undersøgelsen for at checke, om de var enige i tolkningen af, hvorledes

de havde opfattet situationen. Herefter lod Park kolleger komme med kritik af, hvorvidt hun havde overset noget eller lagt for meget vægt på andet. Park konkluderer, at relevans overordnet er et forhold mellem en dokumentrepræsentation og en bruger. Fortolkningen af dokumentrepræsentation er det centrale problem og kan afhænge af flere faktorer.

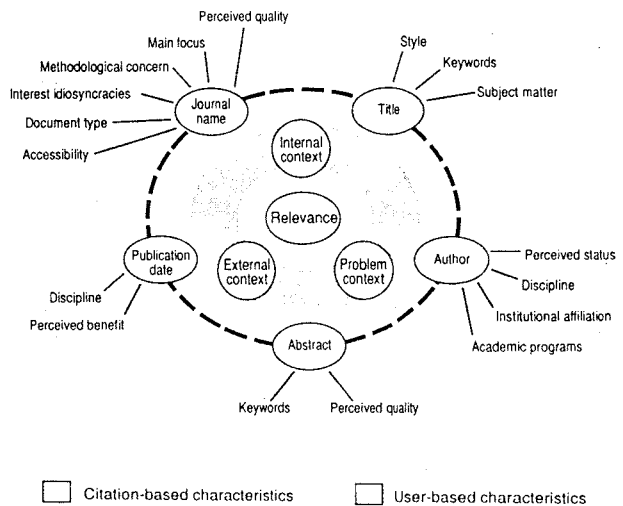
"A user's evaluation of a citation can be analysed in two ways, from the attributes of the citation itself and from the context in which the evaluation takes place. The data analysis represents those two dimensions, which I call citation-based and user-based (Park, 1993, s. 329)".

Resultatet af undersøgelsen var for Park (1993, 1994), at hun i analysefasen skelnede mellem to kategorier af brugerbaserede relevanskriterier. Den første kategori baserer sig på karakteristika ved dokumentrepræsentation og den anden baserer sig på karakteristika ved brugeren selv. De dokumentrepræsentative kriterier går på, hvorledes karakteristika ved disse påvirker brugerens relevansbedømmelsesproces.

Den første del af analysen omhandler brugerens fortolkning af aspekter eller komponenter ved den enkelte dokumentrepræsentation. Fortolkningen af dokumentrepræsentationen skal vise, hvorledes brugeren behandler de enkelte elementer eller sammenhængen mellem disse.

Den anden del af analysen ser på de grundlæggende faktorer, der påvirker valget af den enkelte dokumentrepræsentation. Faktorerne afhænger af brugerens erfaringer, perception og viden om informationsproblemet. Denne del af analysen inddeler kilderne til brugerens relevansbedømmelse i tre kontekster. Den interne kontekst, der rummer brugerens fortolkninger baseret på hans/hendes erfaring og uddannelse. Endvidere den eksterne kontekst, nemlig den situation hvori søgningen finder sted, samt en kontekst, der viser, hvordan

og hvorfor brugeren anvender information til at konstruere og løse sit informationsproblem.



Parks oversigt over kontekstuelle variable og bruger-relevanskriterier (Park, 1994, s.342)

Park konkluderer på baggrund af sin undersøgelse, at forholdet mellem en bruger og et dokument ikke er simpelt, men afhænger af en lang række forskellige kontekstuelle faktorer. Hun mener, at de empiriske resultater af hendes undersøgelse viser, at man ikke kan måle relevans præcist. Hun erkender også, at resultaterne er begrænset til undersøgelsens specielle kontekst og er her helt på linje med Barry (1994).

Problemet i hendes undersøgelse er, som vi ser det, at hun først måler/afdækker brugernes relevansbedømmelse efterfølgende og ikke undervejs i relevansbedømmelsesprocessen. En kritik, som er i tråd med den, vi tidligere i artiklen har fremført overfor Barrys undersøgelse, men Park tager i større grad højde for det dynamiske, situationelle, i og med, at hun interviewer sine brugere løbende, mens de relevansbedømmer resultatet af den søgning, de har fået foretaget. Park vurderer i et forbehold overfor sin egen undersøgelse, at det vil være muligt at følge en brugers relevansbedømmelse på forskellige stadier i hans/hendes informa-

tionssøgning og følge de skiftende mønstre i brugerens opfattelse af relevans.

Operationelt set er det et problem, at en sådan undersøgelse vil producere store mængder af kvalitative data selv for en enkelt bruger, og dermed vil det med mange deltagere i en undersøgelse blive et problem at mangle overskuelighed i undersøgelsesmateriale. Vi er derfor af den opfattelse, at man må skelne mellem undersøgelser, der går på at afdække relevanskriterier, og undersøgelser, som skal teste specifikke informations-systemers effektivitet. Parks undersøgelsesdesign er velegnet til at evaluere et informationssystem med henblik på hvilke krav, der skal stilles til praksis m.h.t. indeksering, katalogisering, interfacedesign, vis-formater samt brugermodelleringer.

3.3 Forslag til undersøgelsesdesign

Vi er af den opfattelse, at den måske sværeste opgave indenfor operationalisering ligger i konstruktionen af et brugbart undersøgelsesdesign til brug ved testning af informationssystemer. Det er nødvendigt at foretage en skarp skelnen mellem undersøgelser, som sigter imod en egentlig effektivitetstest og undersøgelser, som eksempelvis har til formål at identificere relevanskriterier, idet testsituationen for at kunne sige noget om et systems effektivitet, og deraf følgende designudvikling stiller krav om en høj grad af tilnærmet objektivitet i relevansbedømmelsen. Swanson siger herom:

"The testing of an information system, insofar as it can influence the design of the system, must operate with objectified requests and objective relevance of the responding documents (Swanson, 1986, s. 396)".

Vi vil forsøge at skitsere et forslag til et undersøgelsesdesign, som kan illustrere nogle af problemerne, samt forhåbentlig give nogle brugbare forslag til løsning af disse.

Vi mener, at man i forhold til den gældende forskning bør lægge vægt på især to hensyn. Det første er, at man i højere grad må forsøge at lave

trindelte analyser af relevansbedømmelsen, hvor man undersøger brugerens relevansbedømmelse på forskellige stadier i en litteratursøgningsproces. Man skal på hvert enkelt stadie sørge for at registrere relevansbedømmelsen dynamisk, imens brugeren interagerer med systemet. Her er en mulighed for at anvende en blanding af højtænkningens metoden, interview samt observation via videooptagelser af et litteratursøgningsforløb. Dette skal følges op af et efterfølgende interview, hvor brugeren kan redegøre for sine motiver for den konkrete relevansbedømmelse, både i forhold til en udskrift af søgningen og en gennemgang af videooptagelsen af søgningen. Et centralt problem er, hvordan man vælger at markere poster relevante. Hvorledes dette valg foretages afhænger af, hvorvidt undersøgelsen sigter mod at teste det pågældende informationssystems effektivitet, eller sigter mod at afdække forskellige relevanskriterier. Vil man teste systemeffektivitet, må man acceptere en reduktion af problemstillingen, idet man må konstruere en inddeling, hvorefter brugeren kan markere poster relevante. Uanset om man vælger at lade brugerne markere posterne som relevante i modsætning til ikke-relevante, eller lader brugerne markere relevansbedømmelse ved hjælp af en skala, har man i forhold til en teoretisk definition af relevans det problem, at det er svært at anvende brugerens relevansbedømmelse som grundlag for generalisering.

En metode til delvist at overvinde dette problem kunne være at skabe en testpulje af tilnærmet objektivt relevante dokumenter. Ved at lave en pulje af de dokumenter, som på tværs af brugerne i undersøgelsen indeholder eksempelvis fem relevanskriterier eller mere. Puljen, vedtager man, udgør tilnærmet relevante dokumenter, som man herefter overfører til et nyt informationssystem sammen med alle dokumenterne fra den oprindelige søgning. Både de, som er markeret relevante og de, som blev siet fra. Dette frembyder flere problemer, da vi ønsker en realistisk situation og ikke en situation med instruerede brugere og konstruerede forespørgsler. Det ideelle ville være at følge en litteratursøgning, der sker ud fra en skriftligt stillet

opgave på en uddannelsesinstitution, og som kunne gentages senere og stadig være realistisk i forhold til brugerne. Teknisk set skal man herefter overføre alle posterne fra den første søgning til det nye informationssystem og ikke medtage nye dokumenter. Nu vil man have en pulje, som man ved er tilnærmet relevant efter nogle faste relevansparametre, og som man kan måle genfindingsgraden på. Øges genfindingsgraden i gennemsnit for den enkelte bruger i forhold til testpuljen af tilnærmet relevante dokumenter har man et fingerpeg om, at det nye system er bedre til at hjælpe brugeren med at relevansbedømme. En vigtig forudsætning for undersøgelsen er, at brugergruppen såvidt muligt matcher den første brugergruppe. Universiteter eller andre uddannelsesinstitutioner har den fordel i undersøgelsesøjemed, at man som regel kender den samlede population og derfor kan lave et udtræk af studerende, der nøje matcher den oprindelige brugergruppe. Det er vores teori, at der herved skabes en situation, hvor en sammenligning af genfindingsgraden mellem de to brugergrupper giver et mål, som er et adækvat udtryk for de to informationssystemers relative effektivitet. En anden vigtig forudsætning er, at brugergruppen er relativt stor og homogen, hvilket i øvrigt også var en af Cuadra & Katters konklusioner på deres undersøgelse.

Det er vigtigt at understrege, at undersøgelsens resultater er begrænset til den kontekst, hvori den finder sted, eksempelvis uddannelsesinstitutionen, det faglige domæne etc. Undersøgelsen kan yderligere understøttes, hvis man foretager en analyse af motiver for citationspraksis, som vi tidligere så det efterlyst af Harter (1992). Motiverne for citationspraksis må afdækkes gennem en række kvalitative interviews med studerende og forskere, og analysen skal tage hensyn til, at citationspraksis både er afhængig af domænet og af brugerens placering i det akademiske hierarki. Vi ser den domæneanalytiske indgang som vigtig i den forstand, at den med sit overordnede udgangspunkt i domænets faglige vidensstruktur kan angive nogle generelle relevanskriterier, der ikke er begrænset til specifikke kontekster og enkelte brugergrupper. Et samspil

mellem en dynamisk, situationel undersøgelse og en domænespecifik, vil efter vores mening være et bud på, hvordan man kan komme videre i forhold til at udforme undersøgelser af relevans.

4. Konklusion

Vi har i denne artikel argumenteret for, at man på det teoretiske plan må tage udgangspunkt i et multidimensionalt relevansbegreb, som tager hensyn til en lang række situationelle og dynamiske faktorer, for at kunne bevæge sig mod en brugbar relevansdefinition. Desuden mener vi, at man må indtænke domæneanalytiske og systemafhængige tilgange, hvis man vil indkredse relevans.

Vi har fremhævet nødvendigheden af at skelne mellem undersøgelser, der sigter mod at afdække relevanskriterier og undersøgelser, som retter sig mod at teste IR-systemers effektivitet. De dynamiske og situationelle faktorerers vigtighed leder os frem til at konkludere, at det er nødvendigt at operere med et tilnærmet objektivt relevansbegreb i effektivitetsmålingsøjemed. I den forbindelse har vi kort forsøgt at skitsere et oplæg, som vil være et brugbart udgangspunkt for en sådan undersøgelse.

Litteraturliste

Barry, C.L. (1994). User-Defined Relevance Criteria: An exploratory Study. *Journal of the American Society for Information Science*, Vol. 45, nr. 3, s. 149-159.

Cuadra, C.A. & Katter R.V. (1967). Opening the Black Box of relevance. *Journal of Documentation*, Vol. 23, nr. 4, s. 291-303.

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

Fröhlich, T.J. (1994). Relevance Reconsidered-Towards an Agenda for the 21st Century: Introduction to Special Topic Issue on Relevance Research. *Journal of the American Society for Information Science*, Vol. 45, nr. 3, s. 124-134.

Harter, S.P. (1992). Psychological Relevance and Information Science. *Journal of the American Society for information Science*, Vol. 43, nr. 9, s. 602-615.

Hjørland, B. (1992). Informationsvidenskabens Grundbegreber. Khb., Danmarks Biblioteksskole.

Hjørland, B. (1993). Emnerepræsentation og Informationssøgning: Bidrag til en teori på et kundskabsteoretisk grundlag. Göteborg, Bibliotekshögskolan och centrum for biblioteks- och informationsvetenskap.

Hjørland, B. & Albrechtsen, H. (1994). Toward a New Horizon in information Science: Domain-analysis. Upubliceret artikel.

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

Park, T.K. (1993). The Nature of Relevance in Information Retrieval: An Empirical Study. *The Library Quarterly*, Vol. 63, nr. 3, s. 318-351.

Park, T.K. (1994). Toward a Theory of User-Based Relevance: A Call for a New Paradigm of Inquiry. *Journal of the American Society for Information Science*, Vol. 45, nr. 3, s. 135-141.

Schamber, L. 'et al' (1990). A Re-Examination of Relevance: Toward a Dynamic, Situational Definition. *Information Processing and Management*. Vol. 26, nr. 6, s. 755-776

Schamber, L. (1994). Relevance and Information Behavior. *Annual Review of Information Science and technology*, Vol. 29, s. 3-48.

Swanson, D.R. (1986). Subjective versus Objective Relevance in Bibliographic Retrieval Systems. *The Library Quarterly*, Vol. 56, nr. 4, s. 389-398.